nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-020-01341-8

Intercontinental genomic parallelism in multiple three-spined stickleback adaptive radiations

Isabel S. Magalhaes[1,2,9]✉, James R. Whiting[1,3,9]✉, Daniele D'Agostino[1], Paul A. Hohenlohe[4], Muayad Mahmud[1,5], Michael A. Bell[6], Skúli Skúlason[7,8] and Andrew D. C. MacColl[1]

Parallelism, the evolution of similar traits in populations diversifying in similar conditions, provides strong evidence of adaptation by natural selection. Many studies of parallelism focus on comparisons of different ecotypes or contrasting environments, defined a priori, which could upwardly bias the apparent prevalence of parallelism. Here, we estimated genomic parallelism associated with components of environmental and phenotypic variation at an intercontinental scale across four freshwater adaptive radiations (Alaska, British Columbia, Iceland and Scotland) of the three-spined stickleback (*Gasterosteus aculeatus*). We combined large-scale biological sampling and phenotyping with restriction site associated DNA sequencing (RAD-Seq) data from 73 freshwater lake populations and four marine ones (1,380 fish) to associate genome-wide allele frequencies with continuous distributions of environmental and phenotypic variation. Our three main findings demonstrate that (1) quantitative variation in phenotypes and environments can predict genomic parallelism; (2) genomic parallelism at the early stages of adaptive radiations, even at large geographic scales, is founded on standing variation; and (3) similar environments are a better predictor of genome-wide parallelism than similar phenotypes. Overall, this study validates the importance and predictive power of major phenotypic and environmental factors likely to influence the emergence of common patterns of genomic divergence, providing a clearer picture than analyses of dichotomous phenotypes and environments.

Adaptive radiations are rapid branchings on the tree of life, associated with adaptation to distinct ecological niches[1]. As major sources of biodiversity, their study has revealed much about the evolution of phenotypic variation[1,2]. Adaptive radiations also highlight what is unknown about biodiversity evolution. For example, despite abundant phenotypic diversity, not all trait combinations evolve in every radiation, yet organisms in different places sometimes arrive at very similar endpoints[3,4]. The latter has been observed in cichlid fish[5], *Anolis* lizards[6] and Darwin's finches[7]—famous examples of parallel phenotypic evolution. This suggests that Stephen Jay Gould's contention that evolution is contingent and unrepeatable[8] cannot be completely true. Extensive prior work has examined the role that genetic correlations between traits might play in constraining diversity, but the answers provided have not been entirely satisfactory[3,9,10]. Alongside these processes, it is probable that repeatable patterns of evolution are channelled in predictable ways by common environments and selection regimes. Often termed parallel evolution (distinguishable from convergence by shared evolutionary 'start' and 'end' points, but see refs. [11,12]), this process results from environmental similarities within and between radiations. Striking natural examples of phenotypic parallelism[1,6] support this hypothesis, and the persistent appearance of familiar forms in similar ecological niches demonstrates the importance of selection.

However, a limitation with studies on phenotypic parallelism is that they have concentrated largely on the comparison of pairs of strongly different ecotypes or environments[13–15]. This approach might upwardly bias the prevalence of parallelism with comparisons known a priori to occur repeatedly, effectively constraining the evolutionary endpoint. Furthermore, similar environments are typically assumed on the basis of comparable (typically morphological or life history) phenotypes, concealing the role of individual components of environmental variation in driving parallelism. This compromises our ability to understand adaptation, much of which is likely to be broadly physiological. Such drawbacks highlight the importance of combining measures of phenotype, environment and genomics in studies of parallelism. Addressing this gap is needed for a complete understanding of adaptation[16].

Highlighting consistent signatures of adaptation in the genomes of multiple, independent natural populations has also proven to be a valuable tool for studies on the genetic basis of adaptation[17–19]. However, again our comprehension of the relationship between genomic parallelism and continuous phenotypic or environmental variation is surprisingly poor. A major barrier to combining genotype, environment and phenotype has been achieving the necessary biological replication across all three to make broad inferences and shift from description to hypothesis testing. This has rarely been applied (but see refs. [20–22]), and it remains to be shown whether signals of parallelism obtained from continuous measures are comparable to those from ecotypes and previous studies. Here we use such methods to test for environmentally and phenotypically associated genomic parallelism across radiations of three-spined stickleback fish (*Gasterosteus aculeatus*, hereafter stickleback).

Stickleback provide a powerful natural experiment to test parallelism. They are ancestrally marine but, after the colonization of freshwater across the northern hemisphere, are in the early stages of

[1]School of Life Sciences, University of Nottingham, University Park, Nottingham, UK. [2]Department of Life Sciences, Whitelands College, University of Roehampton, London, UK. [3]Biosciences, College of Life and Environmental Sciences, Geoffrey Pope, University of Exeter, London, UK. [4]Institute for Bioinformatics and Evolutionary Studies, Department of Biological Sciences, University of Idaho, Moscow, ID, USA. [5]Erbil Polytechnic University, Kurdistan Region, Iraq. [6]Museum of Paleontology, University of California, Berkeley, CA, USA. [7]Department of Aquaculture and Fish Biology, Hólar University, Sauðárkrókur, Iceland. [8]Icelandic Museum of Natural History, Suðurlandsbraut, Reykjavík, Iceland. [9]These authors contributed equally: Isabel S. Magalhaes, James R. Whiting. ✉e-mail: isabel.magalhaes@roehampton.ac.uk; J.Whiting2@exeter.ac.uk

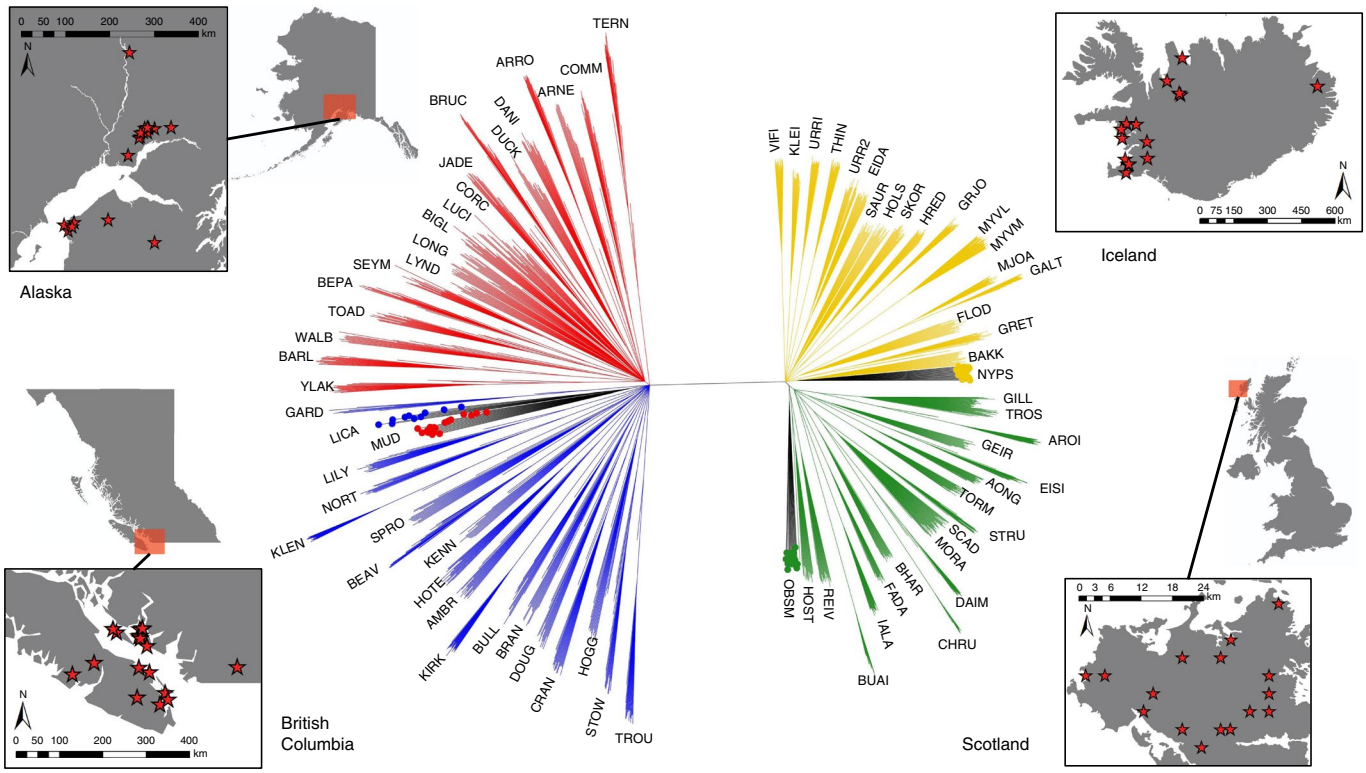

Fig. 1 | Sampling sites and the bootstrapped unrooted NJ tree for stickleback from 73 freshwater lake populations and 4 marine populations from four countries on two continents, based on 11,266 genetic markers for 1,380 individuals. All nodes shown have a bootstrap support of at least 80 (other nodes were collapsed). The freshwater branches are coloured according to the radiation to which they belong. The marine branches are black with the tips coloured according to radiation. The tips represent individual fish, which were generally tightly clustered by population (small labels). The stars represent the lakes sampled.

replicated adaptive radiations (which we define as sets of geographically proximal, closely related populations). Comparing multiple populations derived from the marine ancestor provides a model for exploring both phenotypic parallelism and its genetic basis[23,24] in response to environmental variation. Phenotypic parallelism is well established[25,26], and while it is often considered in dichotomous pairings of marine–freshwater, benthic–limnetic or lake–stream ecotypes, there is substantial continuous phenotypic variation among freshwater populations that has rarely been explored[27] in this context. Parallel genomic loci under selection have been identified across the contrasting ecotype pairs[19,28,29], but the combination of phenotypic and genomic parallelism in one study is rare (although see refs. [20,30,31]) and has not been done at the scale of replicated adaptive radiations across continents.

For this study we sampled 73 freshwater lake populations (1,304 fish) and four marine populations (76 fish) from four adaptive radiations: two from the Pacific coast of North America (Alaska and British Columbia (BC)) and two from Atlantic Europe (Iceland and the island of North Uist (Scotland)) (Fig. 1 and Supplementary Table 1). We consider the variation between lakes within geographic regions as adaptive radiations. Assessing groups of lakes as adaptive radiations distinguishes the stickleback system from other notable systems of adaptive radiations, such as the comparisons of within-lake cichlid radiations[5,14,17]. We quantify parallelism between phenotypes, environments and genomic loci under selection (using restriction site associated DNA sequencing (RAD-Seq) data) in each adaptive radiation and examine how they are associated. Rewards to be gained by connecting the evolution of parallelism more explicitly to the environmental and phenotypic variation include a better grasp of why some traits evolve in concert and a predictive understanding of parallelism and repeatability[4]. This new understanding is essential if we are to reach a consensus on how biodiversity is altered by adaptation.

Results and discussion

Environmental and phenotypic similarity across radiations. We first quantified environmental and phenotypic parallelism across four adaptive radiations to provide an indication of how much genomic parallelism associated with environments and phenotypes to expect. For environments, we quantified the lake area, parasite prevalence (*Gyrodactylus* spp. (Gyro) and *Schistocephalus* spp. (Schisto)) and water chemistry (pH and metal cation concentrations of calcium (Ca), sodium (Na) and zinc (Zn)). For phenotypes, we collected 12 variables associated with shape (3), armour (7) and tropic (2) morphology (Supplementary Table 2). See Methods for detailed information on sampling and measures.

Environment. A principal component analysis (PCA) on all seven environmental variables across all lakes revealed that the first axis of environmental variance (Env_{PC1}) separated lakes along a predominant gradient of pH, with additional minor loadings reflecting Ca and Gyro (Fig. 2a and Supplementary Table 3). This axis did not separate radiations but emphasized the variation from alkaline to acid present in all of them (the most acidic environments were absent in Iceland). Env_{PC2} separated lakes with high and low zinc, with additional minor loadings of Schisto (positive) and Ca (negative). This axis separated European and North American clusters. Interestingly, BC lakes were completely subsumed within the more environmentally variable Alaskan lake cluster (Fig. 2a,b).

Phenotypes—armour. All armour traits loaded positively onto $Armour_{PC1}$ (Fig. 2a and Supplementary Table 3). All radiations

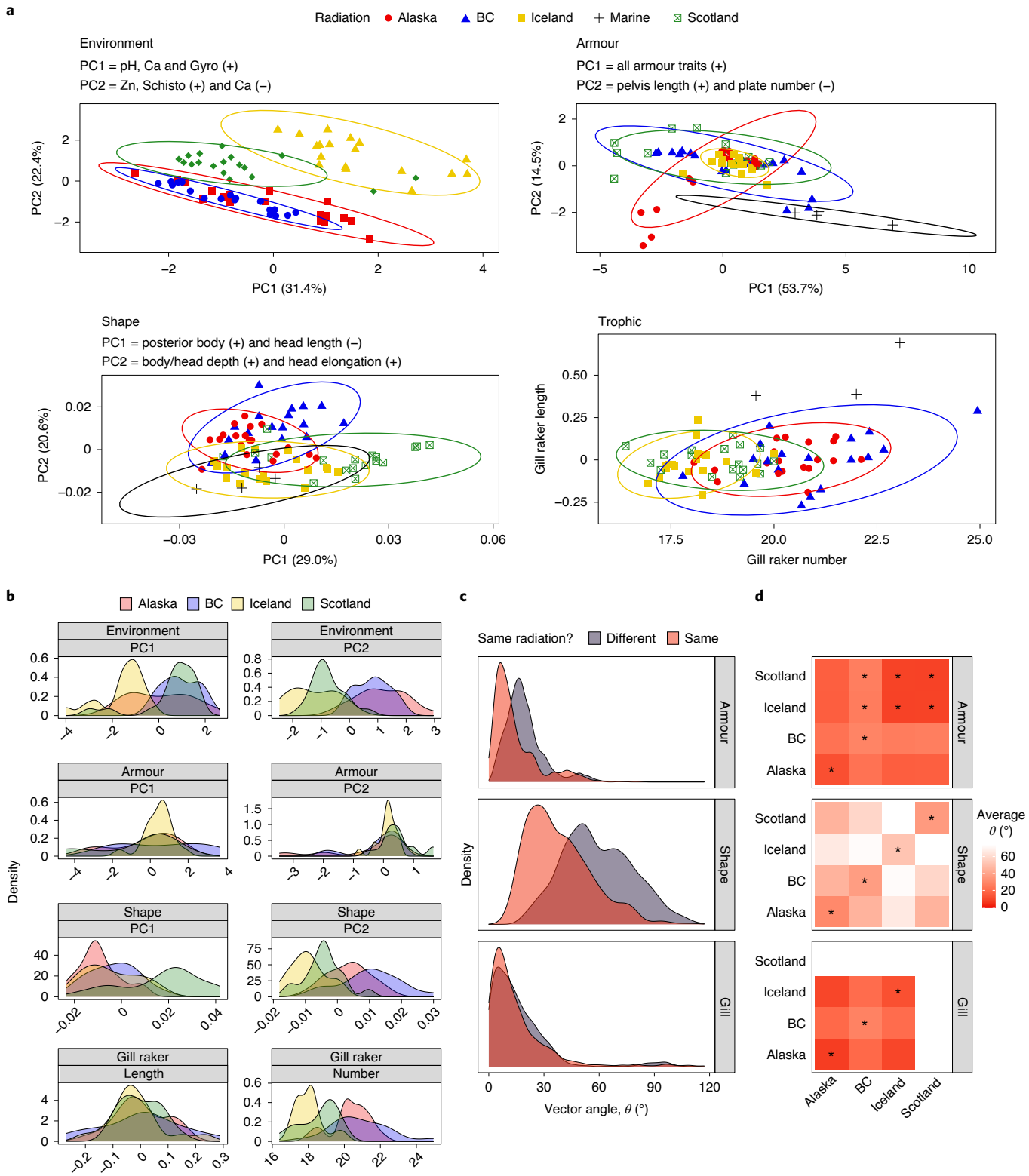

Fig. 2 | Comparisons and analyses of environmental and phenotypic parallelism across adaptive radiations. a, PCAs of environmental variables (environment), regression residuals of Procrustes coordinates against log centroid of body shape (shape), armour traits and trophic traits (gill raker numbers and lengths). Each point represents a population, and the ellipses are 95% confidence ellipses. The names of variables with the highest positive (+) or negative (–) loadings in are subtitles. Marine populations (+) are projected where data were available using PC loadings calculated with freshwater populations only. **b**, Density distributions of populations along each PC axis grouped by adaptive radiation. **c**, Angles between marine–freshwater vectors for armour, shape and gill raker traits. For each category of traits, the density distributions are coloured according to whether the vectors were compared within or between different radiations. **d**, Heat maps of comparisons of vectors between specific radiation pairings coloured inversely according to average angle. Asterisks within panels denote whether between-radiation vectors do not differ significantly ($P > 0.05$) from within-radiation vectors (defined as radiation on the y axis). If between-radiation vectors do not differ significantly from within-radiation vectors, a common trajectory of marine–freshwater phenotypes is assumed.

overlapped on $Armour_{PC1}$ (Fig. 2b), but there were also significant differences in $Armour_{PC1}$ values between radiations (Supplementary Table 4). Scotland and Alaska had populations with extreme armour reduction (low $Armour_{PC1}$), but in Scotland these populations also had complete loss of armour plates (high $Armour_{PC2}$), which were retained in Alaska (low $Armour_{PC2}$). These deviations produce the anomalous relationship between $Armour_{PC1}$ and $Armour_{PC2}$ in Alaska (Fig. 2a). Iceland exhibited minimal variation in armour traits compared with other radiations. Aside from a few populations from BC, there was no overlap in armour traits between marine and freshwater populations. Marine populations have a higher number of lateral plates and generally more exaggerated armour traits. Importantly, however, the projection of marine armour phenotypes suggests that they fall on the same parallel axis, but the marine morphospace is beyond freshwater space (Fig. 2a).

Phenotypic change vectors analysis (PCVA; see Methods) highlighted phenotypic constraint along a common axis for armour (mean angle (θ) within radiations, 14.5°; between radiations, 19.5°) (Fig. 2c and Supplementary Table 5). Indeed, the θ values between vectors from Iceland and Scotland were not significantly larger than the θ values between vectors within Iceland or within Scotland ($P > 0.05$), suggesting a highly parallel axis for marine-to-freshwater transition in Europe. There was some divergence away from this axis, particularly in Alaska, with significant deviations ($P < 0.001$) in trajectory through the trait space (Fig. 2d). This probably results from Alaskan $Armour_{PC2}$ deviations, but these results suggest some non-parallelism between North America and Europe. Despite significant variation among vector trajectories, the average θ remained low (≤28.3°), suggesting that these significant differences in vector angle represent relatively minor idiosyncrasies along an otherwise conserved, parallel axis.

Phenotypes—body shape. Body shape differed significantly between lakes and radiations (Supplementary Table 4), despite overlap within the morphospace (Fig. 2a,b). Scotland exhibited the most extreme body shapes (high $Shape_{PC1}$: elongated, slender bodies and small heads). BC populations had particularly deep bodies and long, deep heads (high $Shape_{PC2}$). Surprisingly, Scotland, the smallest region sampled (303 km^2), had the most shape variation. Marine populations had low $Shape_{PC1}$ and $Shape_{PC2}$ scores, overlapping with some freshwater populations, mostly from Iceland (Fig. 2a). Shape PCVA highlighted greater non-parallelism compared with armour (mean θ within radiations, 38.3°; between radiations, 59.3°), with numerous orthogonal comparisons (Fig. 2c). All between-radiation comparisons of θ were significantly greater than within-radiation comparisons ($P < 0.001$) (Fig. 2d), demonstrating that marine–freshwater shape phenotype transitions have occurred along variable, non-parallel trajectories.

Phenotypes—trophic morphology. Generally, gill raker length increased with gill raker number. Alaskan trophic morphology was the most variable, and BC contained a subset of that variation (Fig. 2a,b). Marine–freshwater trophic PCVA highlighted parallel changes (mean θ within radiations, 15.7°; between radiations, 16.6°), although we lacked marine trophic data from Scotland (Fig. 2c and Supplementary Table 5). Significant deviations among Alaska, BC and Iceland (Fig. 2d), however, suggest some idiosyncrasy among radiations. Freshwater populations had shorter gill rakers for their size relative to marine populations.

Relationship between environmental and phenotypic similarity. We used generalized linear models (GLMs) to test for parallel associations between each environmental variable and morphology (the full results are in Supplementary Table 6). Between armour traits and environmental variables, only $Armour_{PC1}$ was significantly associated with pH in a parallel way across all radiations ($F_{1,71} = 9.20$; false discovery rate (FDR) = 0.006; σ^2 explained, 11.3%). There were also non-parallel, radiation-specific associations (slopes varying between radiations) between armour and several other environmental variables, such as the North American–specific associations between $Armour_{PC2}$ and both Ca and Na. This is consistent with previously reported non-parallel associations between armour and Ca between North American and European radiations[32]. These results highlight parallel reductions of skeletal traits ($Armour_{PC1}$) with lower pH and suggest that pH is a predictor of phenotypic armour parallelism globally. Given that Iceland lacks more acidic freshwater habitats, this may explain why Iceland has comparably limited armour variation.

$Shape_{PC1}$ and Na were associated in parallel across our dataset ($F_{1,71} = 6.51$; FDR = 0.018; σ^2 explained, 46.2%), whereby fish at lower salinity levels were elongated and had small heads. $Shape_{PC1}$ and Gyro were associated, but only in Europe ($F_{3,65} = 3.01$; FDR = 0.045; σ^2 explained, 48.7%). $Shape_{PC2}$ varied in a parallel manner with Schisto ($F_{1,71} = 46.4$; FDR ≤ 0.001; σ^2 explained, 55.5%), potentially because of body shape distortion that can occur as a result of *S. solidus* infection. Although the causative factors driving the link between body shape and water chemistry are largely unknown, similar relationships between salinity and stickleback body shape have been found in other lakes[33].

We expected trophic morphology to evolve in response to zooplankton communities, which can vary in response to water chemistry[34,35], particularly Zn. Accordingly, we found parallel, inverse associations between gill raker number and Zn ($F_{1,71} = 63.6$; FDR ≤ 0.001; σ^2 explained, 53.4%).

The parallel environment–phenotype associations described here (either globally or between specific radiations) might thus be expected to be underwritten by parallel genetic variation. This is particularly true for environmental variables that also overlap between radiations (such as pH and Ca) and the phenotypes associated with them. These predictions are contingent on traits having a simple genetic basis, however[28,36,37], and may not hold for phenotypes with more complex, polygenic architectures.

Phylogenetic relationship among radiations. Older lineages, or lineages not experiencing gene flow, are expected to share less ancestral variation as a function of common ancestry, which may constrain parallelism[18,38,39]. Recent studies have highlighted this as a probable limitation of parallelism in adaptive radiations across continental scales[32,40–42]. Knowing the genetic relationship across all populations is therefore important to interpret patterns of genomic parallelism. A neighbour-joining (NJ) tree based on 11,266 unlinked single nucleotide polymorphisms (SNPs) showed that the four geographic locations form four well-resolved radiations (Fig. 1). TERN in Alaska is slightly anomalous, sitting at a shorter evolutionary distance from European lakes compared with the rest of Alaska, which suggests more recent colonization by the common ancestor. Interestingly, Icelandic and Scottish marine populations clustered separately with their respective freshwater populations, but both North American marine populations clustered together. This may indicate stronger structuring of marine populations in the Atlantic relative to the Pacific or a re-invasion of Alaskan marine regions by BC marine populations. A PCA on freshwater populations confirmed that radiations form independent clusters, with the dominant axis of variation (PC1 = 36.0%) separating Pacific/North American and Atlantic/European radiation pairs (Supplementary Fig. 1). PC2 (7.0%) and PC3 (5.8%) separated North American and European radiations, respectively. Geographically adjacent radiations were also the most genetically similar (Supplementary Table 7; Alaska and BC, mean pairwise $F_{ST} = 0.198$); Scotland and Iceland, $F_{ST} = 0.194$), suggesting that despite independent clustering, lineage splits may be relatively recent, or that gene flow may be occurring within the Atlantic and Pacific groups. Divergence across continents

was stronger and deeper ($0.314 \leq F_{ST} \leq 0.338$) than within oceans/continents, consistent with previous studies that have estimated the time of divergence between Atlantic and Pacific stickleback at approximately 200,000 years[43,44] (but see ref. [45]).

Shared polymorphisms among adaptive radiations were structured predictably (Supplementary Fig. 2), with most shared sites found between Alaska and BC ($N=11{,}524$) and between Iceland and Scotland ($N=8{,}789$). A considerable number of SNPs were common to all four radiations, however (the SNP was polymorphic in all radiations, $N=6{,}339$), suggesting some global retention of ancestral variation. Not including globally shared polymorphisms, sharing among intercontinental comparisons was minimal ($N \leq 933$). Accordingly, between-continent structuring accounts for the largest proportion of molecular variance in our data (analysis of molecular variance (AMOVA): $\sigma = 889.7$, 34.7%). Within continents, populations within radiations ($\sigma = 366.5$, 14.3%) were more genetically variable than radiations themselves ($\sigma = 142.3$, 5.6%) (Supplementary Table 8). Molecular variance, then, is not structured according to geographic scale (continent → radiation → population). This may be due to intracontinental gene flow, bottlenecks at the founding of ancestral Pacific/Atlantic marine lineages[46] or older Pacific/Atlantic divergence times relative to modern freshwater, but it is not possible to differentiate between the scenarios without demographic modelling.

Phenotypically and environmentally associated SNPs and genomic regions within radiations. We associated allele frequency changes with each environmental and phenotypic variable ($N=19$) within each radiation (18–19 populations) using Bayenv2 (ref. [47]) (Methods) and compared outlier genome windows across radiations to identify parallel genome changes.

Several thousand SNPs for each radiation were highly associated (high Bayes factor ($>\log_{10}(1.5)$) and top 5% of Spearman's ρ; see Methods) with environmental and phenotypic variables (Supplementary Table 9). We mapped the SNPs onto non-overlapping 50-kilobase (kb) windows, consistent with approximate linkage disequilibrium (LD) in the stickleback genome[48,49], but we also examined 75-kb, 100-kb and 200-kb windows (Supplementary Data 1) and windows of equivalent genetic distance (0.1 cM), which confirmed that our results were robust and not influenced by variable linkage across the genome (Supplementary Note 1).

We found 1,836 unique 50-kb windows associated with an environmental variable or a phenotypic trait (Supplementary Data 1), ranging from 146 windows associated with pelvic spine length in BC to 11 associated with lake area variation in Scotland. These strong signals of association, even across modest variation, support the adaptive nature of these radiations. Across unique windows, 454 were associated with both an environmental and a phenotypic variable in the same radiation, suggesting that some phenotypically associated regions are also responsible for local adaptation to environments (Supplementary Data 1). This result also suggests that measuring important aspects of the environment may provide profitable ways of identifying candidate regions for adaptation and cryptic phenotypes, such as physiology.

Genomic parallelism associated with environmental and phenotypic variation across radiations. To assess genomic parallelism, we compared observed windows associated in two or more radiations (parallel windows) against randomly permuted (10,000 iterations) null distributions. We first quantified the overall level of genomic parallelism associated with groups of environmental or phenotypic variables. There were no environmentally or phenotypically associated 50-kb windows parallel in all four radiations for individual variables (randomized permutations $N_{\text{Expected-Environmental}} = 0$, $N_{\text{Exp-Pheno}} = 0.0002$), but one window was parallel in a group of three radiations (chrIV: 14400000–14450000 associated with length of pelvis in BC, Iceland and Scotland) ($N_{\text{Exp-Env}} = 0.05$, $N_{\text{Exp-Pheno}} = 0.149$, $P = 0.133$). Many windows, however, exhibited parallelism between pairs of radiations. A total of 39 environmentally associated windows (pooled across all seven variables) ($N_{\text{Exp}} = 12.4$; 95% upper limit, 18; $P < 0.001$) and 65 phenotypically associated windows (pooled across all 12 variables) ($N_{\text{Exp}} = 30.7$; 95% upper limit, 40; $P < 0.001$) were parallel between two radiations (Supplementary Table 10). Parallelism was disproportionately greater for armour (46/65) and gill raker traits (12/65) (mostly number) than for shape (7/65) ($\chi^2 = 6.506$, $P = 0.04$). This is consistent with skeletal traits with simple genetic architectures[28,36,37] being more likely to show evidence of phenotypic parallelism. Interestingly, parallel associated windows (mean SNP $N=7.13$, s.d. = 4.4) had on average more SNPs per window than non-parallel windows (mean SNP $N=6.27$, s.d. = 4.3) (GLM, $\text{LRT}_{1,3867} = 22.4$, $P < 0.001$) and exhibited slightly stronger signals of association with variables (mean residual SNPs above expected, 1.82 parallel, 1.60 non-parallel; GLM, $\text{LRT}_{1,3867} = 5.48$, $P = 0.019$).

We next explored parallelism associated with individual environmental and phenotypic variables. We observed significantly more parallel windows than expected for two environmental variables (Ca and pH) (Fig. 3 and Supplementary Table 10). Reflecting expectations, these were the same variables that load onto the shared Env_{PC1} across all radiations (Fig. 2b) and were involved in parallel environment × phenotype interactions in all (pH) or some (Ca) radiations. Furthermore, we did not detect significant genomic parallelism associated with variables that varied between radiations, such as salinity, Zn and *S. solidus* prevalence. These results highlight that common environmental axes, such as the shared acid–alkali axis, promote signals of parallelism in the genome, although parallelism seems limited to specific pairs rather than extending to all radiations. However, not all environmental variables with parallel environment × phenotype interactions produced signals of genomic parallelism, such as Na–$\text{Shape}_{\text{PC1}}$ and Zn–gill raker number.

Five phenotypic variables were associated with more genomic parallelism than expected by chance (Fig. 3 and Supplementary Table 10): four armour traits (second dorsal spine, pelvic spine length, length of pelvis and armour plate number) and gill raker number. These results are consistent with the observed, constrained marine–freshwater armour phenotype trajectory. This also suggests that variation in armour trajectories (for example, in Alaska) is the result of different genotypes being selected in different environments. Additionally, parallel quantitative trait loci (QTL) have been described for gill raker number[50] but not length, which exhibits more plasticity[51]. Shape traits were not associated with any significant genomic parallelism, despite parallel environment × phenotype interactions. The probable polygenic architecture of shape phenotypes, as has been similarly described in cichlids[52], may result in greater redundancy in the genotype–phenotype map, reducing the likelihood of genetic parallelism. Moreover, the partly plastic nature of body shape[46,53] may lead to environment × shape associations via the reaction norm rather than genomic reuse.

Marine–freshwater (M × F) associated windows were more parallel than those associated with specific variables (Fig. 3) and overlapped well with previously identified M × F quantitative trait loci (proportions overlapping: Alaska = 0.85, BC = 0.85, Iceland = 0.81, Scotland = 0.69)[19,48,54] (Supplementary Table 11). Several parallel M × F windows were also parallel for Ca, pH, Na, armour traits and gill raker number (Supplementary Table 12). These results suggest that our methods and sequencing coverage are robust enough to recover known parallel regions and that, unsurprisingly, genomic parallelism associated with freshwater variables is more modest than marine–freshwater parallelism. The latter may reflect subtler variation between lakes compared with stark marine–freshwater contrasts but demonstrates that reduced parallelism for individual fitness components is probably biological rather than methodological.

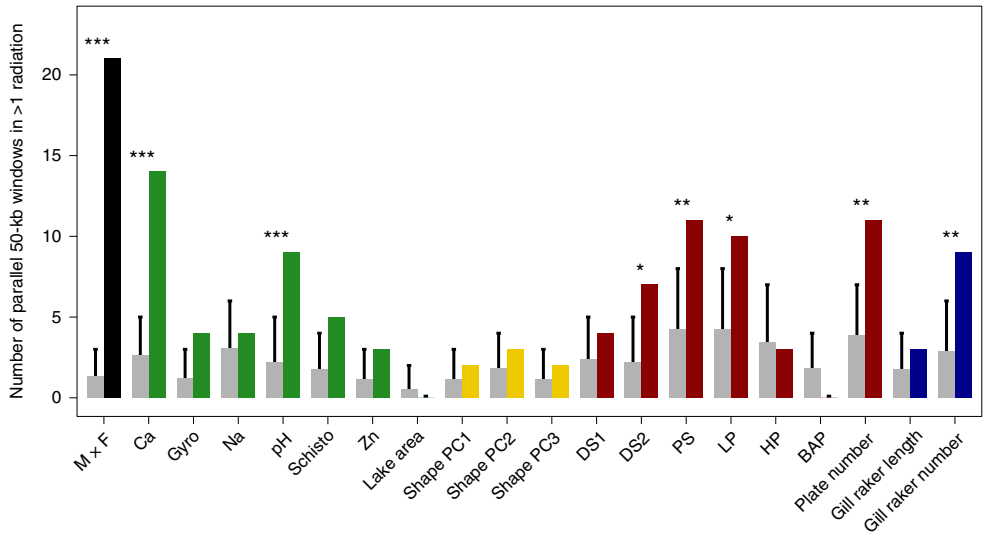

Fig. 3 | Expected and observed counts of 50-kb windows containing an above 99% binomial expectation number of SNPs associated with M × F, environmental variables and phenotypic traits in at least two radiations. Expected bars (grey) represent the mean counts across 10,000 simulated outcomes with 95% confidence intervals per a one-tailed hypothesis. The asterisks denote the significance of FDR-corrected one-tailed tests between the observed counts and the 100,000 simulated counts at the $P < 0.05$ (*), $P < 0.01$ (**) and $P < 0.001$ (***) levels. DS1, first dorsal spine; DS2, second dorsal spine; PS, pelvic spine; LP, pelvis length; HP, pelvis height; BAP, biggest armour plate.

Marine–freshwater comparisons lump together many selective agents without being able to discern which are parallel and which are not. Overall, our results suggest that across a number of comparisons involving two or three (but not all four) freshwater adaptive radiations analysed, the evolution of these phenotypes and environmentally associated traits are disproportionately linked to the same genomic regions.

Finally, we explored genomic parallelism between specific comparisons to identify pairs of radiations with the highest levels of parallelism. We found the greatest number of significantly parallel windows in the comparison between Alaska and BC (Ca, Gyro, pelvic spine length, plate number and gill raker number), followed by Iceland and Scotland (Ca, pH and dorsal spine length) (Supplementary Table 10). Intercontinental parallelism was weaker: Ca and pelvis length for Alaska–Iceland, Schisto for BC–Iceland, pelvic spine length for BC–Scotland and none for Alaska–Scotland. Indeed, restricting the permutations exclusively to Alaska–BC and Iceland–Scotland was enough to recover the significant parallelism observed when all radiations were analysed together (Supplementary Fig. 3). Phylogenetic patterns strongly support the notion that radiations within continents share similar genetic variation, making parallelism through shared standing variation the most parsimonious explanation for these intracontinental biases. Marginal evidence for this was observed, as parallel associated regions had a greater enrichment of shared sites (mean sum of $\log_{10}$ enrichment scores, 1.3) compared with random expectations relative to associated regions that did not overlap among radiations (mean = 1.16, $P = 0.14$) and neutral regions (mean = 1.02, $P < 0.001$). However, comparisons between parallel and outlier regions were not significant. This suggests that radiations may be exploiting pools of ancestral variation for adaptation, but the same ancestral variation is not always adapted in parallel.

Experimental studies of parallelism have increasingly implicated standing genetic variation in the genesis of parallelism in stickleback[55] and other species[56–58]. Coancestry patterns, centred at the focal, causative loci, can discern between parallelisms via de novo mutations, standing variation or introgressed alleles[59]; however, we lack the sequencing resolution to make these comparisons here. Furthermore, the source of parallelism is likely to vary locus by locus and trait by trait, making it difficult to assess with a genome-wide approach. Indeed, all three modes of repeated gene reuse have been observed in the radiation of a wild tomato clade[60].

Linkage and the genomic location of parallel regions. As a reduced-representation approach, selection scans with RAD sequencing depend on LD between markers and functional loci[48,61,62]. LD varies widely across organisms and within genomes but has been relatively well-characterized in stickleback[63,64], and RAD sequencing has been used successfully in this species to test for genomic parallelism[36,48,59,63–65]. LD associates inversely with recombination rate across the genome, so we estimated how recombination may affect our results using a previously published genetic map[65]. Recombination was significantly reduced in associated windows and parallel windows compared with non-associated windows (Supplementary Fig. 4; Kruskal–Wallis, $\chi^2 = 122.21$, $P < 0.001$) but did not differ significantly between associated and parallel windows ($P = 0.55$). Reduced recombination can be an important mechanism in adaptation through maintaining adaptive alleles, as seen in stickleback[66] and cichlids[67].

These patterns may also reflect an increased ability to detect selection in low-recombination windows through increased linkage with causative SNPs. If this possibility is true, our estimates of association (and by extension parallelism) may be conservative if false negatives are pervasive in high-recombination regions. Importantly, our signatures of parallelism cannot be explained by variable recombination. Background selection can produce false-positive signatures of parallelism by reducing local diversity in shared low-recombination regions[68]. This is less of an issue, however, when associating allele frequencies with environmental/phenotypic clines as done here.

Genetic distance (0.1 cM) windows corroborated the 50-kb results, returning significant parallelism for Ca, pH, pelvic spine length, pelvis length, plate number and gill raker number. We also recovered weakly significant parallelism for several other environmental and armour variables (Supplementary Fig. 5 and Supplementary Note 1), suggesting that our 50-kb results may be conservative.

We plotted parallel 50-kb windows (Supplementary Fig. 6) and merged adjacent windows (Supplementary Table 13) to inspect putative causative genes (Supplementary Data 2). Merging adjacently

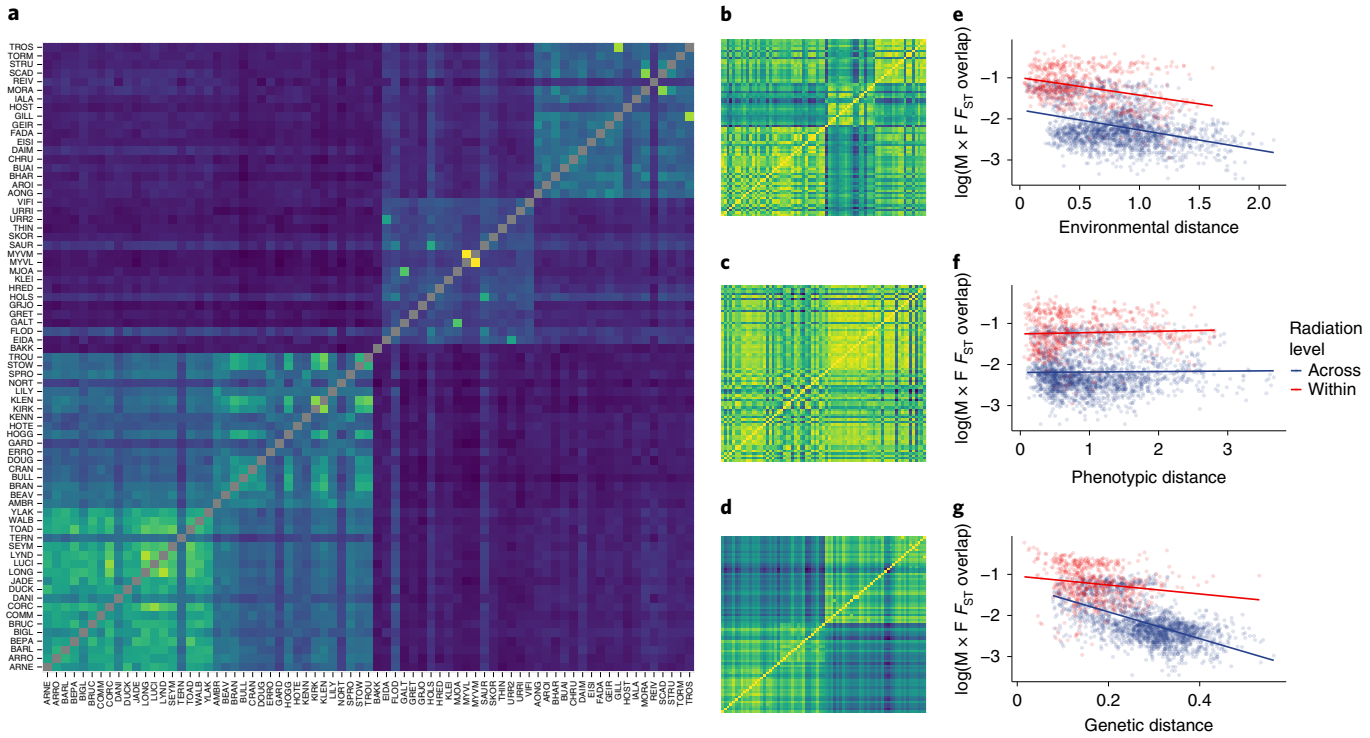

Fig. 4 | Associations between genome-wide M × F F_{ST} and environmental, phenotypic and genetic distance across all pairwise comparisons of 73 freshwater populations. a, Proportion of M × F F_{ST} 50-kb outlier windows that overlap among freshwater replicates. The freshwater populations are ordered as Alaska, BC, Iceland and Scotland, with these locations distinguishable as four clear clusters. **b**, Environmental distances between freshwater populations, recorded as Euclidean distances in PCA space for all seven environmental variables. **c**, Phenotypic (Euclidean) distances between freshwater populations for the 12 phenotypic variables. **d**, Genetic distances between freshwater populations, recorded as genome-wide pairwise F_{ST} on the basis of 8,395 unlinked SNPs. **e**–**g**, Associations between environmental (**e**), phenotypic (**f**) and genetic (**g**) distances and M × F F_{ST} overlap (log-transformed). The points are coloured according to whether the pairwise comparison is being made within a radiation or across radiations.

prior to permutations did not change which variables were significantly parallel (Fig. 3). By doing this, we identified some wider genomic regions with parallelism across multiple radiations. One example was the pooling of plate-number-associated windows in three radiations around the well-known *Eda* gene, with a known functional role in armour plate evolution[19,28,29], which emerged despite the limited plate number variation across freshwater populations.

We also observed adjacent windows around a known inversion[69] region (250 kb) on chromosome I, which contains the genes *igfbp2a*, *stk11ip* and *atp1a1* and was strongly associated with Ca, Na and pH in several radiations (Supplementary Data 2). Removing these windows prior to permutations did not change which variables were significantly parallel (Fig. 3). Inversions can be beneficial for adaptive haplotypes by reducing local recombination, and they have been implicated in genomic parallelism in other systems such as parallel crab/wave ecotypes of *Littorina saxatillis*[70]. Within this region, *atpa1a1* is particularly interesting, given its previously detected association with the major ecological transition from marine to freshwater[71] and its functional role in metal ion management[19,69].

Extensive LD in freshwater populations could result from drift, but it is consistent with strong directional selection after freshwater invasion and has been reported for stickleback populations from Alaska[48]. However, it had not previously been observed for the same regions across several independent adaptive radiations. These results are also consistent with many diverged marine–freshwater SNPs aggregating in just 19 short genomic regions, including three known inversions[69]. Overall, these results suggest that physically linked genomic regions are hitch-hiking in separate radiation pairs, which may contain parallel genes across all radiations but are undetected by our methods.

Relationships between genomic parallelism and phylogenetic, phenotypic and environmental similarity. On the basis of the assumption that freshwater populations radiated from common marine ancestors[1,18], and to leverage statistical power from our biological sampling, we also compared parallel F_{ST} outliers between all M × F comparisons to examine the relative effects of genetic, phenotypic and environmental similarity on genome-wide parallelism at a large geographic scale. To do this, we calculated the top 5% of 50-kb windows on the basis of F_{ST} for each freshwater population and its relevant marine population and assessed all pairwise overlaps ($N = 2{,}628$) (Fig. 4a). Genome-wide F_{ST} outliers are more susceptible to random parallelism through recurrent drift or background selection, although, as discussed above, these processes are unlikely to influence previous results obtained by comparing allele frequency gradients across all populations within a radiation. Nevertheless, broad patterns inferred over all 2,628 comparisons should be apparent despite noise.

Parallel windows were more common in intracontinental than in intercontinental comparisons, again highlighting the importance of these pairings as the major contributors towards pairwise genomic parallelism signals, as discussed previously. This strongly suggests that genomic parallelism at large geographic scales must be partly contingent on shared genetic variation, although exceptions exist, such as recurrent de novo mutation at *Pitx1* (ref. [72]) during freshwater pelvis evolution. There is also the possibility of haplotype sharing between radiations within continents by gene flow through

marine populations, which may be facilitated in North America, despite the greater geographic distance, by a shared coastline connecting Alaska and BC[38].

We used Mantel tests to compare a parallel F_{ST} overlap matrix with genetic, environmental and phenotypic distance matrices (Fig. 4b–d). Across all comparisons, the number of parallel windows was strongly negatively correlated with genetic ($r=-0.67$, $P<0.001$) and environmental dissimilarity ($r=-0.45$, $P<0.001$), but only weakly with phenotypic dissimilarity ($r=-0.07$, $P=0.060$) (Fig. 4e–g). Thus, genomic parallelism increases in populations that are more genetically or environmentally similar, but not in those that are more phenotypically similar. Associations between environmental dissimilarity and F_{ST} parallelism were still strongly negative ($r=-0.41$, $P<0.001$) after correcting for genetic similarity in partial Mantel tests, suggesting that correlations between local environment and local ancestry do not drive this effect. Phenotypic dissimilarity remained unassociated with F_{ST} parallelism after controlling for genetic similarity ($r=-0.026$, $P=0.341$), suggesting that environmental similarity is a better predictor of genomic parallelism than phenotypic similarity, at least in terms of our measured environmental variables and the observable morphological phenotypes in this system. The genotype–phenotype map may be simpler than the equivalent genotype–environment map, potentially leading to overcorrection when including genetic distance matrices, but this would not explain why phenotypic distance associations were weaker before correction.

Early studies of marine and freshwater stickleback[19,29,48] were some of the first to demonstrate the repeatability of the genetic basis of adaptation in natural populations, and that it might be pervasive. These results drove researchers to examine other systems for genomic parallelism, such as cichlids[67,73], periwinkles[74], stick insects[15] and *Arabidopsis*[75]. These diverse systems have highlighted that genomic parallelism is highly variable, and more recent studies of stickleback from outside the original Eastern Pacific populations have similarly shown variability within the system itself[69]. The field has since matured to question how and why this variability persists, and the stickleback system remains at the forefront of this research. Studies on 16 lake–stream stickleback pairs from BC demonstrated that continuous phenotypic and environmental variation predicts genomic parallelism (F_{ST} regions)[20] and suggested that ecotype genomic parallelism may be stronger for certain phenotypic traits than others[76]. These findings are recapitulated and extended here, to demonstrate that similarity of specific environmental and phenotypic variables is a good predictor for signatures of genomic parallelism. Our results also confirm that these results extend beyond BC. Furthermore, our data provide additional statistical power (73 marine–freshwater comparisons versus 16 lake–stream, albeit with reduced independence) to elucidate that environmental similarity is a better predictor of genomic parallelism than is phenotypic parallelism.

A major question of interest concerns the geographic scale to which patterns of genomic parallelism extend. While we found marine–freshwater genomic parallelism to be constrained across all four radiations, we do observe genomic parallelism at a continental scale and within freshwater populations founded from both Eastern Pacific and Atlantic marine populations. Agreeing with our results, a comparison of 'regional' lake–stream parallelism in BC to 'global' lake–stream parallelism in a collection of lake–stream pairs from Europe and North America (one pair per region) highlighted a global constraint on parallelism at the genetic level and for some phenotypes[40]. However, comparable watersheds of multiple lake–stream ecotype pairs are less common beyond BC[40] (but see refs. [77,78]), restricting global comparisons of lake–stream adaptive radiations such as those presented here for marine–freshwater.

A recent comparison of levels of genomic parallelism in North American (Pacific-founded) and European populations suggested that a founding event of Atlantic marine populations limited standing freshwater variation, leading to lower parallelism in European populations[42]. Consistent with the idea that even minor differences in selection may limit genomic parallelism of standing genetic variation[38], this study also speculated that differences in selection homogeneity between North American and European environments could explain variable genomic parallelism. The data to test this hypothesis, however, have hitherto been lacking. Our results on segregating variants and molecular variance confirm distinct North American and European clusters of standing genetic variation, which are consistent with a founding bottleneck or an older divergence time between intercontinental radiations than within each ocean. However, genetic variation in the North American and European radiations was broadly comparable, suggesting similar potential in Europe to produce freshwater parallelism as in North America. This result, combined with our stronger observed environmental homogeneity in North America than in Europe, as well as strong associations between environmental distance and genome-wide parallelism (Fig. 4b,e), suggests that environmental homogeneity is a valid explanation for some of the differences in parallelism observed between the two continents. Overall, our study thus highlights that the variable levels of genomic parallelism observed in marine–freshwater stickleback comparisons at the global scale are probably the result of an interplay between environmental heterogeneity (but not physical distance) at continental scales and a history of founding bottlenecks and segregating genetic variation among founding populations at the global scale. Furthermore, genomic parallelism of specific phenotypes is predictable on the basis of common phenotypic trajectories, probably underwritten by simple versus complex genetic architectures.

Methods

Sampling and environmental data collection. We sampled 18 (19 in Alaska) freshwater lakes and one marine coastal population in North Uist, Scotland (April and June 2013), in Iceland (May and June 2014), in BC (April and May 2015) and in the Cook Inlet basin, Alaska (June 2015). The lake names, geographic coordinates and numbers of samples used in the study are shown in Supplementary Table 1. Each adaptive radiation analysed comprises a variety of different adaptive forms that most likely evolved from closely related ancestral marine lineages in each region. We measured a set of 7 biotic and abiotic environmental variables and a set of 12 phenotypic traits (measures of body shape, armour traits and gill rakers). We measured the pH, concentrations of metallic cation concentrations (Na, Ca and Zn), lake area, and calculated population prevalence of Gyro and Schisto for each lake. The concentrations of cations, pH, lake area and parasite prevalence per lake are shown in Supplementary Table 2. The environmental variables included in our analyses were selected because of their presumed fitness effects and ability to be precisely measured (Supplementary Table 2). For abiotic environmental variables, we chose pH; ionic concentrations of Ca, Na and Zn; and lake area, which have been associated previously with the evolution of body shape, size and armour in stickleback[27,33,79,80]. Many biotic variables are difficult to quantify precisely, so we used the prevalence of two parasites (Gyro (ectoparasitic trematodes) and Schisto (endoparasitic cestode)) that are likely to affect the reproduction and life cycle of stickleback[81–83]. It is important to note that the measured variables might actually be proxies for other, unmeasured variables and not the primary causes of selection. The details of the fish collection and quantification of abiotic and abiotic variables can be found in the Supplementary Methods. As marine–freshwater parallelism is well documented[19,29,84], we compared our results for parallelism across freshwater radiations with well-studied marine–freshwater parallelism in this species, and we used the results as a positive control for the methods used (Supplementary Note 2). Some M × F associated regions detected by our use of Bayenv2 could be the result of differences in allele frequencies between only a few freshwater populations and marine populations (within-freshwater variation rather than explicit marine–freshwater divergence). Such false positives are unlikely, however, as a combination of a high Bayes factor and a high Spearman's ρ requires many freshwater populations to display consistent allele frequency changes relative to marine populations. Some variables that vary among freshwater populations may not vary consistently between marine and freshwater, such that our parallel M × F regions are not expected to be a sum of our single-variable, parallel freshwater regions.

DNA extractions and RAD library preparation and sequencing. Genomic DNA was purified from 10 to 21 individuals from each population, chosen to represent a widely distributed subset of the most environmentally and phenotypically variable lakes (Supplementary Data 3). The extracted genomic DNA was normalized to a concentration of 25 ng μl^{-1} in 96-well plates.

In 2014 we conducted RAD sequencing on samples from Scotland and from Iceland. The sequencing libraries were prepared and processed into RAD following the modified libraries according to ref. [85]. In 2016 we conducted RAD sequencing on samples from BC and from Alaska. The sequencing libraries were prepared following the modified single-digest RAD protocol of ref. [86]. The two RAD-sequencing protocols interrogate the same set of loci across the genome, so that the SNP data are compatible across all four radiations. See the Supplementary Methods for the details of the RAD library preparation and sequencing.

Population genetics statistics and phylogenetic tree. The raw sequence reads were demultiplexed using Stacks v.1.35 (ref. [87]). The numbers of reads per individual are shown in Supplementary Data 3 (see Supplementary Methods for the details on the alignment of reads and Stacks pipeline used). For Bayenv2 and association analyses, autosomal SNPs were called using the following filters in the Populations program within Stacks: SNPs present in a minimum of 8 populations ("-p 8"), in over 50% of the individuals within a population ("-r 0.5"), and with a minimum minor allele frequency (MAF) within a population of 0.05 ("–min-maf 0.05"). After filtering, we retained 26,169, 29,111, 26,937, and 26,990 SNPs for Alaska, BC, Iceland and Scotland, respectively.

For the analyses of population structure and phylogenetics across all radiations, a subset of unlinked SNPs were generated. Here, autosomal SNPs were called using the following filters in Populations: SNPs that were present in all 4 radiations ("-p 4"), in over 50% of individuals within a radiation ("-r 0.5"), with a minimum MAF within a radiation of 0.05 ("–min-maf 0.05"), and only the first SNP per RAD locus was retained ("–write-single-snp"). The values of F_{ST} were bootstrapped and calculated in Populations. This set of SNPs were then pruned for LD in plink using indep-pairwise 50 5 0.2. We also produced an additional linkage-pruned dataset with marine populations (-p 5) with 11,266 SNPs. The unlinked SNP dataset with marine fish was used to construct an NJ tree for all fish in the R package ape, using a distance matrix (*bionj*) computed from the SNP data[88]. The tree was bootstrapped 100 times, and nodes with less than 80% support were collapsed. The PCA analysis of population structure was conducted using plink[89].

Phenotypic and environmental variation—body shape, armour, gill rakers and environmental data analyses. All morphological measurements (body shape, body armour and gill raker traits) were done following ref. [27]. The details of the quantification of phenotypic traits can be found in the Supplementary Methods.

We performed three PCAs: one on the armour traits, another on body shape and another on the six environmental variables. The body shape and armour PCAs were performed on regression residuals of all individuals from all radiations pooled together to extract the common PCs of body shape, armour and environmental variation, and retained axes that explained more than 10% of the total variance. The armour and environmental PCAs were conducted with scaled inputs due to the different units of measurement between variables. The shape PCAs were conducted on morphometric residuals and as such were not scaled. All phenotypic analyses, including analyses of variance and analyses of covariance and plotting were done in R v.3.4.3[90].

For the analyses of covariance, we compared each phenotype (PC1 and PC2 for shape and armour, actual values for gill raker number and length) with each of our seven environmental variables. We explored five possible GLMs: (1) a null model, (2) phenotype varies by environment (linear slope), (3) phenotype varies by radiation (intercepts vary), (4) phenotype varies by environment and intercept varies by radiation (parallel slope, intercepts vary) and (5) phenotype varies by environment in a radiation-specific manner (non-parallel slopes). The best model was chosen by backwards model selection using Akaike's information criterion (AIC), with simpler models (model 1 simplest) preferred if the change in AIC > -2. We then calculated *F* statistics for the resulting GLM. For variable slopes, we used post-hoc Tukey tests to compare the radiations.

Genotype–environment/phenotype associations. For each radiation separately ($N = 18$ or 19 populations), we used Bayenv2 (ref. [47]) to identify associations between genomic allele frequencies ($N = 10$ to 21 individual fish, mean = 17.8), the set of 7 biotic and abiotic environmental variables (Ca, Na, pH, Zn, lake area, prevalence of Gyro and *S. solidus*) and the set of 12 phenotypic traits ($Shape_{PC1}$, $Shape_{PC2}$, $Shape_{PC3}$, first dorsal spine (DS1), second dorsal spine (DS2), pelvic spine (PS), pelvis length (LP), pelvis height (HP) and biggest armour plate (BAP), plate number, gill raker length and gill raker number) mentioned above. For each radiation, a matrix of genetic covariance was calculated using a subset of SNPs limited to a single SNP per RAD locus and pruned for LD ($R^2 < 0.4$) in plink[89]. This cut-off was selected to balance the trade-off between SNPs retained and minimizing the effects of linkage. Covariance matrices were therefore calculated using 9,619, 7,983, 7,300 and 5,705 SNPs for Alaska, BC, Iceland and Scotland, respectively. The covariance matrices were calculated across 100,000 iterations and averaged across 5 independent runs. Bayenv2 was run independently eight times, and the final results were averaged across runs. The purpose of the covariance matrix is to rule out spurious associations between drifting allele frequencies associated with population structure and environmental or phenotypic variation. Some of these correlations did exist in our data, but they poorly explain the signals of genomic parallelism (Supplementary Fig. 7).

Environmentally and phenotypically associated SNPs were selected as having a $\log_{10}$(Bayes factor) > 1.5 and an absolute Spearman's rank coefficient above the 95th percentile. The combination of the Bayes factor and a non-parametric measure of correlation helps avoid selecting SNPs with high Bayes factors due to single or few outlier populations with extreme allele frequencies and environmental or phenotypic variation[91]. The SNPs were grouped into 50-kb, 75-kb, 100-kb, 200-kb and 0.1 cM windows (Supplementary Table 14) to test the robustness of our results across different extents of linkage. Our 50-kb dataset was composed of 4,868 windows with SNPs in all radiations, covering approximately 55% of the 447-Mb genome, with a further 1,940 windows sequenced in two or more radiations providing information on an additional 21.7% of the genome. To evaluate whether the windows were environmentally or phenotypically associated, we adapted the methodology of ref. [92]. We calculated the upper 99% binomial expectation for the number of associated SNPs given the total number of SNPs in a specific window, and we selected windows that had a greater number of associated SNPs than this expectation. We focused on repeated changes within the same genomic regions rather than on reuse of the same mutations. This is because the causal mutations are unknown in most cases and may not be sequenced by reduced-representation sequencing methods such as RAD sequencing. This method also controls for variation in SNP density across windows and ensures that significant windows exhibit consistent allele frequency correlations across multiple SNPs. We visualized the genomic locations of associated windows using Manhattan plots (Supplementary Fig. 6) and plotted the residual number of outlier SNPs above the binomial expectation (Supplementary Fig. 8). Linkage groups I–XXI were visualized (except XIX); windows on scaffolds were not visualized. Finally, we compared these associated windows across radiations to examine those that were parallel.

As a positive control for the methods used, we compared our results for parallelism across freshwater radiations with well-studied marine–freshwater parallelism in this species[19,29] and then examined genomic differentiation between all freshwater populations pooled within a radiation and four marine populations pooled together (one from each geographic region; Supplementary Table 1 and Supplementary Note 2).

Parallelism statistics. We use the term 'genomic parallelism' when referring to repeated changes within the same genomic regions, rather than the strict definition of genomic parallelism that refers to reuse of the same mutations. The use of 'parallelism' terminology is highly variable in the literature[12], but our usage is consistent with stickleback literature[19,48,77] and reflects the parallelism of phenotypes. For all radiation groupings (11 combinations in total: 1 four-radiation grouping, 4 three-radiation groupings and 6 two-radiation groupings), we calculated the significance of parallel window counts using a permutation method. For each environmental or phenotypic variable, we randomly drew *N* windows from each radiation's total pool, where *N* was equivalent to the associated window count for each radiation. We then assessed the overlap of randomly associated windows across radiations and pooled the results over 10,000 iterations. The output from all permutations was used as a null distribution to infer *P* values, which were then FDR-corrected using the R package qvalue[93].

Grouping of adjacent windows and expanding parallelism regions. Windows of 50 kb and above were based on a linkage assumption and to minimize non-independence between windows. There were, however, occasionally adjacent windows associated with the same variable across different groupings. Large regions of relatively strong linkage are plausible if recombination is reduced through processes such as genomic rearrangements. To investigate these, we grouped associated windows that were adjacent as well as those that were direct matches across radiations because the likelihood of adjacent associated windows resulting independently is low, suggesting non-independence and probable linkage. We repeated the above permutations assuming adjacent windows to be single associated regions. These windows are available in Supplementary Table 13.

Multivariate vector comparison of environments and phenotypes. Using the average trait values from marine and freshwater populations, we calculated the vectors of phenotypic change for armour, shape and gill raker variables (three vectors per population) separately between each freshwater population and the marine population from the same radiation. We then calculated the angles (θ) between all vectors from the same radiation (one distribution of θ values per radiation) and the angles between vectors from different radiations (one distribution of θ values per between-radiation comparison ($N = 6$); for example, one distribution of θ values comprising angles between each vector of Iceland versus each vector of Scotland). We then compared radiations pairwise, asking whether the distribution of angles between vectors from different radiations differed significantly from the distribution of angles calculated within a radiation (for example, comparing whether the distribution of θ values between Scottish vectors was significantly different from the distribution of θ values between Scottish and Icelandic vectors). We performed this analysis for all pairwise comparisons, comparing the between-radiation distribution with both within-radiation distributions separately (Fig. 2c,d). The analysis was performed separately for armour, shape and gill raker variables. We lacked data for Scottish marine gill rakers, so these comparisons were not possible.

Comparing relative influences of environment, phenotype and genetics. Values of F_{ST} were calculated between each freshwater population and its relevant marine population (Alaska, MUD1; BC, LICA; Iceland, NYPS; Scotland, OBSM) in 50-kb windows using the R package PopGenome[94]. For each M × F comparison, windows above the 95% quantile were classed as outliers. Outlier windows were compared across all pairwise freshwater comparisons (2,628 comparisons among 73 populations), with overlapping outliers representing M × F F_{ST} parallelism. Dissimilarity matrices of environment and phenotype were calculated as Euclidean distances in PCA space for the 7 environmental and 12 phenotypic variables. The genetic dissimilarity matrix was composed of genome-wide pairwise F_{ST} estimates between freshwater populations. The matrix of M × F parallelism was associated to environmental, phenotypic and genetic dissimilarity matrices using Mantel tests (Spearman's) with 9,999 permutations. Partial Mantel tests were performed with genetic distance as the conditional matrix for environmental and phenotypic effects on M × F parallelism, again with 9,999 permutations.

Ethical compliance. Ethical approval for sampling in the United Kingdom was under Home Office licence no. 40/3486, in BC under Professor Dolph Schluter's UBC animal care certificate no. A11-0402 and in Alaska under University of Alaska Anchorage IACUC protocol no. 739596-1. No ethical approval was required in Iceland.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

BAM files of the aligned reads for each individual and corresponding sample information have been deposited in the European Nucleotide Archive database under the project PRJEB20851, with the sample accession numbers ERS1831811–ERS1833111 and run accession numbers ERR2055459–ERR2056759.

Code availability

The scripts used for all analyses are archived through Github/Zenodo (https://doi.org/10.5281/zenodo.4024117).

Received: 4 December 2019; Accepted: 5 October 2020;
Published online: 30 November 2020

References

1. Schluter, D. *The Ecology of Adaptive Radiations* (Oxford Univ. Press, 2000).
2. Gavrilets, S. & Losos, J. B. Adaptive radiation: contrasting theory with data. *Science* **323**, 732–737 (2009).
3. Arnold, S. J., Bürger, R., Hohenlohe, P. A., Ajie, B. C. & Jones, A. G. Understanding the evolution and stability of the *G*-matrix. *Evolution* **62**, 2451–2461 (2008).
4. Losos, J. B. Adaptive radiation, ecological opportunity, and evolutionary determinism: American Society of Naturalists E. O. Wilson award address. *Am. Nat.* **175**, 623–639 (2010).
5. Elmer, K. R. et al. Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nat. Commun.* **5**, 5168 (2014).
6. Mahler, D. L., Ingram, T., Revell, L. J. & Losos, J. B. Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* **341**, 292–295 (2013).
7. Lamichhaney, S. et al. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**, 371–375 (2015).
8. Gould, S. J. *Wonderful Life: The Burgess Shale and the Nature of History* (W. W. Norton, 1989).
9. Schluter, D. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766–1774 (1996).
10. Roff, D. The evolution of the *G* matrix: selection or drift? *Heredity* **84**, 135–142 (2000).
11. Arendt, J. & Reznick, D. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* **23**, 26–32 (2008).
12. Stuart, Y. E. Divergent uses of 'parallel evolution' during the history of the American naturalist. *Am. Nat.* **193**, 11–19 (2019).
13. Oke, K. B., Rolshausen, G., LeBlond, C. & Hendry, A. P. How parallel is parallel evolution? A comparative analysis in fishes. *Am. Nat.* **190**, 1–16 (2017).
14. McGee, M. D., Neches, R. Y. & Seehausen, O. Evaluating genomic divergence and parallelism in replicate ecomorphs from young and old cichlid adaptive radiations. *Mol. Ecol.* **25**, 260–268 (2016).
15. Soria-Carrasco, V. et al. Stick insect genomes reveal natural selection's role in parallel speciation. *Science* **344**, 738–742 (2014).
16. MacColl, A. D. C. The ecological causes of evolution. *Trends Ecol. Evol.* **26**, 514–522 (2011).
17. Elmer, K. R. & Meyer, A. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol. Evol.* **26**, 298–306 (2011).
18. Conte, G. L., Arnegard, M. E., Peichel, C. L. & Schluter, D. The probability of genetic parallelism and convergence in natural populations. *Proc. R. Soc. B* **279**, 5039–5047 (2012).
19. Jones, F. C. et al. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55–61 (2012).
20. Stuart, Y. E. et al. Contrasting effects of environment and genetics generate a continuum of parallel evolution. *Nat. Ecol. Evol.* **1**, 0158 (2017).
21. Jacobs, A. et al. Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. *PLoS Genet.* **16**, 1008658 (2020).
22. Muschick, M., Indermaur, A. & Salzburger, W. Convergent evolution within an adaptive radiation of cichlid fishes. *Curr. Biol.* **22**, 2362–2368 (2012).
23. Foster, S. & Bell, M. *The Evolutionary Biology of the Threespine Stickleback* (Oxford Univ. Press, 1994).
24. Taylor, E. B. & McPhail, J. D. Historical contingency and ecological determinism interact to prime speciation in sticklebacks, *Gasterosteus*. *Proc. R. Soc. Lond. B* **267**, 2375–2384 (2000).
25. Kaeuffer, R., Peichel, C. L., Bolnick, D. I. & Hendry, A. P. Parallel and nonparallel aspects of ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. *Evolution* **66**, 402–418 (2012).
26. Ravinet, M., Prodöhl, P. A. & Harrod, C. Parallel and nonparallel ecological, morphological and genetic divergence in lake-stream stickleback from a single catchment. *J. Evol. Biol.* **26**, 186–204 (2013).
27. Magalhaes, I. S., D'Agostino, D., Hohenlohe, P. A. & MacColl, A. D. C. The ecology of an adaptive radiation of three-spined stickleback from North Uist, Scotland. *Mol. Ecol.* **25**, 4319–4336 (2016).
28. Colosimo, P. F. et al. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**, 1928–1933 (2005).
29. Jones, F. C. et al. A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Curr. Biol.* **22**, 83–90 (2012).
30. Raeymaekers, J. A. M. et al. Adaptive and non-adaptive divergence in a common landscape. *Nat. Commun.* **8**, 267 (2017).
31. Rennison, D. J., Stuart, Y. E., Bolnick, D. I. & Peichel, C. L. Ecological factors and morphological traits are associated with repeated genomic differentiation between lake and stream stickleback. *Phil. Trans. R. Soc. B* **374**, 20180241 (2019).
32. MacColl, A. D. C. & Aucott, B. Inappropriate analysis does not reveal the ecological causes of evolution of stickleback armour: a critique of Spence et al. 2013. *Ecol. Evol.* **4**, 3509–3513 (2014).
33. Spoljaric, M. A. & Reimchen, T. E. 10,000 years later: evolution of body shape in Haida Gwaii three-spined stickleback. *J. Fish Biol.* **70**, 1484–1503 (2007).
34. De Schamphelaere, K. A. C. et al. Reproductive toxicity of dietary zinc to *Daphnia magna*. *Aquat. Toxicol.* **70**, 233–244 (2004).
35. Martins, C., Jesus, F. T. & Nogueira, A. J. A. The effects of copper and zinc on survival, growth and reproduction of the cladoceran *Daphnia longispina*: introducing new data in an 'old' issue. *Ecotoxicology* **26**, 1157–1169 (2017).
36. Miller, C. T. et al. Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. *Genetics* **197**, 405–420 (2014).
37. Chan, Y. F. et al. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **327**, 302–305 (2010).
38. Thompson, K. A., Osmond, M. M. & Schluter, D. Parallel genetic evolution and speciation from standing variation. *Evol. Lett.* **3**, 129–141 (2019).
39. Nelson, T. C. & Cresko, W. A. Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. *Evol. Lett.* **2**, 9–21 (2018).
40. Paccard, A. et al. Repeatability of adaptive radiation depends on spatial scale: regional versus global replicates of stickleback in lake versus stream habitats. *J. Hered.* **111**, 43–56 (2019).
41. Baldo, L., Riera, J. L., Salzburger, W. & Barluenga, M. Phylogeography and ecological niche shape the cichlid fish gut microbiota in Central American and African lakes. *Front. Microbiol.* **10**, 2372 (2019).
42. Fang, B., Kemppainen, P., Momigliano, P. & Merilä, J. On the causes of geographically heterogeneous parallel evolution in sticklebacks. *Nat. Ecol. Evol.* **4**, 1105–1115 (2020).
43. Mäkinen, H. S. & Merilä, J. Mitochondrial DNA phylogeography of the three-spined stickleback (*Gasterosteus aculeatus*) in Europe—evidence for multiple glacial refugia. *Mol. Phylogenet. Evol.* **46**, 167–182 (2008).
44. Liu, S., Hansen, M. M. & Jacobsen, M. W. Region-wide and ecotype-specific differences in demographic histories of threespine stickleback populations, estimated from whole genome sequences. *Mol. Ecol.* **25**, 5187–5202 (2016).
45. Fang, B., Merilä, J., Ribeiro, F., Alexandre, C. M. & Momigliano, P. Worldwide phylogeny of three-spined sticklebacks. *Mol. Phylogenet. Evol.* **127**, 613–625 (2018).
46. Garduno-Paz, M. V., Couderc, S. & Adams, C. E. Habitat complexity modulates phenotype expression through developmental plasticity in the threespine stickleback. *Biol. J. Linn. Soc.* **100**, 407–413 (2010).

47. Coop, G., Witonsky, D., Di Rienzo, A. & Pritchard, J. K. Using environmental correlations to identify loci underlying local adaptation. *Genetics* **185**, 1411–1423 (2010).
48. Hohenlohe, P. A. et al. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* **6**, e1000862 (2010).
49. Guo, B., DeFaveri, J., Sotelo, G., Nair, A. & Merilä, J. Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biol.* **13**, 19 (2015).
50. Glazer, A. M., Cleves, P. A., Erickson, P. A., Lam, A. Y. & Miller, C. T. Parallel developmental genetic features underlie stickleback gill raker evolution. *EvoDevo* **5**, 19 (2014).
51. Day, T., Pritchard, J. & Schluter, D. A comparison of two sticklebacks. *Evolution* **48**, 1723–1734 (1994).
52. Franchini, P. et al. Genomic architecture of ecologically divergent body shape in a pair of sympatric crater lake cichlid fishes. *Mol. Ecol.* **23**, 1828–1845 (2014).
53. McCairns, R. J. S. & Bernatchez, L. Plasticity and heritability of morphological variation within and between parapatric stickleback demes. *J. Evol. Biol.* **25**, 1097–1112 (2012).
54. Peichel, C. L. & Marques, D. A. The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Phil. Trans. R. Soc. B* **372**, 20150486 (2017).
55. Marques, D. A. et al. Genomics of rapid incipient speciation in sympatric threespine stickleback. *PLoS Genet.* **12**, e1005887 (2016).
56. Burke, M. K., Liti, G. & Long, A. D. Standing genetic variation drives repeatable experimental evolution in outcrossing populations of *Saccharomyces cerevisiae*. *Mol. Biol. Evol.* **31**, 3228–3239 (2014).
57. Kang, L., Aggarwal, D. D., Rashkovetsky, E., Korol, A. B. & Michalak, P. Rapid genomic changes in *Drosophila melanogaster* adapting to desiccation stress in an experimental evolution system. *BMC Genom.* **17**, 233 (2016).
58. Gompert, Z. & Messina, F. J. Genomic evidence that resource-based trade-offs limit host-range expansion in a seed beetle. *Evolution* **70**, 1249–1264 (2016).
59. Berner, D., Moser, D., Roesti, M., Buescher, H. & Salzburger, W. Genetic architecture of skeletal evolution in European lake and stream stickleback. *Evolution* **68**, 1792–1805 (2014).
60. Pease, J. B., Haak, D. C., Hahn, M. W. & Moyle, L. C. Phylogenomics reveals three sources of adaptive variation during a rapid radiation. *PLoS Biol.* **14**, e1002379 (2016).
61. Lowry, D. B. et al. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Mol. Ecol. Resour.* **17**, 142–152 (2017).
62. McKinney, G. J., Larson, W. A., Seeb, L. W. & Seeb, J. E. RADseq provides unprecedented insights into molecular ecology and evolutionary genetics: comment on Breaking RAD by Lowry et al. (2016). *Mol. Ecol. Resour.* **17**, 356–361 (2017).
63. Roesti, M., Kueng, B., Moser, D. & Berner, D. The genomics of ecological vicariance in threespine stickleback fish. *Nat. Commun.* **6**, 8767 (2015).
64. Catchen, J. M. et al. Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. *Mol. Ecol. Resour.* **17**, 362–365 (2017).
65. Roesti, M., Moser, D. & Berner, D. Recombination in the threespine stickleback genome—patterns and consequences. *Mol. Ecol.* **22**, 3014–3027 (2013).
66. Samuk, K. et al. Gene flow and selection interact to promote adaptive divergence in regions of low recombination. *Mol. Ecol.* **26**, 4378–4390 (2017).
67. Meier, J. I., Marques, D. A., Wagner, C. E., Excoffier, L. & Seehausen, O. Genomics of parallel ecological speciation in Lake Victoria cichlids. *Mol. Biol. Evol.* **35**, 1489–1506 (2018).
68. Cruickshank, T. E. & Hahn, M. W. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* **23**, 3133–3157 (2014).
69. Terekhanova, N. V. et al. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. *PLoS Genet.* **10**, e1004696 (2014).
70. Westram, A. M. et al. Clines on the seashore: the genomic architecture underlying rapid divergence in the face of gene flow. *Evol. Lett.* **2**, 297–309 (2018).
71. Shimada, Y., Shikano, T. & Merilä, J. A high incidence of selection on physiologically important genes in the three-spined stickleback, *Gasterosteus aculeatus*. *Mol. Biol. Evol.* **28**, 181–193 (2011).
72. Xie, K. T. et al. DNA fragility in the parallel evolution of pelvic reduction in stickleback fish. *Science* **363**, 81–84 (2019).
73. Henning, F. & Meyer, A. The evolutionary genomics of cichlid fishes: explosive speciation and adaptation in the postgenomic era. *Annu. Rev. Genomics Hum. Genet.* **15**, 417–441 (2014).
74. Kess, T., Galindo, J. & Boulding, E. G. Genomic divergence between Spanish *Littorina saxatilis* ecotypes unravels limited admixture and extensive parallelism associated with population history. *Int. J. Bus. Innov. Res.* **17**, 8311–8327 (2018).
75. Bohutínská, M. et al. Genomic basis of parallel adaptation varies with divergence in *Arabidopsis* and its relatives. Preprint at *BioRxiv* https://doi.org/10.1101/2020.03.24.005397 (2020).
76. Rennison, D. J., Samuk, K., Owens, G. L. & Miller, S. E. Shared patterns of genome-wide differentiation are more strongly predicted by geography than by ecology. *Am. Nat.* **195**, 192–200 (2019).
77. Lucek, K., Sivasundar, A., Roy, D. & Seehausen, O. Repeated and predictable patterns of ecotypic differentiation during a biological invasion: lake–stream divergence in parapatric Swiss stickleback. *J. Evol. Biol.* **26**, 2691–2709 (2013).
78. Berner, D., Roesti, M., Hendry, A. P. & Salzburger, W. Constraints on speciation suggested by comparing lake–stream stickleback divergence across two continents. *Mol. Ecol.* **19**, 4963–4978 (2010).
79. Giles, N. Behavioural effects of the parasite *Schistocephalus solidus* (Cestoda) on an intermediate host, the three-spined stickleback, *Gasterosteus aculeatus* L. *Anim. Behav.* **31**, 1192–1194 (1983).
80. Spence, R., Wootton, R. J., Barber, I., Przybylski, M. & Smith, C. Ecological causes of morphological evolution in the three-spined stickleback. *Ecol. Evol.* **3**, 1717–1726 (2013).
81. Reimchen, T. E. Incidence and intensity of *Cyathocephalus truncatus* and *Schistocephalus solidus* infection in *Gasterosteus aculeatus*. *Can. J. Zool.* **60**, 1091–1095 (1982).
82. MacColl, A. D. C. Parasite burdens differ between sympatric three-spined stickleback species. *Ecography* **32**, 153–160 (2009).
83. Stutz, W. E., Lau, O. L. & Bolnick, D. I. Contrasting patterns of phenotype-dependent parasitism within and among populations of threespine stickleback. *Am. Nat.* **183**, 810–825 (2014).
84. Bassham, S., Catchen, J., Lescak, E., von Hippel, F. A. & Cresko, W. A. Repeated selection of alternatively adapted haplotypes creates sweeping genomic remodeling in stickleback. *Genetics* **209**, 921–939 (2018).
85. Etter, P. D., Preston, J. L., Bassham, S., Cresko, W. A. & Johnson, E. A. Local de novo assembly of RAD paired-end contigs using short sequencing reads. *PLoS ONE* **6**, e18561 (2011).
86. Ali, O. A. et al. Rad capture (Rapture): flexible and efficient sequence-based genotyping. *Genetics* **202**, 389–400 (2016).
87. Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A. & Cresko, W. A. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140 (2013).
88. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
89. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).
90. R Core Team *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2017); https://www.R-project.org/
91. Günther, T. & Coop, G. Robust identification of local adaptation from allele frequencies. *Genetics* **195**, 205–220 (2013).
92. Yeaman, S. et al. Convergent local adaptation to climate in distantly related conifers. *Science* **353**, 1431–1433 (2016).
93. Storey, J. qvalue: Q-value estimation for false discovery rate control. R package version 2.0.0 (2015).
94. Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E. & Lercher, M. J. PopGenome: an efficient Swiss Army knife for population genomic analyses in R. *Mol. Biol. Evol.* **31**, 1929–1936 (2014).

Acknowledgements

We thank S. Robertson, R. Young, A. Rahman, B. Santos, S. Goodacre, P. Halldorsson, B. K. Kristjánsson, D. Schluter, K. Samuk, D. Rennison and S. Miller for help with the sampling and sampling permits. We thank A. Lowe and L. Dean for help with the DNA extractions; C. Wiench, A. Stahlke and S. Hendricks for help making the RAD libraries; and J. Brookfield for discussion of the probability calculations. This work was funded by a NERC grant (no. NE/J02239X/1 to A.D.C.M.), and further support was provided by NIH grant no. P30GM103324.

Author contributions

I.S.M., A.D.C.M. and J.R.W. conceived the project, interpreted the data and wrote the manuscript. I.S.M., D.D. and A.D.C.M. performed the fieldwork. I.S.M., M.M. and D.D. generated the phenotypic data. I.S.M. and P.A.H. generated the RAD data, and J.R.W., I.S.M. and P.A.H. analysed it. P.A.H., M.A.B. and S.S. helped with the sampling and revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-020-01341-8.

Correspondence and requests for materials should be addressed to I.S.M. or J.R.W.

Peer review information Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Isabel S. Magalhaes

Last updated by author(s): Sep 11, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	TPS software was used for collection of landmark data for body shape . All data collection is specified and detailed in the methods and SI section.
Data analysis	Software used for data analyses: MORPHOJ 1.03 , Stacks – 1.35 , R , Bayenv2, GSnap, GENEPOP, PLINK, APE, PGDSpider2. Scripts used for all analyses are archived through Github/Zenodo (DOI:10.5281/zenodo.4024117).. This is also specified and detailed in the SI section and the Methods section.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Bam files of aligned reads for each individual and corresponding sample information have been deposited in the European Nucleotide Archive database under the project PRJEB20851, with the sample accession numbers ERS1831811-ERS1833111, and run accession numbers ERR2055459-ERR2056759. Scripts used for all analyses are archived through Github/Zenodo (DOI:10.5281/zenodo.4024117).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We combined large-scale biological sampling and phenotyping with RAD-sequencing data from 73 freshwater populations (1,300 fish) to associate genome-wide allele frequencies with continuous distributions of environmental and phenotypic variation. Overall, this study validates the relative influences of environment, phenotype and genetic contingency on repeatable signatures of adaptation in the genome.
Research sample	We collected 3-spined stickleback (Gasterosteus aculeatus) from 73 freshwater populations and 4 marine populations. Lake names, geographic coordinates and numbers of fish used in the study are shown in Supplementary Table 1.
Sampling strategy	We aimed at sampling most (if not all) of the important phenotypic and environmental variation in all radiations. Previous to sampling we used literature and our collaborators knowledge to identify lakes and populations with the highest variance in environmental and phenotypic variables.
Data collection	I.S.M, D.D., and A.D.C.M performed field work. I.S.M, M.M. and D.D. generated the phenotypic data. I.S.M. and P.H. generated RAD data. Details of how data was collected can be found in the Supplementary Information.
Timing and spatial scale	Data collection was done between 2013-2015 and analysis was performed during 2016-2020.
Data exclusions	No data were excluded.
Reproducibility	All datasets underlying our analyses are provided in the data supplement or submitted to public archives to facilitate future analyses.
Randomization	Not directly relevant to environmental, morphological or population genomics analyses
Blinding	Not directly relevant to environmental, morphological or population genomics analyses

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	We measured the pH, concentrations of metallic cation concentrations sodium (“Na”), calcium (“Ca”) and zinc (“Zn”) of each lake. Concentrations of cations, pH and parasite prevalence per lake are shown in Supplementary Table 2. Details of collection and quantification of abiotic and abiotic variables can be found in Supplementary Information.
Location	We sampled 18 freshwater lakes and a marine site in North Uist, Scotland between April and June 2013, 18 freshwater lakes and a marine site from Iceland between May and June 2014, 18 freshwater lakes and a marine population from British Columbia (BC) between April and May 2015 and 19 freshwater lakes and a marine population from the Cook Inlet basin, Alaska in June 2015. Lake names, geographic coordinates and numbers of samples used in the study are shown in Supplementary Table 1.
Access & import/export	Fieldwork in has been conducted in collaboration with a different institution in each country and in compliance with appropriate national laws of each country and using permits obtained by our collaborators (Prof. Skuli Skulason in Iceland, Professor Dolph Schluter in British Columbia, and Professor Michael Bell in Alaska).
Disturbance	No disturbance was caused by the study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Not used, all individuals analysed were collected in the field.
Wild animals	Fish were collected by setting between 10 and 30 unbaited minnow traps (Gee traps, Dynamic Aqua, Vancouver, Canada) in water approximately 0.3–3m deep, within 5 m of shore along a 100–400m stretch of shoreline. The fish were haphazardly selected with individuals of all sizes, sex and breeding condition collected (see Supplementary Table S17 for details of sex of each fish collected).
Field-collected samples	3-spined stickleback were transported to the lab and immediately humanely killed the fish by overdose with the anesthetic MS222.
Ethics oversight	Ethical approval for sampling in the UK was under Home Office licence 40/3486, in British Columbia under Professor Dolph Schluter's UBC animal care certificate A11-0402, and in Alaska under University of Alaska Anchorage IACUC protocol 739596-1. No ethical approval was required in Iceland.

Note that full information on the approval of the study protocol must also be provided in the manuscript.