Intercontinental genomic parallelism in multiple adaptive radiations

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

[1] *School of Life Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, U.K.*
[2] *Department of Life Sciences, Whitelands College, University of Roehampton, London SW15 4JD, U.K.*
[3] *Biosciences, College of Life and Environmental Sciences, Geoffrey Pope, University of Exeter, Stocker Road, EX4 4QD, U.K.*
[4] *Institute for Bioinformatics and Evolutionary Studies, Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA.*
[5] *Department of Ecology & Evolution, Stony Brook University, 650 Life Sciences Building, Stony Brook, NY 11794-5245*
[6] *Department of Aquaculture and Fish Biology, Hólar University College, Saudárkrókur, Iceland*

*These authors contributed equally to this work

Abstract

Parallelism, the evolution of similar traits in populations diversifying in similar conditions, provides good evidence of adaptation by natural selection. Many studies of parallelism have focused on comparisons of strongly different ecotypes or sharply contrasting environments, defined *a priori*, which could upwardly bias the apparent prevalence of parallelism. Here, we estimated genomic parallelism associated with individual components of environmental and phenotypic variation at an intercontinental scale across four adaptive radiations of the three-spined stickleback (*Gasterosteus aculeatus*), by associating genome-wide allele frequencies with continuous distributions of environmental and phenotypic variation. We found that genomic parallelism was well predicted by parallelism of phenotype-environment associations, suggesting that a quantitative characterization of phenotypes and environments can provide a good prediction of expected genomic parallelism. Further, we examined the explanatory power of genetic, phenotypic, and environmental similarity in predicting parallelism. We found that parallelism tended to be greater for geographically

proximate, genetically similar radiations, highlighting the significant contingency of standing variation in the early stages of adaptive radiations, before new mutations accumulate. However, we also demonstrate that distance within multivariate environmental space predicts parallelism, after correction for genetic distance. This study thus demonstrates the relative influences of environment, phenotype and genetic contingency on repeatable signatures of adaptation in the genome.

Introduction

Adaptive radiations are well known as rapid branchings on the tree of life ([1]).They may be the source of most biodiversity, and their study has revealed a great deal about the evolution of phenotypic diversity ([1,2]). However, patterns in adaptive radiations also highlight some of the unknowns about how biodiversity evolves. For example, although adaptive radiations are typified by abundant phenotypic diversity, not all trait combinations evolve in every radiation, while on the other hand organisms in different places sometimes arrive at very similar endpoints ([3,4]). This suggests that Stephen Jay Gould's famous contention that evolution is contingent and unrepeatable ([5]) cannot be completely true. A good deal of work has focussed on the role that genetic correlations between traits might play in creating constraints on diversity, but so far the answers provided by this approach have not been entirely satisfactory ([3,6,7]). Alongside these processes, it is probable evolution can be shaped in a predictable way by common environments and shared selective regimes, and that the emergence of repeatable patterns of evolution, a process predominantly known as parallelism (which we distinguish from convergence here by the inclusion of shared evolutionary 'start', as well as 'end', points, although see ([8,9]), is the direct result of environmental similarities within and between radiations. Striking examples of phenotypic

parallelism ([1,10]) in the natural world support this hypothesis, and the persistent appearance of familiar forms in similar ecological niches demonstrates the significance of selection in this process.

However, a significant problem with studies that have focused on phenotypic parallelism is that they have concentrated largely on the comparison of pairs of strongly different ecotypes or widely different environments ([11–13]). Such studies could upwardly bias the apparent prevalence of parallelism because the chosen comparisons were known *a priori* to occur repeatedly in different locations, effectively constraining the evolutionary end point. The approach also conceals the role of individual components of environmental variation in driving parallelism, as similar environments are often assumed based on comparable phenotypes. This gap surely needs addressing for a complete understanding of adaptation ([14]). In addition, traits that are measurable in sufficient numbers from (usually) wild organisms are generally limited to morphological and life history traits ([13]). This seriously compromises our ability to understand adaptation, a great deal of which is likely to be physiological in the broad sense. Such drawbacks highlight the importance of combining measures of phenotype and environment alongside genomics in studies of parallelism. The detection of consistent genomic signatures across multiple, independent natural populations has proven to be a valuable tool for studies of evolutionary patterns and the discovery of genes involved in adaptation ([15–17]). However, yet again our comprehension of the relationship between genomic parallelism and continuous phenotypic or environmental variation is surprisingly poor.

Until recently, a major barrier to combining genotype, environment and phenotype has been the high costs of sequencing, prohibiting large-scale genomic sampling that could be aligned with large-scale ecological sampling. However with dramatic drops in DNA

82 sequencing costs, such genomic data can nowadays be used alongside ecological data to
83 associate allele frequencies ([18]), and determine genomic regions associated with individual
84 components of environment and phenotypes. Such an approach is essential if studies of
85 parallelism are to shift from description to hypothesis testing, but has rarely been applied
86 (but see [12, 19, 20]), and it remains to be shown whether signals of parallelism obtained from
87 continuous measures are comparable to those from ecotypes and from previous studies.
88 Here we make use of such methods to test for environmentally and phenotypically
89 associated genomic parallelism across radiations of three-spined stickleback fish
90 (*Gasterosteus aculeatus*, hereafter 'stickleback').

91 Stickleback provide a powerful natural experiment to test parallelism. They are
92 primitively marine but have undergone replicated adaptive radiations across the northern
93 hemisphere following colonisation of freshwater in widely separated geographical locations.
94 This allows us to compare multiple populations derived from the marine ancestor, and
95 provides a model for exploring both phenotypic parallelism and its genetic basis ([21,22]) in
96 response to environmental variation. Phenotypic parallelism in populations that have
97 evolved independently in similar habitats is well established ([23,24]), and whilst often
98 considered in dichotomous pairings of marine – freshwater, benthic – limnetic, lake -stream
99 ecotypes, there is a huge amount of continuous phenotypic variation among freshwater
100 populations that has hitherto rarely been explored ([25]) in this context. In addition, genome
101 scans have identified loci that have come repeatedly under selection across the contrasting
102 ecotype pairs ([17,26,27]), but the combination of the phenotypic and genomic parallelism in
103 one study is rare (although see [28]) and has not previously been done at the scale of
104 replicated adaptive radiations across continents.

For this study we sampled 73 freshwater lake populations from four adaptive radiations in Alaska, British Columbia ('BC'), Iceland and the island of North Uist ('Scotland') (Fig. 1, Supplementary Table 1), measured a set of six biotic and abiotic environmental variables and a set of 12 phenotypic traits (measures of body shape, armour traits and gill rakers) (Supplementary Table 2), and performed a genome-wide scan of a total of 1,304 individual stickleback using restriction-site associated DNA (RAD) sequencing. We first assessed environmental and phenotypic parallelism among radiations, and their phylogenetic relationship, which provided a reference for how much genomic parallelism to expect. We then scanned the genome for associations with continuous environmental and phenotypic axes of variation within radiations, and identified genomic parallelism by looking at the presence of allele frequency associations in the same genomic regions across radiations. As marine-freshwater parallelism is well-documented ([17,27,29]), we compared our results for parallelism across freshwater radiations with well-studied marine-freshwater parallelism in this species, and used the results as a positive control for the methods used. Finally, we examined how the prevalence of parallel genomic regions is associated with the phylogenetic histories of our adaptive radiations and how well it can be explained by multivariate quantification of environmental and phenotypic similarity between radiations. 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.

Environmental and phenotypic similarity across radiations

129 Interpreting inter-radiation patterns of genomic parallelism associated with phenotypic
130 traits and environmental variables requires that we first assess the environmental and
131 phenotypic similarity across all four radiations analysed here, as well as the relationship
132 between the two types of variables. Environmental similarity is predicted to influence both
133 phenotypic and genomic parallelism, as similar selection regimes are imposed on organisms
134 ([13,30,31]). These analyses provide insights into some major factors that are likely to influence
135 the emergence of a common pattern of genomic divergence at such a large geographical
136 scale, and indicate how much genomic parallelism associated with environments and
137 phenotypes to expect.

138 *I) Environment.* Adaptation of stickleback to freshwater in these radiations has
139 happened rapidly, over the course of the past few thousand years, and has most likely been
140 driven by strong selection favouring pre-existing alleles present at low frequencies in the
141 ancestral marine population ([32]). Therefore, we might expect to observe shared features of
142 diversification when multiple groups of organisms experience similar environmental
143 circumstances, as similar selection regimes are imposed on organisms ([12,33]). A Principal
144 Component Analysis (PCA) on six measures of water chemistry (pH, calcium (Ca), sodium
145 (Na) and zinc (Zn), and parasitism (prevalence of the parasites *Gyrodactylus spp.* (Gyro) and
146 *Schistochephalus solidus* (Schisto)) across all lakes revealed that the first axis of
147 environmental variance (Env_{PC1}) separated lakes along a gradient of pH and calcium
148 concentration (Fig. 2; Supplementary Table 3). This axis did not separate radiations but
149 emphasised the variation from alkaline to acid present in all of them, with the widest range
150 in Alaska. The second axis of environmental variance (Env_{PC2}) distinguished mostly between
151 lakes with high and low zinc and largely separated the European from the North American

lakes. The two groups of European lakes overlapped very little environmentally with each other along the second axis or with the North American lakes, but the variation in BC lakes was completely subsumed within that of the more environmentally variable Alaskan lakes.

An analysis of the similarity of the direction of the major PC vectors (θ) of environmental variation (PCs that explained at least 10% of the variation) and of the magnitude of their variation revealed that Alaska and BC exhibited the strongest environmental similarity in terms of direction ($\theta = 8.71°$, $p = 0.077$) (Supplementary Table 4), as lower values of θ are indicative of more common environmental covariance. However, they also showed the greatest difference in magnitudes of variation, as BC had the lowest amount of environmental variation of all radiations while Alaska had the highest. Iceland on the other hand was the most unique radiation in terms of environmental covariance, a pattern likely driven by Iceland's uniquely volcanic nature.

II) *Phenotypes - Armour.* The major axis of armour variation ($Armour_{PC1}$) represented a correlated axis of all traits, particularly dorsal spine length and pelvic characteristics (Fig. 2; Supplementary Table 3). Although there were significant differences in $Armour_{PC1}$ among both radiations and lakes (Supplementary Table 5), the direction of this vector was highly conserved across all radiations ($4.13° \leq \theta \leq 8.37°$; $0.004 \leq p \leq 0.053$) (Supplementary Table 4). Nonetheless, we found populations with extreme armour, such as absence or extreme reduction of dorsal spines, pelvic spines, pelvis size, and generally low $Armour_{PC1}$, that occur in Scotland and Alaska but not elsewhere. However, while in Scotland pelvic armour reductions were accompanied by complete or almost complete loss of armour plates (high $Armour_{PC2}$), Alaskan populations retained those even when other features of armour were highly reduced. These deviations produce the somewhat anomalous relationship between $Armour_{PC1}$ and $Armour_{PC2}$ in Alaska (Fig. 2). Iceland exhibited particularly low variation in

armour traits compared to other radiations, while BC had the highest. These results indicate that whilst the direction of armour variation occurred predominantly on a shared axis across adaptive radiations, the amount of armour variation along that axis was variable. When comparing all freshwater populations with four marine ones (one from each country), we found that, aside from a few populations from BC, there was no overlap in armour traits between marine and freshwater populations. This is because marine populations have a higher number of lateral plates and more exaggerated armour traits in general. Importantly however, projection of marine armour phenotypes suggests they fall on the same axis but beyond freshwater space (Fig. 2).

III) *Phenotypes -Body shape.* Despite some overlap in body shape across all radiations suggesting similar morphologies have evolved repeatedly across continents (Fig. 2), there were significant differences in body shape morphologies among both radiations and lakes (Supplementary Table 5). The most extreme body shapes were found in Scotland, where some populations have very elongated and slender bodies with small heads ($Shape_{PC1}$ axis), and in BC where some populations had the deepest bodies and heads and the longest heads ($Shape_{PC2}$ axis). Scotland had the largest amount of variation in body shape, which is a surprising result given this variation is found within 1000 km^2 of North Uist, a much smaller area than the other three radiations. The orientation of the major PC vectors (θ) revealed that North American radiations were the most similar (θ = 3.49°, p=0.19), followed by the European radiations (θ = 8.48°, p=0.63) (Supplementary Table 4), which highlights a potential effect of geography/shared ancestry on the major axis of shape variation. Marine populations overlapped in body shape with some of the freshwater populations that had lower ShapePC1 and ShapePC2 scores, most of them found in Iceland (Fig. 2).

199 *IV*) *Phenotypes - Trophic morphology*. The largest variation in gill raker numbers and
200 length was found in Alaska, while BC populations contained a subset of that variation (Fig.
201 2). Across the two dimensions of gill raker number and length, variation within radiations
202 was generally constrained to a major axis in which populations with more gill rakers also
203 tended to have longer gill rakers. All freshwater populations generally had shorter gill rakers
204 for their size relative to marine populations.

205 *V) Relationship between environmental and phenotypic similarity*. We used GLMs to
206 test for parallel associations between the first two environmental PCs (Env_{PC1} and Env_{PC2})
207 and armour, shape, and gill raker morphology across radiations. We found that $Armour_{PC1}$
208 and $Armour_{PC2}$ were also significantly associated with Env_{PC1} ($Armour_{PC1}$: F=6.13, p=0.016;
209 ArmourPC2: F=4.11, p=0.047), but not with Env_{PC2} ($Armour_{PC1}$: F= 1.12, p=0.293; $Armour_{PC2}$:
210 F=1.10, p=0.298) (Supplementary Table 6). $Armour_{PC1}$ correlations with Env_{PC1} were similar
211 across radiations but significant slope variation was observed between radiations for
212 $Armour_{PC2}$ and Env_{PC1} associations (F= 2.76, p=0.050). The result highlights parallelism in the
213 way general skeletal traits ($Armour_{PC1}$) are reduced with lower pH and calcium (Env_{PC1}), but
214 non-parallel associations between these measures and the relationship between pelvic size
215 and plate number ($Armour_{PC2}$). The latter is consistent with previous findings of a lack of
216 convergence in plate number in response to Ca between North American and European
217 radiations ([34]), while the former suggests that common pH and calcium environments,
218 experienced in all radiations as a shared acid-alkali axis, promotes phenotypic parallelism.

219 $Shape_{PC1}$ and $Shape_{PC2}$ were significantly associated with both Env_{PC1} ($Shape_{PC1}$: F=
220 4.85, p=0.031; $Shape_{PC2}$: F=28.20, p<0.001) and Env_{PC2} ($Shape_{PC1}$: F= 7.56, p=0.008; $Shape_{PC2}$:
221 F=48.01, p<0.001), suggesting body shape variation is strongly linked to specific measures of
222 water chemistry. Intercepts across radiations varied significantly but slopes did not,

indicating parallelism across radiations in the way shape and environment are associated (Supplementary Table 6). The result highlights parallelism in the way bodies become more elongated and slender and heads are reduced with lower pH and calcium and higher Zinc.

As for gill rakers, we expected trophic morphology to evolve in response to zooplankton communities, which in turn have been shown to vary in response to measures of water chemistry ([35,36]), in particular Zn which was the dominant component of Env_{PC2}. Consistent with this, we found gill raker number to be strongly significantly associated with both Env_{PC2} ($F=43.74$, $p<0.001$) and Env_{PC1} ($F=18.17$, $p<0.001$). Slopes varied between radiations for the association of gill raker number and Env_{PC1}, demonstrating non-parallelism in trophic response to pH and calcium. In contrast, common slopes across radiations highlight parallelism in trophic association between gill raker number and Zn variation (Env_{PC2}).

Corroborating predictions that selection imposed by similar environments can lead to phenotypic parallelism ([1]), we found that similar variation from alkaline to acid environments (Env_{PC1}) present in all radiations was significantly associated with parallel variation in body shape and armour. If phenotypic adaptation happens via natural selection on genetic variation, we would then expect these associations to translate into genomic parallelism associated with both environmental variables that show a similar range across radiations (pH and Ca) and phenotypic traits associated with them.

Phylogenetic relationship among radiations.

The probability of genomic parallelism has been linked to time since lineages split, highlighting the probable influence of genetic similarity and shared ancestral variants on

parallelism ([16]), it is therefore important that we know the genetic relationship across all populations before interpreting patterns of genomic parallelism. Genomic parallelism across populations within any one radiation is likely to evolve from shared genetic variation, but the probability of that should decline in geographically distant radiations as a function of common ancestry ([16,37]). Indeed, recent studies have highlighted the probable limitation of parallelism in this system across continental scales ([20,38,39]). A Neighbour Joining (NJ) tree based on 8,395 unlinked SNPs showed that, with the exception of a likely recently formed population in Alaska (TERN), which is basal to both the Alaskan and British Columbian radiations, the four geographic locations form four well-resolved radiations (Fig. 1). A PCA on the same dataset confirmed that radiations form independent clusters, and also revealed that the dominant axis of variation (PC1 = 36.0%) separates North American and European radiations pairs from one another (Supplementary Fig. 1). North American radiations then separated on PC2 (7.0%) and European radiations on PC3 (5.8%). In addition to being close together in the NJ tree, we found that the geographically adjacent radiations were also the most genetically similar (Supplementary Table 7): Alaska and BC (mean pairwise F_{ST} = 0.198), and Scotland and Iceland (F_{ST} = 0.194), suggesting that although these form independent clusters, the lineage split between them is relatively recent, or that any gene flow between radiations is occurring within the Atlantic and Pacific groups. Genetic divergence between inter-continental pairings was found to be stronger and deeper (0.314 ≤ F_{ST} ≤ 0.338) than within continents, which is consistent with previous studies that have estimated the time of divergence between stickleback from Europe and North America to be approximately 200,000 years ([40,41] but see [42]). Further, between-continent structuring accounts for the largest proportion of molecular variance in our data (AMOVA: σ = 889.7, 34.7%). Within continents, populations within radiations (σ = 366.5, 14.3%) were more genetically variable

than radiations (σ = 142.3, 5.6%) (Supplementary Table 8). This highlights that molecular variance isn't structuring according to geographic scale (Continent > Radiation > Populations), but rather gene flow may be occurring between intra-continental radiations.

Following the idea that the probability of parallelism at the genetic level is linked to time since lineages split (i.e. genetic similarity), we would expect groups with closer evolutionary histories, or indeed present gene flow, to show the strongest genomic parallelism if adaptation occurs through shared standing genetic variation or the exchange of beneficial freshwater alleles between intra-continental radiations via marine populations ([29,32]).

Phenotypically and environmentally associated SNPs and genomic regions within radiations.

We scanned the genome to identify regions associated with individual components of phenotypic and environmental variation within each of the four adaptive radiations. 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. For each radiation separately (N=18 or 19 populations) we used Bayesian linear models implemented in Bayenv2 ([43]) to identify associations between population SNP allele frequencies (N=10 to 21 individual fish, mean= 17.8), the set of six biotic and abiotic environmental variables and the set of 12 phenotypic traits (Supplementary Table 2). 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 ([17,27,29]), and then examined genomic differentiation between all freshwater populations

294 pooled within a radiation and four marine populations pooled together (one from each
295 country, Supplementary Table 1).
296 Our analyses identified population allele frequencies of several thousand SNPs for
297 each radiation as being highly correlated (high Bayesfactor [>$\log_{10}(1.5)$] and top 5% of
298 Spearman's ρ, see methods) with the abiotic and biotic environmental variables
299 ('environmentally associated SNPs') or with phenotypic traits ('phenotypically associated
300 SNPs' (Supplementary Table 9). SNPs were then mapped onto non-overlapping sliding
301 windows of 50kb, 75kb, 100kb or 200kb, which allowed us to test the robustness of our
302 results across different extents of linkage. Further, we repeated our analysis across windows
303 of equivalent genetic distance (0.1 cM windows), which confirmed that our results were not
304 influenced by variable linkage across the genome (Supporting Information). Here we report
305 only results for 50kb windows, given that this is consistent with approximate linkage
306 disequilibrium within the stickleback genome ([44,45]). Results for window sizes of 75kb,
307 100kb, 200kb can be found in Supplementary Dataset 1. Our 50kb dataset was composed of
308 4,868 windows with SNPs in all locations, covering approximately 55% of the 447 Mb
309 genome, with a further 1,940 windows sequenced in 2 or more locations providing
310 information on an additional 21.7% of the genome.
311 Windows were classified as 'environmentally associated' and 'phenotypically
312 associated' if they contained more associated SNPs than expected under a 99% binomial
313 expectation ([46]). We found 1791 unique 50kb windows associated with an environmental
314 variable or a phenotypic trait (Supplementary Dataset 1), ranging from 146 windows
315 associated with pelvic spine length in BC, to 21 associated with Na variation, also in BC. It is
316 striking how many windows show strong signals of association with phenotypic and

environmental variables, even when their variation is modest, clearly supporting the adaptive nature of these radiations.

Out of all the unique windows, 431 were associated with both an environmental and phenotypic variable in the same radiation, suggesting several regions associated with phenotypic traits might also be responsible for local adaptation to environments (Supplementary Dataset 1). It also suggests that directly measuring important aspects of the environment may provide profitable ways of discerning the genomic basis of adaptation and of identifying agents of selection, while bypassing the often-difficult measurement of phenotypes.

Genomic parallelism associated with environmental and phenotypic variation across radiations.

To quantify genomic parallelism we identified environmentally- and phenotypically-associated windows that were shared across two or more of our radiations, i.e. parallel windows (Supplementary Fig. 2). We then compared the overall observed numbers of parallel windows to a null distribution of randomly associated windows permuted over 10,000 iterations.

We quantified genomic parallelism at three levels: 1) gross, general levels of parallelism associated across all radiations for groups of phenotypes or environment; 2) genomic parallelism associated with individual variables across all comparisons to understand contributions of individual variables; 3) parallelism for individual variables in specific pairings, to identify pairs of radiations having the highest levels of parallelism, and for which variables. The latter is important as patterns of parallelism may be lost when we

pool more than two radiations or variables together, if for example parallelism is very strong in one radiation-pairing but not others.

When quantifying the overall level of genomic parallelism associated with groups of environmental or phenotypic variables we found no environmentally- or phenotypically-associated 50kb windows parallel in all four radiations for individual variables (Randomised permutations $N_{Expected\text{-}Environmental}$= 0.0002, $N_{Exp\text{-}Pheno}$ = 0.0001), but one window was parallel in a group of three radiations (chrIV: 14400000-14450000 associated with length of pelvis in BC, Iceland and Scotland) ($N_{Exp\text{-}Env}$ = 0.05, $N_{Exp\text{-}Pheno}$ = 0.161, p = 0.149). Many windows however exhibited parallelism between pairs of two radiations. A total of 39 environmentally associated windows (pooled across all 6 environmental variables) (N_{Exp} = 11.9, 95% Upper limit (UL) = 18, $p < 0.001$) and 65 phenotypically associated windows (pooled across all 12 phenotypic variables) (N_{Exp} = 30.9, 95% UL = 40, $p < 0.001$) were parallel between two radiations (Supplementary Fig. 2). Parallelism was disproportionately greater for armour and gill raker traits (number mostly) than for shape, with the 65 phenotypically associated windows split into 46, 12 and seven associated windows for armour, gill raker and shape variables respectively (χ^2= 6.506, P = 0.04). This is consistent with the fact that skeletal traits, several of which are known to have simple genetic architectures ([26,47,48]), are particularly likely to show evidence of phenotypic parallelism. Interestingly, parallel associated windows (mean SNP N = 7.13) had on average more SNPs per window than non-parallel windows (mean SNP N = 6.27) (GLM, $LRT_{1,3867}$ = 22.1, $p <$ 0.001) and exhibited slightly stronger signals of association with variables (mean residual SNPs above expected = 1.82 parallel; 1.61 non-parallel; GLM, $LRT_{1,3867}$ = 5.05, p = 0.025).

Random permutations indicated that there were statistically significant levels of parallelism for the number of windows ('significantly parallel windows') associated with two

364 environmental variables (Ca and pH) (Fig. 3; Supplementary Table 10). Consistent with our
365 expectations from patterns of environmental parallelism, these were the same variables
366 that share an axis of variation (Env_{PC1}) across freshwater environments in all radiations. In
367 addition, we did not detect significant genomic parallelism associated with variables that
368 exhibit variation between radiations, such as salinity, zinc and *S. solidus* prevalence. These
369 results together confirm that common environmental axes, such as the one experienced in
370 all radiations as a shared acid-alkali axis, likely promote signals of parallelism in the genome.
371 We also found more genomic parallelism than expected by chance associated with a
372 total of five phenotypic variables (Fig. 3; Supplementary Table 10): four armour traits (2nd
373 dorsal spine, pelvic spine length, length of pelvis and armour plate number) and gill raker
374 number. These results are consistent with the high heritability of skeletal traits ([26,47,48]), and
375 suggest that variation in armour is the result of different genotypes being selected in
376 different environments. Further, parallel QTLs have been described for gill raker number
377 ([49]), but not length, which exhibits more plasticity ([50]). Body shape traits were not associated
378 with any significant genomic parallelism, despite parallelism across radiations in the way
379 shape is associated with the environment. It is probable the partly plastic nature of body
380 shape ([51,52]) leads to an association between environment and body shape via the reaction
381 norm rather than genomic re-use.
382 Our analyses also showed much stronger parallelism across marine-freshwater (MxF)
383 comparisons than across freshwater variables (Fig. 3), and 44/158 of MxF associated
384 windows overlapped with previously identified genomic regions contributing to marine-
385 freshwater divergence ([17,44]) (Supplementary Table 11). Several regions found parallel for
386 MxF analyses were also parallel for Ca, pH, Na, armour traits and gill raker number
387 (Supplementary Table 12). These results suggest our methods and sequencing coverage are

388 robust enough to recover known parallel regions, and interestingly that genomic parallelism
389 associated with freshwater variables is more modest than marine-freshwater parallelism.
390 The latter likely reflects subtler variation between habitats within radiations in comparison
391 to stark marine-freshwater contrasts. Further, these results highlight that binary ecotype
392 pairings, which likely include variation in many environmental and phenotypic traits, lump
393 together parallelism of many components of fitness without being able to discern which are
394 parallel and which are non-parallel. Overall, our results suggest that across these four
395 freshwater adaptive radiations, evolution of these phenotypes and environmentally
396 associated traits are disproportionately linked to the same genomic regions.

397 Within specific pairings, we found the greatest number of significantly parallel
398 windows in the comparison between Alaska and BC (two environmental variables: Ca and
399 *Gyrodactylus spp.*, and three phenotypic traits: pelvic spine length, plate number and gill
400 raker number), followed by Iceland and Scotland (two environmental variables: Ca and pH,
401 and one phenotypic trait: dorsal spine length) (Supplementary Fig. 2; Supplementary Table
402 10). Two environmental variables and one phenotypic trait were also associated with
403 significantly more parallel windows in comparisons between Iceland and the North
404 American radiations: Ca and length of the pelvis (Alaska and Iceland) and *Schistocephalus*
405 *solidus* (BC and Iceland). Pelvic spine length was the only variable of any category found to
406 be parallel between Scotland and BC, and no significant parallelism was detected between
407 Scotland and Alaska (suggesting observed overlap may be the result of chance).
408 Phylogenetic patterns and the segregation of molecular variance strongly support the
409 notion that radiations within continents share similar genetic variation, making parallelism
410 through shared standing variation the most parsimonious explanation for our intra-

411 continental parallelism biases. Experimental studies with stickleback have demonstrated
412 rapid morphological adaptation from standing genetic variation, even recovering diverse
413 morphologies from variation found within phenotypically-derived freshwater populations
414 ([53]). Coancestry patterns, centred at the focal, causative loci, can discern between whether
415 parallel evolution occurs on *de novo* mutations, standing variation, or introgressed alleles,
416 however we lack the sequencing resolution in our current data to make these comparisons.

417

418 **Linkage and the genomic location of parallel regions.**

419 As with any reduced-representation genomic approach, the power of RAD
420 sequencing to detect loci associated with phenotypic or environmental variables depends
421 on linkage disequilibrium (LD) between markers and functional loci([44,54,55]). The scale of LD
422 varies widely across organisms and within genomes, but has been relatively well-
423 characterized in stickleback ([56,57]), and RAD sequencing has been used successfully in this
424 species specifically to test for genomic parallelism ([30,44,53,56,57]). One explanation for variable
425 linkage across the genome is the negative relationship it shares with local recombination
426 rate. To examine its influence on our results, we estimated recombination rate using a
427 previously published genetic map ([31]) for our 50kb windows, and marked windows that were
428 associated with any variable and associated windows that were also parallel across
429 radiations. Recombination was significantly reduced in associated windows and parallel
430 windows compared with non-associated windows (Supplementary Fig. 3; Kruskal-Wallis, χ^2 =
431 121.43, $p < 0.001$), but did not differ significantly between outlier and parallel windows (p =
432 0.55). Reduced recombination can be an important process in adaptation through

maintaining adaptive alleles, as has been demonstrated in this species ([58]) and others (eg. [59]). However, we cannot rule out that these patterns are produced by an increased ability to detect selection in low-recombination windows as a product of increased linkage with causative SNPs. The latter suggests that our estimates of association, and by extension parallelism, may be conservative if false-negatives are pervasive in high recombination regions. Importantly however, our signatures of parallelism cannot be explained by variable recombination.

Windows based on genetic distance (0.1 cM) corroborated 50kb results, returning strong signals of parallelism for calcium, pH, pelvic spine length, pelvis length, plate number and gill raker number. Interestingly, we also recovered weakly significant parallelism for several other environmental and armour variables (Supplementary Fig. 4, Supplementary Information), suggesting potentially stronger parallelism than we report at 50kb windows.

To assess wider, linked parallel regions, we plotted 50kb windows across the genome to examine clustering of all associated windows (Supplementary Fig. 5). Adjacent windows (two or more, Supplementary Table 13) were pooled together to inspect putative causative genes (Supplementary Dataset 2). Using these two methods we identified a number of wider genomic regions that exhibited parallelism across multiple radiations. An example, and good positive control for our method, involves the pooling of windows associated with plate number in three radiations (Alaska, BC and Iceland) on chromosome IV around the well-known *Eda* gene, which has a well-established role in producing variation in the number of armour plates ([17,26,27]). This pattern emerges despite the limited variation in plate number across freshwater populations.

Pooling adjacent windows also identified a large cluster (250 kb) on chromosome I containing genes *igfbp2a*, *stk11ip* and *atp1a1*, and strongly associated with calcium, sodium

457 and pH in several radiations (Supplementary Data 2). The clustering of windows in this
458 region is perhaps unsurprising given it contains a known inversion ([60]), which as discussed
459 should be beneficial for adaptive haplotypes by reducing local recombination ([61]). Within
460 this region specifically, it is likely that the *atpa1a1* gene causes large effects on fitness, given
461 its previously detected association with the major ecological transition from marine to
462 freshwater ([62]) and functional role in metal ion management ([17,60]). Its apparent role in
463 adaptation to much smaller cation variation between freshwater environments in this study
464 is interesting in light of Fisher's geometric model of adaptation, since this predicts that only
465 alleles of smaller effect should fix as a fitness optimum is approached ([63]). Thus, it may be
466 that different mutations in this gene have a spectrum of effect sizes, or that changes cause
467 subtler differences in expression, rather than larger coding differences.

468 Our results support the existence of genomic regions of physically linked genes that
469 are hitch-hiking in separate radiation pairs, and may contain genes that are parallel across
470 all radiations but undetected by our genomic methods. Extensive linkage disequilibrium in
471 freshwater populations is consistent with what is expected under strong directional
472 selection after colonization from marine populations and has been reported for stickleback
473 populations from Alaska ([44]), but it had not previously been observed for the same regions
474 across several independent adaptive radiations. These results are also consistent with
475 previous findings of large numbers of SNPs highly divergent between marine and freshwater
476 stickleback aggregating in just 19 short genomic regions, including three known inversions
477 ([60]), one of which we also detected and highlighted above.
478

Relationships between genomic parallelism and phylogenetic, phenotypic and environmental similarity.

Freshwater populations have radiated from marine common ancestors ([1,16]), thus the parallelism patterns described in this study are putatively the result of multiple marine - freshwater transitions. Based on this assumption, and using our SNP data from a marine population within each radiation, we performed genome-wide F_{ST} outlier analyses comparing each freshwater population with the marine population within that radiation. We then took the top 5% of 50kb windows according to F_{ST} in each marine – freshwater comparison and looked for overlapping outlier windows across all comparisons (N = 2,628). This quantified the extent of repeated genome-wide differentiation for MxF transitions within and across all radiations (Fig. 4a).

We then used Mantel tests to test statistically the effects on MxF genomic parallelism of relative genetic divergence and environmental/phenotypic similarity alongside one another. We compared the matrix of overlapping marine - freshwater outlier F_{ST} windows to equivalent matrices for environmental dissimilarity (Euclidean distances), phenotypic dissimilarity (Euclidean distances), and genetic dissimilarity (F_{ST} values) (Fig. 4). Across all population comparisons the number of parallel windows was strongly negatively correlated with genetic dissimilarity ($r = -0.61$, $p < 0.001$), and also negatively correlated with environmental dissimilarity ($r = -0.42$, $p < 0.001$) and phenotypic dissimilarity ($r = -0.11$, $p = 0.022$). These results highlight that genomic parallelism increases in populations that are more genetically, environmentally and phenotypically similar, though to varying extents.

Parallel windows were more common in intra- rather than inter-continental comparisons. This highlights the significance of our European and North American pairings as the major contributors towards pairwise signals of genomic parallelism, as we discussed previously, and strongly suggests that genomic parallelism at large geographic scales is contingent on shared genetic variation. Further, at an intra-continental scale, there exists the possibility of haplotype sharing between radiations 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 ([64]). Recent research has also highlighted a probable genetic bottleneck in the founding of Atlantic marine populations that restricts shared freshwater alleles between Atlantic and Pacific freshwater populations ([39]).

We also conducted partial mantel tests for the effects of environmental and phenotypic similarity whilst controlling for genetic similarity, given this is likely to correlate with environment and phenotype in some cases due to geographic proximity. Effects of environmental dissimilarity were marginally reduced when controlling for genetic similarity, but were still strongly negative (r = -0.35, $p < 0.001$), suggesting that similar environments promote genomic parallelism irrespective of genetic similarity. Phenotypic dissimilarity was no longer associated with genomic parallelism after controlling for genetic similarity (r = -0.10, p = 0.097). This latter result suggests that environmental similarity is a better predictor of genomic parallelism than phenotypic similarity (at least in terms of observable morphometric phenotypes) in this system.

In conclusion, our study is the largest to date in this system addressing the relative effects of environment, phenotype and genetics in predicting parallel evolution. Genetic

similarity is the best predictor of genomic parallelism here, in line with recent results and expectations regarding sharing of ancestral variants and introgression between populations ([29,39]). However, even whilst controlling for this, environmental variation (most likely Ca and pH variation in particular) is a good predictor of genome-wide parallelism among freshwater populations on a wide geographic scale. In particular, the higher environmental similarity among North American radiations, compared to European populations, provides a good explanation for the strong phenotypic and genomic parallelism among those freshwater populations, alongside the probable genetic bottleneck in the founding of Atlantic marine populations ([39]). The fact that, despite the larger geographic distance between the two North American radiations, they have more similar phenotypes and exhibit stronger genomic parallelism demonstrates that environment, and by extension selection, can counteract the effects of distance to some extent. Explicit measurements of environment are thus important in predicting parallel evolution even across large geographic scales, and distance alone can be misleading. Phenotypic parallelism was observed for several traits across continents, and populations with similar phenotypes also exhibit stronger genomic parallelism. However, this relationship is weaker than for commonality of environment, and is likely confounded by variable heritability and genetic architecture among phenotypes. Most importantly, our results highlight that quantitative analyses of phenotypes and environments and of their relationship can provide a good prediction of expected genomic parallelism, and provide a much clearer picture of major factors that are likely to influence the emergence of a common pattern of genomic divergence than analyses of dichotomous phenotypes and environments.

Methods

Sampling and environmental data collection.

We sampled 18 lakes in North Uist, Scotland between April and June 2013, 18 lakes from Iceland between May and June 2014, 18 lakes from British Columbia (BC) between April and May 2015 and 19 lakes 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. We measured the pH, concentrations of metallic cation concentrations sodium ("Na"), calcium ("Ca") and zinc ("Zn") of each lake, and calculated population prevalence of *Gyrodactylus spp.* and *Schistochephalus solidus*. Concentrations of cations, pH and parasite prevalence per lake are shown in Supplementary Table 2. Details of the fish collection and quantification of abiotic and abiotic variables can be found in Supplementary Information.

DNA extractions, RAD library preparation and sequencing.

Genomic DNA was purified from 10 to 21 individuals from each of the populations, chosen to represent a widely distributed subset of the most environmentally and phenotypically variable lakes (Supplementary Dataset 3). Extracted genomic DNA was normalized to a concentration of 25 ng /μL in 96-well plates.

In 2014 we conducted RAD sequencing on samples from Scotland and from Iceland. Sequencing libraries were prepared and processed into RAD following the modified libraries according to ([65]). In 2016 we conducted RAD sequencing on samples from British Columbia and from Alaska. Sequencing libraries were prepared following the modified single-digest RAD protocol of ([66]). The two RADseq protocols interrogate the same set of loci across the genome, so that the SNP data are compatible across all four radiations. See Supplementary Information for details of the RAD library preparation and sequencing.

Population genetics statistics and phylogenetic tree.

Raw sequence reads were demultiplexed using Stacks – 1.35 ([67]). Number of reads per individual are shown in Supplementary Dataset 3 (see Supplementary Information for details on the alignment of reads and Stacks pipeline used). For Bayenv2 data, autosomal SNPs were called as being in >7 populations, >50% of the individuals within a population, and with a MAF-filter of 0.05. After filtering we retained 26,990, 26,937, 29,111, 26,169 SNPs for Scotland, Iceland, British Columbia and Alaska respectively. For analyses of population structure across all radiations, a subset of unlinked SNPs were generated. Here, autosomal SNPs were called that were present in all radiations and in >50% of individuals within a radiation, with a MAF-filter of 0.05 (within a radiation). Only the first SNP per RAD locus was retained. F_{ST} was bootstrapped and calculated in POPULATIONS. This set of SNPs were then pruned for linkage disequilibrium in plink using indep-pairwise 50 5 0.2. The set of unlinked SNPs were used to construct a neighbour-joining tree for all fish in the R package 'adegenet', using a distance matrix computed from the SNP data ([68]). The tree was bootstrapped 100 times and nodes with less than 80% support were collapsed. The tree was plotted using the 'ape' package in R ([69]). PCA analysis of population structure was conducted using plink ([70]).

Phenotypic and environmental variation - body shape, armour, gill rakers and environmental data analyses.

All morphological measurements (body shape, body armour and spine traits) were done following ([25]). Details of the quantification of phenotypic traits can be found in Supplementary Information.

We performed three Principal Component Analyses (PCAs): one on the armour traits, another on body shape, and another on the 6 environmental variables. 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 and armour variation (Shape_PCs and ArmourPCs) and environmental variation and retained axes that explained more than 10% of the total variance. Armour and environmental PCAs were conducted with scaled inputs due to different units of measurement between variables. Shape PCAs were conducted on morphometric residuals, and as such were not scaled. All phenotypic analyses, including ANOVAs and ANCOVAs and plotting were done in R version 3.4.3.

Genotype-Environment/Phenotype Associations.

For each radiation separately (N=18 to 19 populations) we used Bayenv2 ([43]) to identify associations between genomic allele frequencies (N=10 to 21 individual fish, mean= 17. 8), the set of six biotic and abiotic environmental variables (Ca, Na, pH, Zn, prevalence of *Gyrodactylus spp.* and *Schistochephalus solidus*) and the set of 12 phenotypic traits (Shape_PC1, Shape_PC2, Shape_PC3, DS1, DS2, PS, LP, HP, BAP, Plate_N, Gill_Raker_L and Gill_Raker_N) 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 linkage disequilibrium ($R^2 < 0.4$) in plink ([70]). This cut-off was selected to balance the trade-off between SNPs retained and minimising the effects of linkage. Covariance matrices were therefore calculated using 9619, 7983, 7300 and 5705 SNPs for Alaska, BC, Iceland and Scotland respectively. Covariance matrices were calculated across 100,000 iterations and averaged across 5 independent runs. Bayenv2 was run independently 8 times and final results were averaged across runs.

Environmentally and phenotypically associated SNPs were selected as having a $\log_{10}$-BayesFactor > 1.5 and an absolute Spearman's rank coefficient above the 95th percentile. The combination of BayesFactor and non-parametric measure of correlation helps to avoid selecting SNPs with high BayesFactors due to spurious populations ([18]). SNPs were grouped into 50kb, 75kb, 100kb, 200kb and 0.1 cM windows (Supplementary Table 14) to test the robustness of our results across different extents of linkage. To evaluate whether windows were environmentally or phenotypically associated, we adapted the methodology of ([46]). We calculated the upper 99% binomial expectation for the number of associated SNPs given the total number of SNPs in a specific window, and selected windows that had a greater number of associated SNPs than this expectation. This method controls for variation in SNP density across windows and ensures that significant windows exhibit consistent allele frequency correlations across multiple SNPs. We visualised the genomic locations of associated windows using Manhattan plots (Supplementary Fig. 5) and plotting the residual number of outlier SNPs above the binomial expectation (Supplementary Fig. 6). Linkage groups I-XXI were visualised with the exception of XIX; windows on scaffolds were not visualised. Finally, we compared these associated windows across radiations to examine those that were parallel.

Parallelism statistics.

For all radiation groupings (11 combinations in total: one four radiation grouping, four three radiation groupings, and six 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* ([71]).

Grouping of adjacent windows and expanding parallelism regions.

Windows of 50kb and above were based on a linkage assumption and to minimise 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 based on the likelihood of adjacent associated windows resulting independently being low, suggesting non-independence and probable linkage. These windows are available in Supplementary Table 13.

Multivariate vector comparison of environments and phenotypes.

Vectors for environmental, shape and armour PCs (>10% variance) and gill raker data were calculated as the difference between the linearly predicted maximum and minimum values per radiation. Angles (θ) and difference in length (ΔL) were calculated for each vector between radiations. Significance of vectors was determined through permutations by simulating random traits from a normal distribution with mean and s.d. equivalent to the observed data and assessing vectors of random traits/variation as above (Supplementary Information).

Comparing relative influences of environment, phenotype and genetics.

F_{ST} was calculated between each freshwater population and its relevant marine population (Alaska = MUD1, BC = LICA, Iceland = NYPS, Scotland = OBSM) in 50kb windows using the R package 'PopGenome' (REF). For each MxF 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 MxF F_{ST} parallelism. Dissimilarity matrices of environment and phenotype were calculated as Euclidean distance in PCA space for the 6 and 12 environmental and phenotypic variables respectively. The genetic dissimilarity matrix was composed of genome-wide pairwise F_{ST} estimates between freshwater populations. The matrix of MxF parallelism was associated to environmental, phenotypic and genetic dissimilarity matrices using Mantel tests with 9,999 permutations. Partial Mantel tests were performed with genetic distance as the conditional matrix for environmental and phenotypic effects on MxF parallelism, again with 9,999 permutations.

Acknowledgements

We thank Shaun Robertson, Rebecca Young, Abdul Rahman, Brian Santos, Sara Goodacre, Petur Halldorsson, Bjarni K. Kristjánsson, Dolph Schluter, Kieran Samuk, Diana Rennison and Sara Miller for help with the sampling and sampling permits. We are grateful to Ann Lowe and Laura Dean for help with the DNA extractions, to Cody Wiench, Amanda Stahlke and

Sarah Hendricks for help making the RAD-libraries and to John Brookfield for discussion of probability calculations. This work was funded by a NERC grant (NE/J02239X/1 to A.D.C.M), and further support was provided by NIH grant P30GM103324.

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 field work. I.S.M, M.M. and D.D. generated the phenotypic data. I.S.M. and P.H. generated RAD data and J.R.W., I.S.M. and P.H. analysed it. P.H., M.B. and S.S. helped with the sampling and revised the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Data Accessibility

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 available at https://github.com/JimWhiting91/stickleback_adaptive_radiations/

References

1. Schluter, D. *The ecology of adaptive radiation*. (OUP Oxford, 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. Gould, S. J. *Wonderful life: The Burgess shale and the nature of history*. (New York: W W Norton and Company) (1989).

6. Schluter, D. Adaptive Radiation Along Genetic Lines of Least Resistance. *Evolution* **50**, 1766 (1996).

7. Roff, D. The evolution of the G matrix: Selection or drift? *Heredity* **84**, 135–142 (2000).

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

9. Stuart, Y. E. Divergent uses of "parallel evolution" during the history of the American naturalist. *Am. Nat.* **193**, 11–19 (2019).

10. 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).

11. Oke, K. B., Rolshausen, G., LeBlond, C. & Hendry, A. P. How parallel is parallel evolution? A comparative analysis in fishes. *Am. Nat* **190**, 1-6 (2017).

12. Stuart, Y. E. *et al.* Contrasting effects of environment and genetics generate a continuum of

parallel evolution. *Nat. Ecol. Evol.* **1,** 0158 (2017).

13. Moser, D., Roesti, M. & Berner, D. Repeated Lake-Stream Divergence in Stickleback Life History within a Central European Lake Basin. *PLoS One* **7,** e50620 (2012).

14. MacColl, A. D. C. The ecological causes of evolution. *Trends in Ecology and Evolution* **26**, 514–522 (2011).

15. Elmer, K. R. & Meyer, A. Adaptation in the age of ecological genomics: Insights from parallelism and convergence. *Trends Ecol. Evol.* **26**, 298–306 (2011).

16. 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 Biol. Sci.* **279**, 5039–5047 (2012).

17. Jones, F. C. *et al.* The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55–61 (2012).

18. Günther, T. & Coop, G. Robust identification of local adaptation from allele frequencies. *Genetics* **195**, 205–220 (2013).

19. Jacobs, A. *et al.* Convergence in form and function overcomes non-parallel evolutionary histories in Arctic charr. *BioRxiv*, 265272 (2018).

20. 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.* (2019).

21. Bell, M. A. & Foster, S. A. *The evolutionary biology of the threespine stickleback* (Oxford University, 1994).

22. Taylor, E. B. & McPhail, J. D. Historical contingency and ecological determinism interact to prime speciation in sticklebacks, Gasterosteus. *Proc. R. Soc. B Biol. Sci.* **267**, 2375–2384 (2000).

23. 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).

24. 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).

25. 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).

26. Colosimo, P. F. *et al.* Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**, 1928–1933 (2005).

27. 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).

28. Raeymaekers, J. A. M. *et al.* Adaptive and non-adaptive divergence in a common landscape. *Nat. Commun.* **8**, (2017).

29. 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).

30. Ferchaud, A. L. & Hansen, M. M. The impact of selection, gene flow and demographic history on heterogeneous genomic divergence: Three-spine sticklebacks in divergent environments. *Mol. Ecol.* **25**, 238–259 (2016).

31. Roesti, M., Moser, D. & Berner, D. Recombination in the threespine stickleback genome - Patterns and consequences. *Mol. Ecol.* **22**, 3014–3027 (2013).

32. Schluter, D. & Conte, G. L. Genetics and ecological speciation. *Proc. Natl. Acad. Sci.* **106**, 9955–9962 (2009).

33. Renaut, S., Owens, G. L. & Rieseberg, L. H. Shared selective pressure and local genomic landscape lead to repeatable patterns of genomic divergence in sunflowers. *Mol. Ecol.* **23**, 311–324 (2014).

34. 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).

35. De Schamphelaere, K. A. C. *et al.* Reproductive toxicity of dietary zinc to Daphnia magna. *Aquat. Toxicol.* **70**, 233–244 (2004).

36. 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).

37. Nelson, T. C. & Cresko, W. A. Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. *Evol. Lett.* **2**, 9–21 (2018).

38. Rennison, D. J., Delmore, K. E., Samuk, K., Owens, G. L. & Miller, S. E. Shared patterns of genome-wide differentiation are more strongly predicted by geography than by ecology. (2019). doi:10.1086/706476

39. Fang, B., Kemppainen, P., Momigliano, P., & Merilä, J. Oceans apart: Heterogeneous patterns of parallel evolution in sticklebacks. *BioRxiv*, 826412 (2019).

40. 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).

41. 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).

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

43. Coop, G., Witonsky, D., Di Rienzo, A. & Pritchard, J. K. Using environmental correlations to identify loci underlying local adaptation. *Genetics* **185**, 1411–1423 (2010).

44. Hohenlohe, P. A. *et al.* Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* **6**, e1000862 (2010).

45. 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).

46. Yeaman, S. *et al.* Convergent local adaptation to climate in distantly related conifers. *Science* **353**, 1431–1433 (2016).

47. Miller, C. T. *et al.* Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. *Genetics* **197**, 405–420 (2014).

48. Chan, Y. F. *et al.* Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a pitxl enhancer. *Science* **327**, 302–305 (2010).

49. 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).

50. Day, T., Pritchard, J., & Schluter, D. A comparison of two sticklebacks. *Evolution*, *48*(5), 1723-1734 (1994).

51. 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).

52. 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).

53. Marques, D. A. *et al.* Genomics of Rapid Incipient Speciation in Sympatric Threespine Stickleback. *PLoS Genet.* **12**, (2016).

54. 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).

55. 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. Resources* **17**, 356–361 (2017).

56. Roesti, M., Kueng, B., Moser, D. & Berner, D. The genomics of ecological vicariance in threespine stickleback fish. *Nat. Commun.* **6**, (2015).

57. Catchen, J. M. *et al.* Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. *Mol. Ecol. Resources* **17**, 362–365 (2017).

58. Samuk, K. *et al.* Gene flow and selection interact to promote adaptive divergence in regions of low recombination. *Mol. Ecol.* **26**, 4378–4390 (2017).

59. 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).

60. 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).

61. Ravinet, M. *et al.* Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology* **30**, 1450–1477 (2017).

62. 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).

63. Fisher, R. A. *The genetical theory of natural selection*, 1st edn. Clarendon. (1930).

64. Ken A. Thompson, Osmond, M. M. & Schluter, D. Patterns of speciation and parallel genetic evolution under adaptation from standing variation. *biorxiv.org* **18**, 292–296 (2010).

65. 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).

66. Ali, O. A. *et al.* Rad capture (Rapture): Flexible and efficient sequence-based genotyping. *Genetics* **202**, 389–400 (2016).

67. 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).

68. Jombart, T. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405 (2008).

69. Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).

70. Chang, C. C. *et al.* Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **4** (2015).

71. Storey, J. qvalue: Q-Value estimation for false discovery rate control. R package version 2.0.0. *uni-bayreuth.de* (2015).

Fig. 1. Sampling sites and bootstrapped NJ tree for stickleback from 73 freshwater populations from four countries on two continents, based on 8,395 genetic markers for 1,304 individuals. All nodes shown have bootstrap support of at least 80 (other nodes were collapsed). Branches are coloured by the radiation to which they belong. Tips represent individual fish, which were generally tightly clustered by population (small labels). Stars represent lakes sampled.

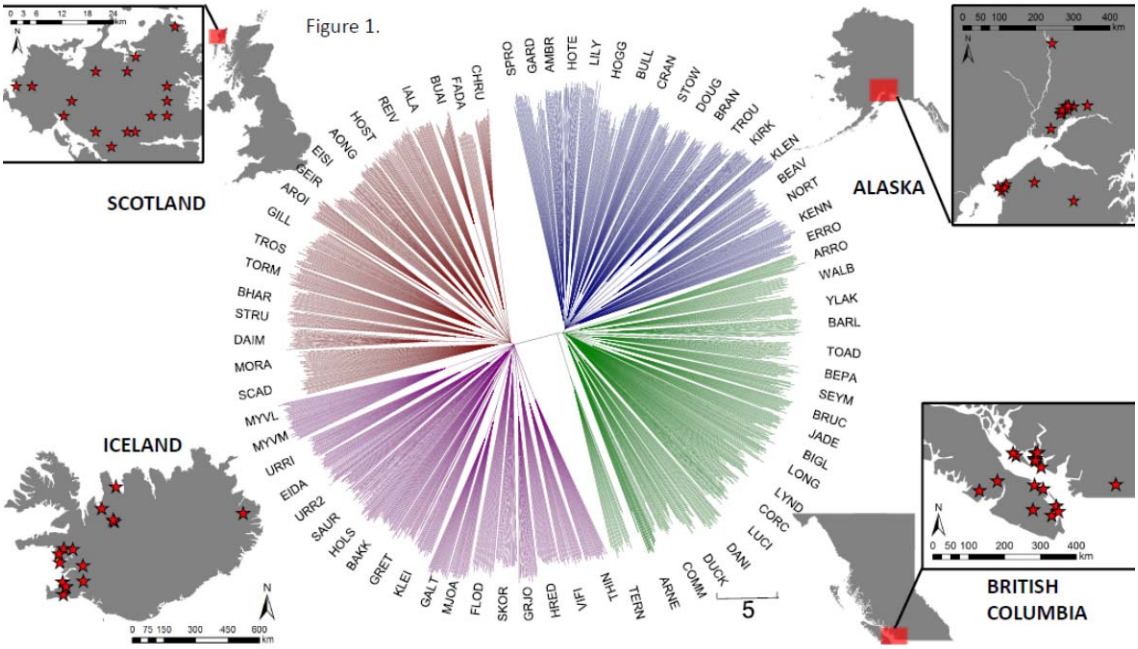

Fig. 2. Principal Component Analyses of environmental variables (Environment); regression residuals of Procrustes coordinates against log centroid of body shape (Shape); armour traits (Armour: length of dorsal spines 1 (DS1) and 2 (DS2), length of pelvic spine (PS), length and height of pelvis (LP and HP), length of biggest armour plate (BAP) and number of armour plates (Plate_N); and gill raker numbers (Gill_Raker_N) vs gill raker length (Gill_Raker_L) (Trophic). Each point represents a population and ellipses are 95% confidence ellipses. Names of variables with the highest positive (+) or negative (-) loadings on each axes are on legends of each axes. All loadings of variables on the first 3 PCs are in Supplementary Table 3. Marine populations (+) are projected where data was available using PC loadings calculated with freshwater populations only.

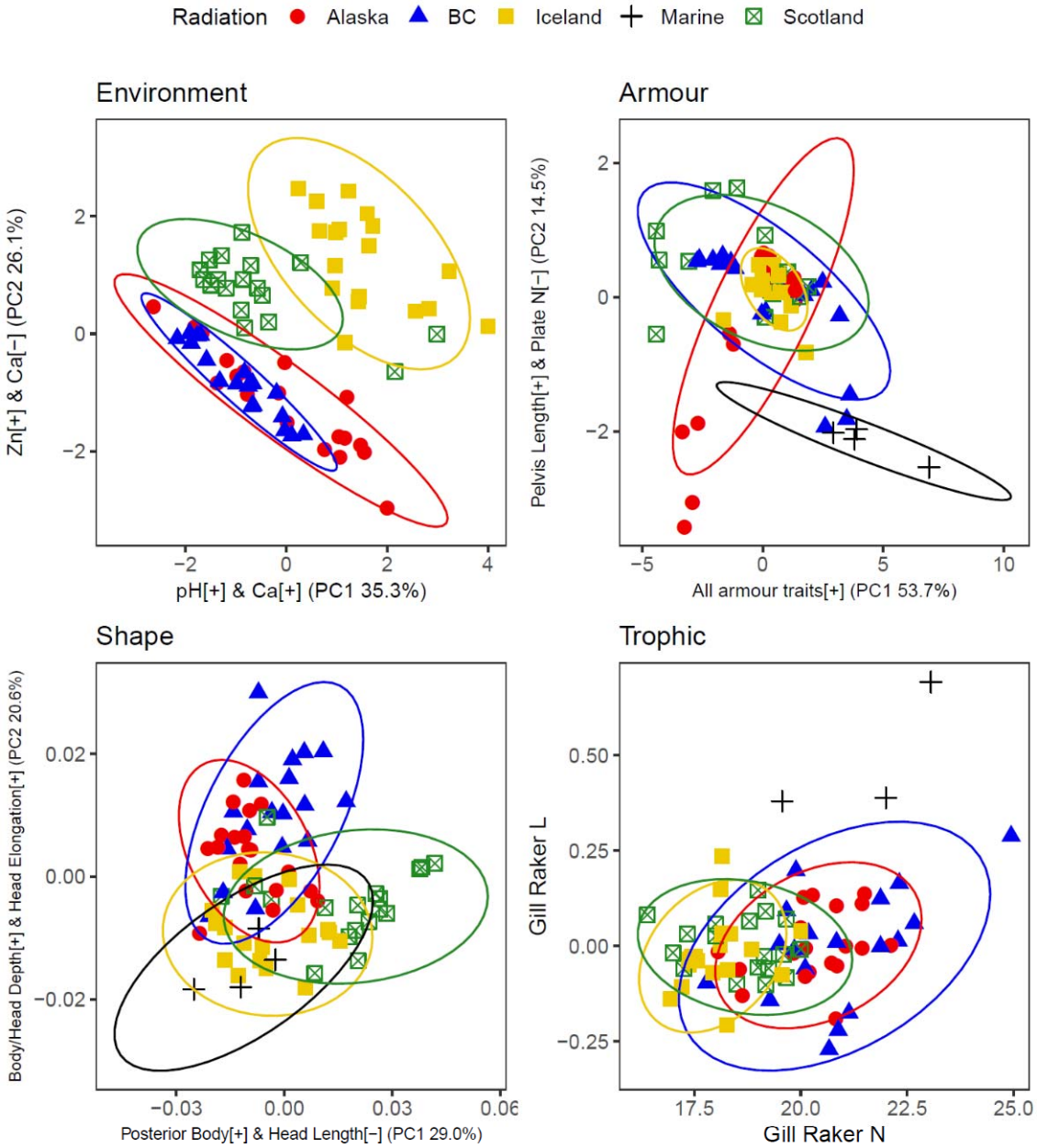

Fig. 3. Expected and observed counts of 50kb windows containing an above 99% binomial expectation number of SNPs associated with marine x freshwater (MxF), environmental variables and phenotypic traits in at least 2 radiations. Expected bars (grey) represent mean counts across 10,000 simulated outcomes with 95% confidence intervals per a one-tailed hypothesis. Asterisks denote significance of FDR-corrected one-tailed tests between the observed counts and the 100,000 simulated counts at the <0.05 (*), <0.01 (**) and <0.001 (***) levels.

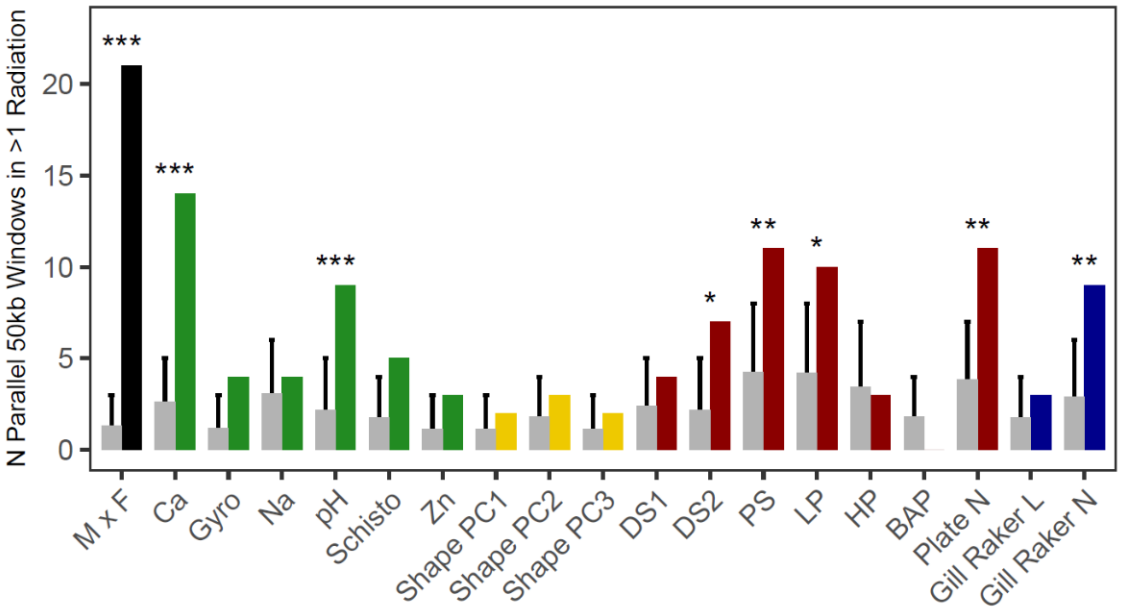

Fig. 4. Associations between genome-wide marine - freshwater F_{ST} and environmental, phenotypic and genetic distance across all pairwise comparisons of 73 freshwater populations. A) Proportion of MxF F_{ST} 50kb outlier windows that overlap among freshwater replicates. Freshwater populations are ordered as Alaska, British Columbia, Iceland and Scotland, with these location distinguishable as four clear clusters. B) Environmental distance between freshwater populations, recorded as euclidean distance in PCA space for all 6 environmental variables. C) Phenotypic distance between freshwater populations, recorded as for environmental distance for the 12 phenotypic variables. D) Genetic distance between freshwater populations, recorded as genome-wide pairwise F_{ST} based on 8,395 unlinked SNPs. E-G) Associations between environmental (E), phenotypic (F) and genetic (G) distance and MxF F_{ST} overlap (log-transformed). Points are coloured according to whether pairwise comparison is being made within a radiation or across radiations.