Genome-phenotype-environment associations identify signatures of selection in a panmictic population of threespine stickleback

Kasha Strickland[1], Katja Räsänen[2], Bjarni Kristjánsson[3], Joseph Phillips[4], Arni Einarsson[5], Ragna Snorradóttir[6], Mireia Bartrons[7], and Zophonías Oddur Jónsson[8]

[1]The University of Edinburgh
[2]Eawag
[3]Hólar University College
[4]Creighton University
[5]Myvatn research Station
[6]Holar University College
[7]UVic-UCC
[8]University of Iceland

July 15, 2022

Abstract

Adaptive genetic divergence occurs when selection imposed by the environment causes the genomic component of the phenotype to differentiate. However, genomic signatures of natural selection are usually identified without information on which trait is responding to selection by which selective agent(s). Here, we integrate whole-genome-sequencing with phenomics and measures of putative selective agents to assess the extent of adaptive divergence in threespine stickleback occupying the highly heterogeneous lake Mývatn, NE Iceland. We find negligible genome wide divergence, yet multiple traits (body size, gill raker structure and defense traits) were divergent along known ecological gradients (temperature, predatory bird densities and water depth). SNP based heritability of all measured traits was high (h2 = 0.42 – 0.65), indicating adaptive potential for all traits. Whilst environment-association analyses identified thousands of loci putatively involved in selection, related to genes linked to neuron development and protein phosphorylation, only allelic variation linked to pelvic spine length was concurrently linked to environmental variation (water depth) - supporting the conclusion that divergence in pelvic spine length occurred in face of gene flow. Our results suggest that whilst there is substantial genetic variation in the traits measured, phenotypic divergence of Mývatn stickleback is mostly weakly associated with environmental gradients, potentially as a result of substantial gene flow. Our study illustrates the value of integrative studies that combine genomic assays of multivariate trait variation with landscape genomics.

Genome-phenotype-environment associations identify signatures of selection in a panmictic population of threespine stickleback

Strickland, K.,[1,2*], Räsänen, K.[3,4], Kristjánsson, B.K.[2], Phillips, J.S.[2,5], Einarsson, A.[6], Snorradóttir, R.G.[2], Bartrons, M.[7], Jónsson, Z.O.[8]

1. Institute of Ecology and Evolution, School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom
2. Department of Aquaculture and Fish Biology, Hólar University, 551 Sauárkrókur, Iceland
3. Department of Aquatic Ecology, EAWAG and Institute of Integrative Biology, ETH Zurich, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland

4. Department of Biological and Environmental Science, University of Jyväskylä, Finland
5. Department of Biology, Creighton University, Omaha, NE 68178, USA
6. Mývatn Research Station, 660 Mývatn, Iceland
7. Aquatic Ecology Group, University of Vic (UVic-UCC), 08500-Vic, Catalonia, Spain
8. Faculty of Life and Environmental Sciences, University of Iceland, 101 Reykjavík, Iceland *author to whom all correspondence should be addressed: kasha.strickland@ed.ac.uk**Running head:** adaptive divergence in Icelandic stickleback

Abstract

Adaptive genetic divergence occurs when selection imposed by the environment causes the genomic component of the phenotype to differentiate. However, genomic signatures of natural selection are usually identified without information on which trait is responding to selection by which selective agent(s). Here we integrate whole-genome-sequencing with phenomics and measures of putative selective agents to assess the extent of adaptive divergence in threespine stickleback occupying the highly heterogeneous lake Mývatn, NE Iceland. We find negligible genome wide divergence, yet multiple traits (body size, gill raker structure and defence traits) were divergent along known ecological gradients (temperature, predatory bird densities and water depth). SNP based heritability of all measured traits was high ($h^2 = 0.42 – 0.65$), indicating adaptive potential for all traits. Whilst environment-association analyses identified thousands of loci putatively involved in selection, related to genes linked to neuron development and protein phosphorylation, only allelic variation linked to pelvic spine length was concurrently linked to environmental variation (water depth) - supporting the conclusion that divergence in pelvic spine length occurred in face of gene flow. Our results suggest that whilst there is substantial genetic variation in the traits measured, phenotypic divergence of Mývatn stickleback is mostly weakly associated with environmental gradients, potentially as a result of substantial gene flow. Our study illustrates the value of integrative studies that combine genomic assays of multivariate trait variation with landscape genomics.

Keywords: adaptive divergence, gene flow, environmental gradients, genome scans, landscape genomics, *Gasterosteus aculeatus*

Introduction

Elucidating the genetic basis of adaptive divergence in natural populations is an enduring goal of evolutionary biology . Doing so can provide insight into evolutionary processes occurring in the wild, including the mechanisms associated with adaptive divergence, and the extent to which divergence takes place in the face of gene flow . Genetically, adaptive divergence is expected to manifest as blocks of differentiation across the genome, at regions containing genes that contribute to adaptation to divergent local environments . Genome scan studies that test these expectations have identified genomic regions associated with adaptation to divergent ecological niches in numerous species (e.g.,). This has been termed a “reverse ecology” approach, whereby loci associated with adaptation may be identified without measuring the traits themselves . However, genome scan studies on wild populations are seldom able to provide precise information on which aspects of the phenotype selection is acting on, or which environmental factors are imposing selection .

A comprehensive view on the genomic mechanisms associated with adaptive divergence requires studies that combine phenotypic, environmental and genomic data. Accordingly, integrative approaches that combine association mapping with landscape genomics or selection scans to map gene-phenotype-environment associations could be a powerful means to infer the genomic basis of adaptation . Association mapping studies (e.g., genome-wide-associations, GWA) identify specific loci that underlie divergent traits, whereas landscape genomic studies can aid in determining loci associated with adaptive divergence, under the assumption that loci should be correlated with environmental variation that is directly or indirectly causing selection . Combining association mapping with landscape genomics can strengthen the identification of genomic signatures of selection by allowing inference on whether causal variants of phenotypic variation are concurrently associated with environmental variation. This would be especially true in cases where correlations between

phenotype and environment are mirrored in genetic polymorphisms, where at some quantitative trait loci, allele frequencies differ between groups that inhabit different environments.

In the absence of dispersal barriers, many populations remain connected by gene flow during the process of adaptive divergence, often along environmental clines . Gene flow is expected to constrain divergence, swamping locally adapted alleles and breaking up favourable allele combinations through recombination . Whilst in cases of substantial gene flow there may be little genome-wide divergence, responses to natural selection may be present at specific genomic regions (islands of divergence;). Identifying genomic divergence in the presence of gene flow is a major challenge because most genome scan approaches require grouping individuals, which is not usually possible when individuals remain connected . Our perspective on adaptive divergence may therefore be biased towards studies where physical barriers to gene flow have facilitated divergence. Although such studies have provided great insight into evolutionary processes, studying processes of divergence in populations connected by gene flow can greatly improve our understanding of the relative roles of natural selection and gene flow in adaptive divergence .

Here, we employ GWA and landscape genomic approaches to map gene-phenotype-environment associations in threespine stickleback that inhabit Mývatn, a highly environmentally heterogeneous lake in NE Iceland. Threespine stickleback is a well-established model system in evolutionary biology . Within freshwater systems, there is evidence for repeated adaptive divergence at both phenotypic and genomic levels , most commonly across the benthic-limnetic axis (e.g.) but also across a range of other selective agents (e.g., predation). However, most of the studies focus on simple environmental contrasts (e.g., benthic vs limnetic or lake vs stream), and only few studies have aimed to test intralacustrine divergence across environmental gradients.

Mývatn is a large (37 km^2) lake, where temperature, water depth, invertebrate, and vertebrate (including stickleback) densities vary over space and time . Stickleback habitats in this lake can crudely be divided to five main types, across which stickleback vary phenotypically . Previous work found that male stickleback had relatively larger brains in a ´lava´ (warm) than a ´mud´ (colder) habitat , relatively longer spines in the north basin than the south basin , and divergence in gill raker morphology and diet among some of the habitats . Evidence for population genetic divergence of stickleback across the lake is mixed. Using samples collected between 1999 and 2002, found evidence for genetic divergence using a suite of nuclear and mitochondrial markers between stickleback inhabiting the ´lava´ and ´mud´ habitats (microsatellites: F_{ST} = 0.08; mtDNA: $F_{ST} = 0.223$), suggesting the presence of two contrasting morphs. In contrast, using samples collected in 2009 and 12 nuclear microsatellite loci (seven of which were the same as in Ólafsdóttir *et al.* 2007), found little evidence for neutral genetic divergence of stickleback across five habitat types (average pairwise $F_{ST} = 0.004$), suggesting extensive gene flow.

Given the known phenotypic divergence in traits typically under selection in stickleback, coupled with spatial variation in possible selective agents, our main goal here was to identify genomic signatures of selection in Mývatn stickleback occupying different environments. Genomic signatures of selection are typically defined as genomic regions which are disproportionately divergent between groups compared to the rest of the genome . We extended this definition to strengthen our identification of signatures of selection: we expected that genomic regions that bear a signature of selection should be both divergent across ecological axes, and contain loci associated with variation in divergent traits. We further measured SNP-based additive genetic variation of divergent traits to gain insight into the evolutionary potential of traits that are spatially divergent.

Methods

Study system and sampling

Lake Mývatn is composed of two basins (North and South basin) that are connected by two narrow channels and vary in a range of abiotic and biotic conditions . The lake is spring fed, with geothermal hot springs (up to ca. 23°C) feeding the north-east of the lake and cold-water springs (ca. 5°C) feeding the south-eastern parts. Most part of the lake follows the ambient temperature, which in summer is around 12-13°C The lake is shallow (1 – 4 m), but with some deeper areas (up to ca. 7 m) caused by historical diatomite mining in some parts of the North basin . Productivity, as well as benthic, epibenthic and pelagic invertebrate abundance

and community structure, also varies through space (Phillips*et al.* in review; Bartrons *et al.* 2015). Based on the combination of water temperature and depth, as well as vegetation and substrate, the habitats occupied by stickleback have previously been classified as warm, rocky shore, cladophorales, pondweed and mined). In addition, long-term monitoring data shows that stickleback population density varies in both space and time with the North basin having higher densities than the South, and with periodically strong dispersal from the North to the South basin (Phillips et al. in review).

The stickleback population of Mývatn has been surveyed each year since 1991 as part of an ongoing long-term monitoring of population demographics . This sampling is done during the third week of June and August each year by laying five unbaited minnow traps at pre-determined locations over two 12 hr periods (see Millet et al. 2013 for details). During monitoring, stickleback are counted to estimate catch-per-unit-effort (CPUE) and frozen for later analysis. For phenotyping and genotyping, a random subset of individuals (ca. N = 100 per site for each of the day and night catches) have been stored since 2009. To study patterns of spatial divergence, we used stickleback from 9 sites collected in June of 2012 due to the availability of detailed ecological data for this time point.

Ecological data

To characterize the environment that stickleback occupy, we focused on a set of ecological variables which represent putative selective agents. First, we used the same five habitat classifications (warm, mined, pondweed, cladophorales, and rocky shore) previously described in (see supplement for details). We also collated data on ecological variables likely to reflect selective agents. These were: water temperature, water depth, stickleback CPUE, piscivorous bird density, and zooplankton abundances and community composition. These were chosen because temperature can affect metabolic processes, development, tolerance to parasite infections , as well as key life history traits , whilst depth can affect sensory processes invertebrate availability, and stickleback visibility to predators . Stickleback CPUE was used as a measure of intraspecific competition , piscivorous bird density as a measure of predation pressure , and invertebrate data as a measure of prey abundance and composition . It should be noted that our measures are only proxies for selection imposed by correlated ecological factors.

Temperature and water depth of each site were used as per . Average temperature at each site was measured between 30 June 2011 and 18 August 2011 with a temperature logger (iButton Maxim Integrated Products, San Jose, CA, USA), placed at mid-depth and recording at three-hour intervals. CPUE for each site was estimated using count data from the long-term monitoring study from June 2012. To measure piscivorous bird density, we used data collected during the waterfowl census conducted each year at Mývatn, during which all waterfowl observed from pre-determined vantage points with known survey areas are counted . We used count data collected between 15th May and 10th June 2012 on the following species known to predate on stickleback: horned grebe (*Podiceps auritus*), red-breasted merganser (*Mergus serrator*), great northern diver (*Gavia immer*), red throated diver (*Gavia stellata*) and goosander (*Mergus merganser*). Note that the Arctic tern (*Sterna paradisaea)* is abundant at Mývatn, and predates on stickleback, but this species is not counted during the bird census. We calculated the density of piscivorous birds (summed across all taxa) in each surveyed segment of the lake (number/m2) (Figure S1).

We used invertebrate data from , which were collected by conducting surveys of the epibenthic and zooplanktonic community. Crustaceans (incl. *Daphnia* , copepods and epibenthic cladocerans) as well as rotifers are important food sources for stickleback in Mývatn (e.g.). Although chironomid larvae are a main food source for Mývatn stickleback, data on midge larval abundance were not of sufficient spatial resolution to be used (see). However, the benthic community is spatially correlated with the epibenthic and zooplankton community in the South basin , suggesting that measures of pelagic and epibenthic zooplankton may serve as a proxy measure for the benthic community in Mývatn. Briefly, three transects were conducted between June - July 2012, during which integrated vertical tows of the whole water column were made at each of 31 sites, spaced 500 – 600m apart (see Figure S2 for distribution, and for more details on sampling and sample processing). Each pooled sample of 15 L was filtered through 63-μm mesh and counted in entirety under a binocular microscope. We used data from the 2nd transect (25th July 2012) as the spatial resolution in

this transect was the greatest. We used data collected from the closest site to each stickleback sampling site (distance to closest zooplankton site: min = 290 m, max = 1365 m). All stickleback sites were within 2.55 km of the nearest site used to collect invertebrate data, which was the distance at which zooplankton communities were found to be spatially autocorrelated. We used number per litre (n/L) of each taxon at the sites closest to stickleback sites.

To summarise variance between sites for use in downstream analyses, we ran a principal components analysis (PCA) using the native stats package in R version 4.1.2 . This summarised the invertebrate data in four general axes, described in detail in Table S1. PC1 (zPC1) described the overall abundance of crustaceans and rotifers, and explained 29% of the variation; PC2 (zPC2) described the negative covariance of rotifer sp. and *Alona* sp. with planktonic and epibenthic crustaceans, which explained 17% of the variation; PC3 (zPC3) described the negative covariance of the rotifer *Keratella* and the cladocerans*Acroperus harpae* and *Chydorus sphaericus* with*Daphnia longispina* , and explained 15% of the variation; and PC4 (zPC4) described the negative covariance of the cladocerans*Eurycercus lamellatus* and *Macrothrix hirsuticornis* with*Daphnia longispina* , *Cyclops abyssorum* and*Asplancha* , explaining 11% of the variation in the data. Overall invertebrate abundance (described in zPC1), was highly negatively correlated with stickleback CPUE. We therefore used only CPUE and not zPC1 for downstream analyses.

Phenotypic data

We randomly selected individuals from each habitat type for phenotyping and genome sequencing from the frozen subset of the long-term monitoring samples. Individuals were thawed, weighed on an electronic balance (wet mass, nearest mg) and their total length (TL) measured using a ruler (to the nearest mm). The right pectoral fin was then cut and stored in 96% ethanol for DNA isolation. We measured traits typically under selection in stickleback: body size, defence traits (armour plate number and length of spines) and dietary traits (gill raker morphology and gut length) . Specifically, for each individual we measured the following 10 traits: total length (TL), total gut length (gut length), number of lateral armour plates (plate number), length of the first dorsal spine (DS1), length of the second dorsal spine (DS2), length of the pelvic spine (PS), length of the second gill raker on the first gill arch (GRL2), length of third gill raker on the first gill arch (GRL3), gap width between second and third gill rakers (GRW), and number of long gill rakers on the first gill arch (GRN) (see below). Note that we measured the second and third gill rakers, rather than the first (which is usually used in studies of stickleback trophic phenotype), because in some cases gill arches broke during dissection. After measurement of TL, each individual was dissected to remove the stomach and the gut, and any tapeworm (*Schistocephalus solidus)* parasites. Gut length (from the sphincter at the end of the oesophagus to the end of the digestive tract) was measured (to the nearest mm) using a ruler.

To aid morphological measurements, fish were stained with alizarine red using standard protocols (Millet et al. 2013). Fish were bleached using a 1:1 ratio of 3% H_2O_2 and 1% KOH and then stained in a solution of alizarin red and 1% KOH . After staining, digital images were taken of the left side of the fish with a Canon EOS 600D digital camera, with mm paper for scale. From these images, plate number was counted and the length of the spines (DS1, DS2 and PS) measured to the nearest hundredth of a millimetre. After imaging, we dissected the first gill arch and, where necessary, re-stained before mounting it between two glass plates and photographing using a digital camera (Nikon Coolpix 4500) mounted to a stereomicroscope (Leica MZ12) with mm paper for scale. We used the digital images of gill arches to measure GRL2, GRL3 and GRW (in mm) and counted GRN. All measurements were taken from the digital images were done using segmented tool in ImageJ .

Whole genome resequencing and SNP detection

Genomic DNA was isolated and purified from the ethanol stored fin clips using Macherey-Nagel nucleomag tissue kit, following the manufacturer's protocol. Paired end, PCR-free 150-bp insert libraries were then prepared for whole genome sequencing using the DNBSeqTM platform by BGI-Hongkong, generating an average of 40 million cleaned reads per individual (min = 33.4 million, max = 41.8 million, equating to an average depth of coverage of 10X). The clean paired-end reads were aligned to the threespine stickleback

genome assembly v5 with BOWTIE2 (version 2.4.1) using default parameter settings , and sorted and indexed using SAMTOOLS (version 1.10) . Variants were discovered using the short variant discovery pipeline of GATK . SNPs and indels were quality filtered according to GATKs best practices guidelines, using the following hard filters: QualByDepth > 2.0, FisherStrand bias < 60.0, RMSMappingQuality < 40.0, MappingQualityRankSumTest < -12.5, ReadPosRankSumTest < -8.0, StrandOddsRatio > 3.0, variant quality score < 30.0. For all following analyses, we removed mitochondrial variants, indels and multiallelic variants, as well as variants identified on either of the sex chromosomes . We then filtered the remaining autosomal SNPs for genotype depth less than six or greater than 100; minor allele counts less than four; missingness less than 20%. The sex of individuals was confirmed using the proportion of reads with depth greater than eight mapped to the X vs Y chromosome .

Statistical analyses

We conducted a series of analyses to test for i) phenotypic and genomic divergence in relation to geographic location and ecological variables, ii) genomic architecture of traits, and iii) associations of genomic variation with environment.

Phenotypic divergence

Phenotypic divergence was analysed using multivariate and univariate Bayesian linear mixed models. All traits were standardised to have a mean of zero and standard deviation of one to improve comparability and model convergence. All models described below were fit with a Gaussian error distribution in the MCMCglmm package in R . Fixed effects were given weakly informative flat priors, random effects were given default inverse Wishart priors, and we fit full unstructured covariance matrices for the multivariate models. Each model was run for a total of 1,020,000 iterations with a burn-in of 20,000 iterations and thinning of 1000 iterations, which resulted in low autocorrelation. Convergence of models was assessed by examining traceplots to visualise sampling mixing and by assessing effective sample sizes and autocorrelation.

Effects of site and habitat –To investigate the extent of phenotypic variation among sites and habitats, we compared three linear mixed effects models with different random effects structures using the deviance information criterion (DIC) to identify the model with the most support, at ΔDIC = 2 : a model with no random effects (i.e., a "null" model), a model that included site as a random effect, and a model that included habitat type as a random effect. These models were designed to identify whether there was phenotypic variance between all habitats or all sites. We grouped traits in the response variable to investigate phenotypic divergence in functionally correlated traits. TL and gut length were each fit as a univariate response; the four defence traits (plate number, DS1, DS2 and PS) and the four gill raker traits (GRL2, GRL3, GRW, GRN) were fit as multivariate response traits, respectively. This resulted in a total of 12 models (three models with different random effects structures for each of the four responses), all fitted with sex as a fixed factor. Models for gut length, defence traits, and gill raker traits included TL as a covariate. Fitting the interaction between sex and TL did not improve model fits, so we present results from models with sex and TL fit as single term effects. Owing to varying levels of replication for each sex at each site (Table 1) we cannot estimate or exclude sex differences in the patterns of divergence. The multivariate models fit full variance-covariance matrices for random effects, allowing us to estimate phenotypic covariances at both individual (residual) and habitat/site levels . Note that because of the different sample sizes for each group of traits (see below), it was not possible to run a single multivariate model with all traits to measure the full covariance matrix. Fish from sites were considered phenotypically divergent if 95% credible intervals of the posterior distributions of predicted trait values did not overlap.

Association with ecological parameters – We next ran a suite of univariate, linear mixed effects models that investigated the effect of ecological predictors on each phenotype independently. All models were fitted with TL and sex as fixed effects and site as a random effect (except for model on total length, which only had sex as a fixed factor). Because many of our eight ecological variables were highly correlated (Table 1), we ran one model per ecological predictor per phenotype (total = 7 models per phenotypic trait) and compared each to a null model without any ecological variables using DIC (as above). In these models, all predictor variables

were standardised to have a mean of 0 and standard deviation of 1 to improve comparability between models. We present the mode and 95% credible intervals of the posterior distribution for the linear coefficients for each ecological predictor, unless ΔDIC to the null model was > 2 because this indicated that the null model was not improved by fitting the ecological variable.

Genomic divergence

To identify the extent of genetic divergence among sites, we used two approaches: 1) Principal Component Analysis (PCA) was used to explore genetic clustering of Mývatn stickleback, and 2) a model-based admixture analysis was used to determine population genetic structure by calculating the proportion of an individual's genome that originates from different hypothetical ancestral gene pools (i.e., admixture coefficients). PCA was conducted in ADEGENET package , and as suggested by 100% of the initial PCs were retained when identifying the number of clusters. Admixture analyses were run using SNMF, which estimates admixture coefficients using non-negative matrix algorithms and makes no assumptions about drift or Hardy-Weinberg equilibrium (HWE). SNMF was run using the LEA package in R for number of ancestral populations (K) $1 - 10$. Pairwise genome-wide Nei's F_{ST}s were calculated between all sites and habitats using VCFTOOLS. Contemporary effective population size (N_e) was estimated using the LD-based method, as implemented in NeEstimator v2 . To remove physical linkage (as required in this method), we first thinned the SNPs across 10KB windows using VCFTOOLS. To provide confidence intervals of estimates of *Ne* , we calculated *Ne* across each chromosome independently.

Genome-wide-association analyses

To estimate SNP-based heritability for each trait, we ran Bayesian mixture models for each trait independently using the BAYESR software . We selected this method as it has been found to be more accurate in estimating additive genetic variance explained by SNPs than alternative methods (e.g., BSLMM or LMM). All models included TL and sex as covariates (except when modelling TL explicitly, which included only sex as a factor), and all traits were standardised to a mean of 0 and standard deviation of 1. We also aimed to estimate pairwise genetic correlations between traits using bivariate models in BAYESR. However, none of these models converged, likely owing to the large sample sizes usually required to estimate genetic covariances, and hence are not reported further.

To identify SNPs underlying phenotypic variation, we performed a GWA using a linear mixed effects model approach with the program GEMMA . This program fits models that control for relatedness among samples and/or population stratification, which reduced false positives even with small sample sizes. GEMMA was selected because it implements generalized linear mixed models for traits with non-normal distributions, and therefore robustly handles data on different scales. These models estimate the linear coefficient for the relationship between each SNP in turn and a given trait. P -values for each SNP were calculated using Wald tests. We used a false discovery rate of 5% to correct for multiple testing and a P -value cut off of 10^{-5} when identifying whether a SNP was putatively associated with trait variance. All models included sex as a fixed factor and length as a covariate, except for the model with total length as a response variable which just fix sex. The number of SNPs used in each model was dependent on the number of individuals available for each trait, because gemma only models SNPs with less than 5% missingness. We then identified whether there was overlap between regions that were associated with trait variation in our data, and quantitative trait loci (QTLs) previously mapped on the stickleback genome , using LIFTOVER and custom R scripts.

Genome-environment association analyses

We investigated gene-environment associations using latent-factor mixed models (LFMM) in the R package LEA , which models loci against environmental variables while controlling for unobserved latent variables (i.e. spatial structure or autocorrelation). Models were run for 100 000 iterations, with 10 000 burn-in cycles and 5 replicates. z-scores were combined from five runs, and we used adjusted$P$ -values using the genomic control method (a recalibration procedure which decreases the false discovery rate). Following guidance of , the resulting P -values for the outlier tests were adjusted for multiple testing to Q-values using the false-discovery rate method as implemented in the "qvalue" 2.6.0 package. A Q-value for SNPs of < 0.05 was considered

significant (i.e. a "candidate SNP"). We compared DIC of models that included one ecological variable at the time to try to identify which ecological variable best predicted the genomic data. To identify whether gene-environment associations were linked to phenotypic variation, we then investigated whether any of the candidate SNPs identified in any of the LFMM analyses were linked (within 5kb) to SNPs identified in GWA analyses.

To explore the molecular function of genomic regions that showed signatures of selection, we first analysed candidate genes for enrichment of molecular functions. To do this, we identified genes that the candidate SNPs were within 5kb of and compared these candidate genes with the reference set of 20 805 genes across the stickleback genome ('gene universe'). Gene ontology (GO) information was obtained from the stickleback reference genome on ENSEMBL using the R package BIOMART , and functional enrichment was investigated using the package TOPGO 2.42 and the Fisher's exact test (at $P < 0.01$). To reduce false positives, we pruned the GO hierarchy by requiring that each GO term had at least 10 annotated genes in our reference list ("nodeSize = 10"). Secondly, physical overlap on the genome between candidate SNPs identified in environment-association analyses and GWA SNPs indicate regions of the genome under selection. For regions of the genome containing both environmentally associated candidate SNPs and GWAs SNPs, we identified (1) the genes in this region (within 5kb), (2) whether haplotypes on that gene in our dataset are predicted to cause variation in protein translation, and (3) the function of the gene. We used the program "snpEff" to detect whether SNPs fall on coding regions changing amino acid sequence.

Results

Out of 200 individuals originally sent for sequencing, 14 samples did not pass the quality control and were therefore not sequenced, resulting in a total of 186 sequenced individuals. Quality filtering after variant discovery resulted in a dataset of 1 205 604 SNPs. During dissections, some samples were broken resulting in slightly different sample sizes for each trait; total length: N = 186, gut length: N = 106, plate number: N = 160, DS1 and DS2: N = 159, PS: N = 158, GRN: N = 133, GRW: N = 159, GRL2 and GRL3: N = 159.

Phenotypic divergence

Total length – Model comparisons showed that TL varied across sites rather than habitats (Table 2). This pattern was predominantly driven by stickleback from HS2 site being shorter than stickleback from other sites (Figure 1). TL was negatively correlated with both temperature (Table 3) and bird density (Table 3), but the model with temperature as a predictor fit the data better (i.e., had a lower DIC; Table 3). All traits, except plate number and GRN, were positively correlated with TL (Table 2), and all results presented hereafter refer to effects on size-corrected traits.

Defence traits - There were no sex differences in the relative length of either dorsal or pelvic spines, but males had more armour plates than females (Table 2). Model comparisons suggested that defence traits tended to vary according to habitat rather than site, although there was only very weak statistical support for this effect (Table 2, Figure 1). Whilst none of the ecological variables predicted relative length of DS1 and armour plate number (Table 3), relatively length of DS2 increased as the density of piscivorous birds increased (Table 3) and relative length of PS increased in deeper water (Table 3, Figure 2). However, statistical support for these environmental associations was weak (the ΔDIC to the null model was within 2, and the lower credible interval of the posterior distribution of the linear coefficient was only just above zero; Table 3).

Individuals with relatively longer DS1 had correspondingly longer DS2 and PS (pairwise phenotypic covariances (CoV), posterior mode and 95%$_{\text{CI}}$: DS1:DS2 = $0.224_{(0.165,\ 0.289)}$; DS1:PS = $0.133_{(0.077,\ 0.190)}$; DS2:PS = $0.122_{(0.070,\ 0.190)}$), indicating that spine traits covaried at the individual level. However, relative length of spines did not seem to covary with armour plate number (see Table S2). There was no evidence for phenotypic covariance across the habitats between any of the defence traits, suggesting that spatial divergence in defence traits was not correlated at the habitat level (see Table S2). However, to increase statistical power for detecting divergence in phenotypic covariances, greater level of replication is likely needed.

Trophic traits - Males had relatively longer GRL2 and GRL3, and relatively more gill rakers (GRN), than

females (Table 2), but there were no sex differences in GRW (Table 2). Model comparisons showed that gill raker traits varied according to site rather than habitat (Table 2), but this variation was not associated statistically with any of the ecological variables (Table 3). Specifically, stickleback from the CS site had relatively longer and fewer gill rakers than stickleback from other sites, and stickleback from site CS and DN had fewer gill rakers than stickleback from all other sites (Figure 1). There was evidence for phenotypic covariances at the residual (i.e., individual) level for some gill raker traits. For instance, individuals with relatively longer GRL2 had correspondingly longer GRL3 (CoV, posterior mode and 95%$_{\mathrm{CI}}$: GRL2:GRL3 = $0.288_{(0.188,\ 0.403)}$). Furthermore, individuals with more gill rakers also had narrower GRW (CoV, posterior mode and 95%$_{\mathrm{CI}}$: GRW:GRN = $-0.184_{(-0.311,\ -0.052)}$)). There was no evidence for phenotypic covariances at the site level between any of the gill raker traits, suggesting that spatial divergence in gill raker characteristics was not correlated (see Table S2).

Gut length - Males had relatively shorter guts than females (Table 2). Model comparisons suggested that relative gut length varied among sites (Figure 1, Table 2), but this variation was not associated to any of the ecological variables (Table 3).

Genomic divergence

Using 1 205 604 SNP markers across 186 individuals, principal component analysis (PCA), pairwise F_{ST} and admixture analyses all revealed little genome-wide differentiation. First, the first two principal components (PC1 and PC2) explained only 0.009 and 0.008% of variance, respectively (see supplement). There was some clustering of individuals along PC1 (Fig S4), which could be indicative of clustering based on genotype (associated with, for example, spatial structure, sex or phenotype), on areas of low recombination or, alternatively, arise as a statistical artefact. Our PCA did not cluster individuals according to any *a priori* hypotheses for which we had data (i.e., according to basin, habitat type, site, sex, phenotype or tapeworm presence). Second, pairwise Nei's F_{st}were very low at both habitat and site level (Table S3). Third, admixture analyses using snmf revealed that the model with K = 1 had the lowest cross-entropy score, suggesting that K = 1 best predicted the population genetic structure in our dataset (Figure 3). This suggests that Mývatn stickleback form one panmictic population. The effective populations size *(N_e)* for the population was estimated as 1976 (± 640).

Genome-wide-association analyses

Additive genetic variance (V_A) ranged from 0.10 for relative gut length to 0.62 for GRN and 0.69 for plate number (Table 4). SNP based heritability ranged from 0.42 (for relative gut length) to 0.62 (for plate number) and 0.65 (for GRL2) (Table 4). We found between five and 100 putatively causal SNPs associated with variation in each trait (Table 4). In nine cases, pairs of SNPs that were associated with trait variation were also closely linked (within 1kb) on the genome. All of these were between functionally similar traits. Specifically, one 1kb region (on chr VII) had SNPs associated with DS1 and DS2, seven regions (on chr I, IV, XIV and XVIII) had SNPs associated with GRL2 and GRL3, and one region (chr VI) had SNPs associated with both GRL3 and GRW. Genomic regions associated with the relative lengths of GRL2, GRL3 and the gut were linked to the same 10 QTLs for various feeding traits, and the relative length of GRL3 was linked to a further 28 QTLs for feeding traits, including gill raker length, spacing and number. Genomic regions associated with relative GRW overlapped with all those for GRL2 and GRL3, as well as a further 68 QTLs for feeding traits, including for lateral, medial and middle raker spacing. Regions of the genome associated with lateral plate number in our analysis were physically linked to 17 regions with QTLs previously identified for armour plating phenotypes, nine of which were for lateral plate number. Most notably, the major peak on the Manhattan plot for lateral plate number in our analyses (see Fig S5) fell on the major effect *EDA* gene on chr IV, which is known to control the recurrent plate loss in freshwater stickleback . All trait related SNPs were linked to regions on the genome with previously mapped QTLs for landmark positions of body shape (; Table S4).

Genome - environment association analyses

The number of candidate SNPs significantly correlated with each ecological variable ranged between 14 000

and 16 650 (Table 5), on 6000 - 7500 genes. Many of the candidate SNPs were the same across models, likely due to the highly correlated nature of the ecological variables. The model for zPC2 (which described the negative covariance of*Rotifers* and *Alona* sp. with other cladocerans and copepods) had the lowest DIC, suggesting that genomic variation was best predicted by invertebrate composition, followed by water depth and temperature (Table 5).

We found between 18 and 34 enriched molecular functions per ecological variable (Table 5). The biological process occurring most commonly in enriched terms across all environments was the development of the central nervous system, followed by other developmental processes (i.e., the development of circulatory system, kidneys, tissue, embryo development, fin development, and ossification) and metabolic processes (most commonly involving protein metabolism). For temperature, protein phosphorylation was the most significantly enriched; for depth, it was ion transmembrane transport; for bird density and zPC3, it was the cell surface receptor signalling pathway category (i.e., signal transduction); for CPUE, it was cell adhesion; for zPC2 it was neuron development; and for zPC4 it was nervous system development (Table 5). Interestingly, we found that terms associated with sensory systems (including response to abiotic stimulus and perception of sound) were enriched for zPC3 and zPC4, which each reflect different components of the community structure of potential prey species. Loci associated with aspects of kidney function, including renal filtration, were enriched for temperature, depth and bird density, and terms associated with immune system and function were enriched in relation to temperature, depth and zPC4.

We found that candidate SNPs identified in LFMM analyses were linked on the genome to putative QTLs (as identified in GWAs analyses) for DS1 (for all environments), DS2 (for water depth, zPC2, zPC3 and zPC4), and PS (for water depth, zPC2, zPC3 and zPC4). There was a total of 23 genes on regions of the genome where there were both candidate SNPs identified in environmental association analyses, and QTLs associated with phenotypic variance (see Table S5). Of those genes, *ULK2* and*CCNB1* genes had haplotypes causing changes in amino acid sequence.

Discussion

Inferring the genomic basis of adaptive trait variation, including the interplay between adaptive divergence and gene flow, remains a central endeavour in evolutionary biology. To increase rigor, data on both selective agents and organisms' multivariate phenotypes are needed when identifying regions of the genome potentially implicated in adaptation . We show that there was phenotypic divergence in several traits of threespine stickleback from Mývatn, Iceland, but little evidence for genome wide population genomic structure, suggesting extensive gene flow. However, a combination of GWA and landscape genomics approaches allowed us to isolate genomic regions associated with both environmental and phenotypic variation, suggesting genomic divergence in response to natural selection in face of gene flow.

Phenotypic divergence in the face of gene flow

Mývatn stickleback showed spatial divergence in several traits, some of which was associated with ecological variation. Notably, stickleback were smaller at the warm shore site, had more armour plates and relatively longer first dorsal spines in the North than in the South basin, and relatively longer, but fewer, gill rakers in the rocky shore site than in the rest of the lake. These results align partially with previous findings from the same population sampled in 2009 , which found spatial phenotypic divergence in body size (N basin: larger), spine length (N basin: longer spines) and gill raker gap width, but no divergence in lateral plate number. The number of armour plates and length of spines are important predator defence traits in stickleback , with increased predation pressure by birds or fish selecting for more armour plates and/or longer spines . Mývatn stickleback represent a low plated morph (range 3 to 10) typical of freshwater environments , but the higher plate number and longer spines in the N than the S basin indicates stronger predator induced selection in the N basin. This was further supported by phenotype-environment association analyses that showed that Mývatn stickleback had longer second dorsal spines where piscivorous waterfowl density was higher, and relatively longer pelvic spines in deeper water. The latter may reflect variation in predation pressure by altering the occurrence and visibility to visual predators (e.g., salmonids;).

Lake fish, including stickleback, often diverge along the benthic-limnetic axis (reviewed in Wagner & Seehausen 2014), whereby fish that specialise on a benthic diet have fewer and shorter gill rakers than those on a limnetic diet . Interestingly, we show that gill raker divergence (i.e., shorter and more *versus* longer and less gill rakers) did not follow a typical benthic-limnetic divergence. Phenotypic variation in gill raker morphology often reflects diet mediated trait variation and can facilitate resource polymorphism, from individual specialization (Bolnick *et al.* 2003) to spatial divergence among contrasting environments. Whilst we found some spatial divergence in gill raker number and length, gill raker morphology was not associated with any of our environmental measures (including prey abundances) suggesting that gill rakers may be responding along an unmeasured axis of divergence. Moreover, our findings contrast to some extent with who found divergence between Mined and Warm habitats in gill raker number and gap width rather than gill raker number and length. This discrepancy might suggest that spatiotemporal variation in availability of stickleback prey in Mývatn , may induce fluctuating selection – an interesting target for future research.

Despite observed phenotypic divergence, we found no evidence for genome-wide divergence, suggesting a single panmictic population and phenotypic divergence in face-of-gene flow. Although this aligns with results of , they contrast with those of who found genetic divergence between two sites and suggested the presence of two morphs 'Lava' and 'Mud'. As above, this discrepancy in genetic divergence may reflect temporal changes in the extent of spatial genetic divergence and the loss of previous ecotypes, potentially as a result of strong population fluctuations in stickleback population size . Notably, population demographic analyses indicate that the high-density North basin periodically subsidises the low-density South basin through dispersal , which could result in periodic gene flow. However, it is also possible that differences in the resolution of both spatial sampling and sequencing may have resulted in the discrepancies between studies.

Gene-phenotype-environment associations: evidence for a role of natural selection?

Gene flow is thought to constrain divergence by swamping locally adapted alleles (but see), and while phenotype-environment associations can indicate responses to natural selection , they may also reflect phenotypic plasticity and non-random dispersal . Evidence for selection is therefore strengthened when observed trait divergence has a genomic component (i.e., "gene-phenotype-environment" associations). Importantly, all traits we measured had substantial additive genetic variation, suggesting high levels of potential for traits to respond to selection. This was further reflected at the genomic level, where we were able to successfully recover similar genomic architecture as previously reported for this species. Using the mapped genomic architecture, we show that while divergence of most traits was not linked to a genomic component, relative spine lengths was. Specifically, genomic regions that were associated with pelvic spine length (that were longer in deeper waters), were divergent across water depth strengthening the conclusion that the observed divergence in pelvic spine length is a result of divergent natural selection – likely due to predation.

Mapping the causative pathways from genes to phenotypes is a notoriously challenging task , but identifying the genes on regions involved in gene-phenotype-environment associations provide insight into the molecular mechanisms underlying adaptation. On the genomic regions associated with the pelvic spine length - depth relationship, we found two genes (ULK2 and CCNB1) with haplotypes coding for alternate amino acids. These are functionally interesting as ULK2 is a serine/threonine kinase which, along with its homolog ULK1, interacts with the master regulator of metabolism (mTOR) and regulates apoptosis in response to starvation, and *CCNB1* codes for Cyclin B1, a major kinase regulator which activates mitosis and regulates the dynamics of the cell cycle. Whilst these findings suggest candidate molecular functions underlying responses to natural selection, explicit follow up studies are needed to get at the causal relationships with pelvic spine length variation.

Many of the regions of the genome that were correlated with environmental variation in our study were not linked to observed trait divergence. This is not surprising given that we only measured a selected subset of traits within specific functional categories. An understanding of what biological functions these regions are associated with can therefore provide hypotheses about targets of divergent natural selection for future studies. In our study, the biological processes implicated in genomic regions included developmental processes, such as development of nervous and sensory systems. Notably, visual response to abiotic stimulus

and perception of sound were enriched biological functions that were associated with the community structure of invertebrates. This is particularly interesting for Mývatn stickleback due to potential for influencing ability to respond to prey stimulus, and because the sensory drive hypothesis posits that an organisms' communication system should be especially sensitive to ecological variation .

Although adaptation can occur across exceptionally short time frames , selection often acts multifariously and environments are rarely stable temporally, resulting in fluctuating selection pressures . Mývatn is a highly dynamic ecosystem, where multiple dimensions of its ecology and, importantly, stickleback population size fluctuate substantially through time . Our environmental, phenotypic and genomic data were collected at a single point in time, and may therefore not accurately reflect the selective environment experienced by Mývatn stickleback within or across generations. Whilst using ecological data collected at the same (single) time point as genomic and phenotypic data is common practice in studies that investigate selection in wild populations , tracking gene-phenotype-environment associations through time would allow inferences on how patterns of phenotypic and genomic variation withstand, or respond to, temporal fluctuations in selective pressures. In context of Mývatn stickleback, this would also facilitate the exploration of the spatiotemporal balance between adaptive divergence and gene flow.

Conclusions

Our results provide evidence for substantial genetic variation underlying functionally relevant traits, and suggest adaptive divergence in face of gene flow in a large, panmictic population inhabiting a heterogeneous lake. In particular, gene-phenotype-environment association analyses allowed us to identify genomic signatures of selection by testing which phenotypic traits and genomic variants are associated with putative selective agents. Whilst we found evidence for genome-phenotype-environment correlations for spine length, we also found evidence for phenotypic divergence in body size (total length) and trophic morphology (gut length, gill raker length and gill raker number) without apparent genomic divergence – despite substantial additive genetic variation in these traits. The lack of genomic trait divergence across environments could reflect a combination of phenotypic plasticity and/or habitat choice , both of which can constrain or accelerate adaptive divergence . Our study sets the stage for a holistic understanding of patterns of divergence and the maintenance of genomic and phenomic variation in spatiotemporally varying wild populations.

Acknowledgements

We thank all field volunteers who have helped to collect the data presented in this manuscript over the years. We especially thank Antoine Millet for sampling and processing stickleback, and collecting temperature data. These surveys were conducted under the auspices of the Mývatn Research Station, which has government approval for collecting fish specimens from the lake. This work was supported by the Icelandic Research Fund, grant of excellence (195571-052).

References

Data Accessibility and Benefit-Sharing Section: All data used in this manuscript will be made available in the dryad repository upon acceptance for publication.

Author contributions: KS conceptualised, designed and led the research, collected and analysed data, wrote the manuscript; KR contributed to conceptual development and manuscript writing; BKK contributed to generation of phenotype data; JP contributed to designing phenotypic analyses; AE collected field data; RGS generated the phenotype data; MB contributed ecological data; ZOJ contributed to conceptualisation, all molecular work and manuscript writing. All authors contributed to editing the manuscript and approved the final version.

Hosted file

`Tables and figures_GEP_final.docx` available at https://authorea.com/users/495327/articles/577103-genome-phenotype-environment-associations-identify-signatures-of-selection-in-a-panmictic-population-of-threespine-stickleback