Genetic population structure constrains local adaptation in sticklebacks

Petri Kemppainen[1,*], Zitong Li[1,2], Pasi Rastas[1,3], Ari Löytynoja[3], Bohao Fang[1], Jing Yang[1,4], Baocheng Guo[5], Takahito Shikano[1], Juha Merilä[1]

[1]Ecological Genetics Research Unit, Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, FI-00014 University of Helsinki, Finland

[2]CSIRO Agriculture & Food, GPO Box 1600, Canberra, ACT 2601, Australia;

[3]Institute of Biotechnology, FI-00014 University of Helsinki, Finland

[4]Chinese Sturgeon Research Institute, Three Gorges Corporation, Yichang, 443100, China

[5]The Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China

*Corresponding author: petrikemppainen2@gmail.com

Abstract

Repeated and independent adaptation to specific environmental conditions from standing genetic variation is common. However, if genetic variation is limited, the evolution of similar locally adapted traits may be restricted to genetically different and potentially less optimal solutions, or prevented from happening altogether. Using a quantitative trait locus (QTL) mapping approach, we identified the genomic regions responsible for the repeated pelvic reduction (PR) in three crosses between nine-spined stickleback populations expressing full and reduced pelvic structures. In one cross, PR mapped to linkage group 7 (LG7) containing the gene *Pitx1*, known to control pelvic reduction also in the three-spined stickleback. In the two other crosses, PR was polygenic and attributed to ten novel QTL, of which 90% were unique to specific crosses. When screening the genomes from 27 different populations for deletions in the *Pitx1* regulatory element, these were only found in the population in which PR mapped to LG7, even though the morphological data indicated large effect QTL for PR in several other populations as well. Consistent with the available theory and simulations parameterised on empirical data, we hypothesise that the observed variability in genetic architecture of PR is due to heterogeneity in the spatial distribution of standing genetic variation caused by strong population structuring and genetic isolation by distance in the sea.

Keywords: convergent evolution, epistasis, local adaptation, pelvic reduction, Pitx1, *Pungitius pungitius*

Introduction

Failure to evolve in response to changing environmental conditions may lead to extirpation or even extinction (Orr and Unckless 2008). If heritable variation necessary for adapting to environmental change already exists in the form of standing genetic variation, genetic adaptation may proceed swiftly, at least compared to the time it would take for populations to adapt from novel mutations (Orr and Unckless 2008; Barrett and Schluter 2008; Thompson *et al*. 2019). Furthermore, when exposed to novel yet similar environments, populations derived from the same ancestral population – hence carrying the same pool of alleles – can often be expected to respond to similar selection pressures in a similar fashion, leading to parallel phenotypic evolution (Arendt and Reznick 2007; Schluter and Conte 2009; Elmer and Meyer 2011; Stern 2013; Conte *et al*. 2015; Bolnick *et al*. 2018; Hermisson and Pennings 2017). However, in genetically highly structured species, potentially advantageous rare alleles may be lost due to founder events and random genetic drift, thus preventing adaptation. Alternatively, due to heterogeneity in the distribution of standing genetic variation, adaptation to given selection pressures could more likely be acquired with phylogenetically independent alleles (rather than alleles that are identical by descent) at the same or different loci influencing the same trait, even if they may differ significantly in their fitness effects (Cohan 1984; Merilä 2013, 2014; Rosenblum *et al.* 2014). Thus, the demographic history of populations likely plays a central role in determining the likelihood of local adaptation, and hence also parallel phenotypic evolution.

The three-spined stickleback (*Gasterosteus aculeatus*) is a widely used model

system to study adaptive evolution in the wild (Bell and Foster 1994; Gibson 2005). Its ability to rapidly adapt to local environmental conditions has often been shown to stem from a global pool of ancestral standing genetic variation (Schluter and Conte 2009; Jones *et al*. 2012; Terekhanova *et al.* 2014, 2019). The nine-spined stickleback (*Pungitius pungitius*) has been emerging as another model system for the study of repeated evolution in the wild (Merilä 2013). In general, it differs from the three-spined stickleback by having smaller effective population sizes (N_e), reduced gene flow in the sea, and a tendency to occur in small landlocked ponds in complete isolation from other populations (Shikano *et al.* 2010a; DeFaveri *et al*. 2012; Merilä 2013; this study). Given their contrasting population demographic characteristics, three- and nine-spined sticklebacks can thus be expected to respond differently with respect to local adaption to newly colonised freshwater habitats.

Regressive evolution of the pelvic complex has occurred in freshwater populations of three (*viz*., *Gasterosteus*, *Pungitius*, and *Culaea*) of the five recognised stickleback genera since the last glacial period (reviewed in Klepaker *et al*. 2013). While marine populations of three- and nine-spined sticklebacks usually have complete pelvic structures with fully developed pelvic girdles and lateral pelvic spines, partial or even complete pelvic reduction is common in freshwater populations (Blouw and Boyd 1992; Shapiro *et al*. 2004, 2006; Herczeg *et al*. 2010; Klepaker *et al*. 2013). Several factors may contribute to this, including the absence of gape-limited predatory fish and limited calcium availability, as well as the presence of certain insect predators (Reist *et al*. 1980; Reimchen *et al*. 1983; Giles 1983; Bell *et al*. 1993; Karhunen *et al*. 2013; Chan *et al*.

83 2010) in freshwater environments. Collectively, sticklebacks provide an important model
84 system to study the genetic mechanisms underlying the adaptive parallel pelvic reduction
85 at both intra- and inter-specific levels, under a wide range of population demographic
86 settings. However, studies of parallel patterns of marine-freshwater divergence in nine-
87 spined sticklebacks are still scarce (Herczeg *et al*. 2010; Shikano *et al*. 2010b; Wang *et*
88 *al*. 2020), especially at the genetic level, precluding any comprehensive and conclusive
89 comparison of the two species.

90 The genetic basis of pelvic reduction in the three-spined stickleback is well
91 understood. Quantitative trait locus (QTL) mapping studies have identified a single
92 chromosomal region containing the gene *Pituitary homeobox transcription factor 1*
93 (*Pitx1*) that explains more than two thirds of the variance in pelvic size in crosses
94 between individuals with complete pelvic girdles and spines, and pelvic-reduced
95 individuals (Cresko *et al*. 2004; Shapiro *et al*. 2004; Coyle *et al*. 2007). Pelvic loss in the
96 marine-freshwater three-spined stickleback model system is predominantly caused by
97 expression changes of the *Pitx1* gene due to recurrent deletions of the pelvic enhancer
98 (*Pel*) upstream of *Pitx1* (Chan *et al*. 2010; Xie *et al.* 2019; in benthic-limnetic and lake-
99 stream pairs of three-spined sticklebacks, the genetic architecture of pelvic reduction is
100 more diversified; Peichel *et al.* 2001, 2017; Deagle *et al*. 2012; Stuart *et al.* 2017). While
101 pelvic reduction in freshwater is much less common in nine- than three-spined
102 stickleback (Klepaker *et al*. 2013; Fig. 1), two independent QTL studies also identified
103 *Pitx1* in linkage group (LG) 7 as a major cause for pelvic reduction in a Canadian
104 (Shapiro *et al.* 2006) and a Finnish (Shikano *et al.* 2013) population. Another large effect

105 region in LG4 was found to be associated with pelvic reduction in an Alaskan population
106 (Shapiro *et al.* 2009). Similar to the three-spined stickleback, pelvic spine and pelvic
107 girdle sizes are strongly correlated in the population from the Finnish pond Rytilampi,
108 since *Pitx1* controls both phenotypes (Shikano *et al.* 2013; Fig. 1 and Supplementary
109 Table 1). In contrast, a considerable amount of phenotypic variation with respect to these
110 traits and their inter-correlations exists among different freshwater pond populations in
111 northern Europe (Herczeg *et al.* 2010; Karhunen *et al.* 2013, 2014; Fig. 1). Given their
112 high heritability (Blouw and Boyd 1992; Leinonen *et al.* 2011), the lack of correlation
113 between spine and girdle lengths suggests that they can be independently controlled by
114 different QTL. Thus, the genetic underpinnings of pelvic reduction in nine-spined
115 sticklebacks (when it occurs) appear to be more diversified than those in the marine-
116 freshwater three-spined stickleback model system (Merilä 2013, 2014).
117 Starting with a survey of previously published phenotypic data on pelvic
118 reduction in the wild, we aimed to investigate the possible genetic heterogeneity
119 underlying pelvic reduction in different nine-spined stickleback populations by mapping
120 QTL for pelvic reduction in three independently colonised ponds. One was the previously
121 studied Rytilampi (earlier analysed only with microsatellites, Shikano *et al.* 2013), now
122 re-analysed along with two new populations (Bynästjärnen and Pyöreälampi) using >75
123 000 SNPs. This data was subjected to a cutting-edge mapping approach (Li Z. *et al.* 2017,
124 2018) that can provide more information on the source of the QTL effects than has been
125 previously possible. To further assess the extent to which *Pel* could be responsible for
126 pelvic reduction in the nine-spined stickleback, we screened the whole genomes of

127 individuals from 27 wild populations for deletions in the genomic region spanning the *Pel*
128 element and the *Pitx1* gene. Finally, utilising more comprehensive geographic sampling
129 than in previous studies (Shikano *et al.* 2010a; DeFaveri *et al*. 2012; Merilä 2013), along
130 with high-quality SNP data, we re-assessed the differences in population structuring and
131 genetic isolation by distance (IBD) between nine-and three-spined sticklebacks. We
132 hypothesised that, in contrast to three-spined sticklebacks, both the scarcity and
133 variability in the genetic architecture of pelvic reduction in nine-spined sticklebacks is
134 due to heterogeneity in the spatial patterns of standing genetic variation caused by strong
135 population structuring – both among adjacent pond populations as well as in the ancestral
136 sea population(s). Building on previous theoretical work, this hypothesis was tested using
137 simulated data parameterised on population demographic parameters obtained from
138 empirical data.
139

140 **Materials and Methods**

141 *Fish collection, crossing, and rearing*

142 For the QTL crosses, three different marine F_0 generation females from the Baltic Sea
143 (Helsinki, Finland; 60°13′N, 25°11′E) were crossed with a freshwater F_0 generation male
144 from either Bynästjärnen (64°27′N, 19°26′E), Pyöreälampi (66°15′N, 29°26′E) or
145 Rytilampi (66°23′N, 29°19′E) ponds. Fish crossing, rearing, and sampling followed the
146 experimental protocol used in the earlier study of the Rytilampi population (Shikano *et*
147 *al.* 2013; Laine *et al.* 2013). For Rytilampi, the F_0 generation fish were artificially mated
148 in the lab during the early breeding season of 2006 (Shikano *et al.* 2013), and the

149 resulting full-sib F_1-offspring were group-reared in aquaria until mature, as explained in
150 Shikano *et al.* (2013). Two randomly chosen F_1 individuals were mated repeatedly (seven
151 different clutches) to produce the F_2 generation between September and October 2008.
152 The offspring were reared in separate aquaria. The same procedure was followed for
153 Pyöreälampi (F_0 mating: Jun 2011; F_1 mating: Jul–Sep 2012; F_2 rearing: Jul 2012–Apr
154 2013; 8 different clutches) and Bynästjärnen (F_0 mating: Jun 2011; F_1 mating: Nov 2013–
155 Jan 2014, F_2 rearing: Nov 2013–Aug 2014; 6 different clutches). The F_2 fish were
156 euthanized at 187, 238, and 238 days post-hatch for Rytilampi, Pyöreälampi and
157 Bynästjärnen, respectively. At this point, the fish were on average 52.3 mm in standard
158 length (Rytilampi = 48.6 mm; Pyöreälampi = 52.3 mm; Bynästjärnen = 53.6 mm), and
159 weighed on average 1.34 g (wet weight; Rytilampi = 1.07 g; Pyöreälampi = 1.49 g;
160 Bynästjärnen = 1.44 g). In total, 308, 283 and 279 F_2 offspring were available for
161 analyses from Helsinki × Bynästjärnen (HEL × BYN), Helsinki × Pyöreälampi (HEL ×
162 PYÖ) and Helsinki × Rytilampi (HEL × RYT) crosses, respectively.
163 The experiments were conducted under licenses from the Finnish National
164 Animal Experiment Board (#STH379A and #STH223A). Experimental protocols were
165 approved by the ethics committee of the University of Helsinki, and all experiments were
166 performed in accordance with relevant guidelines and regulations.
167
168 *Morphological data*
169 To visualise bony elements, all of the F_2-progeny were stained with Alizarin Red S
170 following Pritchard and Schluter (2001). Pelvic spine and girdle lengths from both sides

171 of the body, as well as standard body length, were measured with digital calipers to the
172 nearest 0.01 mm. Although it is known that the left-right asymmetry of the pelvic girdle
173 is heritable in sticklebacks (Blouw and Boyd 1992; Bell *et al.* 2007; Coyle *et al.* 2007),
174 we did not specifically study this. To reduce the number of tests, the mean of the left and
175 right-side measurements was used (analyses conducted for the two sides separately
176 always yielded qualitatively similar results as the tests conducted with the averages;
177 results not shown). All measurements were made by the same person twice; the
178 repeatability (R; Becker 1984) ranged between 0.80 and 0.84 for girdle lengths, and
179 between 0.98 and 0.99 for spine lengths. The QTL analyses were performed on both
180 absolute and relative (scaled to the total body length) trait values, but for all of the
181 analyses that compared phenotypic data between populations (which also differ in body
182 sizes), only relative trait values were used. Previously published phenotypic data from 19
183 wild populations were obtained from Herczeg *et al*. (2010) and Karhunen *et al.* (2013).
184 These included data on pelvic spine and girdle lengths of wild-collected individuals from
185 ten pond populations (Abbortjärn = ABB, Bolotjone = BOL, Karilampi = KAR,
186 Kirkasvetinen lampi = KRK, Mashinnoje = MAS, Lil-Navartjärn = NAV, Hansmytjärn =
187 HAN, Rytilampi = RYT, Bynästjärnen = BYN), four lake populations (Iso Porontima =
188 POR, Riikojärvi = RII, Joortilojärvi = JOR, Västre-Skavträsket = SKA) and five marine
189 populations (Fiskebäckskil = FIS, Trelleborg = TRE, Bölesviken = BÖL, Helsinki = HEL
190 LEV = Levin Navolak), as well data on common garden-reared F_1 generation individuals
191 from two marine (HEL and LEV) and two pond (BYN and PYÖ) populations. Visibly
192 broken spines were treated as missing data. A map showing the geographic location of

193 these populations is provided in Supplementary Fig. 1.

194

195 *Genotyping and linkage map construction*

196 The RAD sequencing protocol used to obtain the SNP data was the same as in Yang *et al.*
197 (2016) and Li Z. *et al*. (2017). In short, genomic DNA was extracted from ethanol
198 preserved fin clips using the phenol-chloroform method (Taggart *et al*. 1992). DNA was
199 fragmented with PstI restriction enzyme and the resulting 300–500 bp long fragments
200 were gel purified. Illumina sequencing adaptors and library specific barcodes were
201 ligated to the digested DNA fragments, and the barcoded RAD samples were pooled and
202 sequenced on 24 lanes of the Illumina HiSeq2000 platform with 45 bp single-end
203 strategy. RAD library construction and sequencing were conducted by BGI
204 HONGKONG CO., LIMITED. After eliminating adapters and barcodes from reads, a
205 quality check was done using FastQC
206 (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/).

207 Linkage mapping for the three crosses was conducted using Lep-MAP3 (LM3;
208 Rastas 2017), as described in detail in Li H. *et al.* (2018). LM3 can infer the
209 parental/grandparental phase based on dense SNP data, which allowed us to utilise the
210 four-way cross QTL mapping method described below. Input data were generated by first
211 mapping individual fastq files to the nine-spined stickleback reference genome using
212 BWA mem (Li H. 2013) and SAMtools (mpileup; Li H. *et al.* 2009), followed by
213 pileupParser.awk and pileup2posterior.awk scripts from the LM3 pipeline using default
214 parameters (see Supplementary File 2 for details).

215

216 *Dimensionality reduction by linkage disequilibrium network clustering*

217 It is essential in QTL mapping to correct for multiple testing in order to reduce the rate of
218 false positives. Moreover, in large genomic datasets, physically adjacent SNPs –
219 particularly those from experimental crosses – are often in linkage disequilibrium (LD),
220 i.e. correlated. Since a group of SNPs in high LD explains similar amounts of genetic
221 variation in a given trait, it is reasonable to apply a dimensional reduction procedure
222 before QTL-mapping to exclude the redundant information from the data. Here we used
223 linkage disequilibrium (LD) network clustering (LDn-clustering) and PC regression as
224 dimensionality reduction tools prior to single-locus QTL mapping (Li Z. *et al.* 2018). The
225 first step of this approach involves an extension of a method developed by Kemppainen
226 *et al.* (2015), which uses LD network analysis for grouping loci connected by high LD.
227 The second step involves principal component analysis (PCA) as a method for
228 dimensionality reduction in each cluster of loci connected by high LD (LD-clusters). This
229 was achieved by the function LDnClustering in the R-package LDna (v.0.64; Li Z. *et al*.
230 2018). The method used here differs slightly from the original method described in Li Z.
231 *et al.* (2018) to increase efficiency in computational speed and reduce complexity (see
232 Supplementary File 3 for details).

233

234 *QTL mapping: four-way crosses model*

235 In some circumstances, such as a four-way cross (Xu 1996), F_1 hybrids of two
236 heterozygous parents (Van Ooijen 2009), and an outbred F_2 design (Xu 2013a), it is

237 possible that up to four different alleles – A and B from the dam, and C and D from the

238 sire – segregate in the population. In such cases, a linear regression model for QTL

239 analysis of the outbred F_2 data (Li Z. *et al.* 2018) is defined by:

240 $$y_i = \beta_0 + x_{dij}\beta_{dj} + x_{sij}\beta_{sj} + z_{ij}\gamma_j + \varepsilon_i, \varepsilon_i \sim N(0, \sigma^2), (1)$$

241 where y_i is the phenotype value of individual i (i=1,…,n), x_{dij}, x_{sij}, z_{ij} are genotypes coded

242 as

243 $$\begin{vmatrix} +1 + 1 + 1 \\ +1 - 1 - 1 \\ -1 + 1 - 1 \\ -1 - 1 + 1 \end{vmatrix} \begin{matrix} for genotype AC, \\ for AD, \\ for BC, \\ for BD. \end{matrix} ,(2)$$

244 (Xu 2013b), β_0 is the parameter of the population mean, β_{dj} is the substitution effect of

245 alleles A and B of the dam (i.e. the grandfather in F_0) at the locus j (j=1,…,p; p is the

246 total number of SNPs), β_{dj} is the substitution effect of alleles C and D of the sire (i.e. the

247 grandmother in F_0), γ_j is the dominance effect, and ε_i is the residual error term mutually

248 following an independent normal distribution.

249 The model (1) requires the knowledge of the grandparental phase (produced by

250 LM3) with the benefit that the source (*viz*. F_0 female or F_0 male) of the observed QTL

251 effect can be inferred, as described in more detail in Supplementary File 1. A multiple

252 correction on the basis of permutation tests was further conducted to control for false

253 positives due to multiplicity (Li Z. *et al.* 2017) with $1e^5$ replicates.

254

255 *Estimating the proportion of phenotypic variance explained by SNPs*

256 The overall proportion of phenotypic variance (PVE) explained jointly by all SNPs (an
257 approximation of the narrow sense heritability) was obtained using LASSO regression
258 (Tibshirani 1996), which incorporates all the SNPs into a multi-locus model:
259

260 $$\frac{1}{2n}\sum_{i=l}^{n}\left(y_i - \beta_0 - x_{dij}\beta_{dj} - x_{sij}\beta_{sj} - z_{ij}\gamma_j\right) + \lambda\sum_{j=l}^{p}\left(\left|\beta_{dj}\right| + \left|\beta_{sj}\right| + \left|\gamma_j\right|\right), (3)$$

261 where the l_1 penalty term $\lambda\sum_{j=l}^{p}\left(\left|\beta_{dj}\right| + \left|\beta_{sj}\right| + \left|\gamma_j\right|\right)$ $(\lambda > 0)$ shrinks the regression
262 parameters towards zero; all other symbols are defined in the same way as in Equation
263 (1).

264

265 Following Sillanpää (2011), the PVE can be estimated by the formula:
266

267 $$PVE_{total} = \frac{var\left(\sum_{j=l}^{p} x_j \hat{\beta}_j\right)}{var(y)} \approx \frac{\mathrm{var(y)} - \sigma_0^2}{var(y)}, (4)$$

268 where $\hat{\beta}_j$ is the effects of the SNPs, and σ_0^2 is the LASSO residual variance estimated by a
269 cross-validation-based estimator introduced by Reid *et al*. (2016). The PVE explained by
270 each linkage group was estimated on the basis of the LASSO estimates using the
271 following formula:

272 $$PVE_{LG} \approx PVE_{total} - \frac{var\left(\sum x_j \hat{\beta}_j\right)}{var(y)}, (5)$$

273 where $\hat{\beta}_j (j \notin G)$ represents all the effects estimated by the LASSO of the SNPs that do not
274 belong to a set of SNPs (e.g. to a chromosome/linkage group). A similar approach was

275 used to estimate the contribution of grandparental alleles and to evaluate the dominance
276 component by treating the coding systems [x_{dij}, x_{sij}, z_{ij}] (2) as different groups of SNPs. A
277 custom R-code for PVE estimation is available from DRYAD (DOI
278 https://doi.org/10.5061/dryad.76hdr7str).

279
280 *Scanning for Pel deletions in full-genome sequence data*

281 A minimum of 20 samples from populations RYT, MAS, BOL, BYN and PYÖ, and 10–
282 31 samples from an additional 22 populations (from Northern Europe and USA;
283 Supplementary Fig. 1 and Supplementary Table 2) were sequenced to 10× coverage by
284 BGI HONGKONG CO., LIMITED. Reads were mapped to the nine-spined stickleback
285 reference genome (Varadharajan *et al*. 2019) with BWA mem (Li H. 2013), and site-wise
286 sequencing coverages were computed with SAMtools (depth; Li H. *et al*. 2009). Relative
287 sequence depths across the *H2afy-Pitx1* intergenic region were estimated for the five
288 focal populations by first computing the median depths for 1000 bp sliding windows, and
289 then normalising these by the median depth for the full intergenic region. The Pel-
290 2.5kb^{SALR} region was extracted from the original BAC contig (GenBank accession
291 number GU130433.1) and mapped to the nine-spined stickleback intergenic region with
292 minimap2 (Li H. 2018). The sequencing depths for the Pel-2.5kb^{SALR} were normalised by
293 dividing the mean depths of the *Pel* region by the mean depths of the full intergenic
294 region. Gene annotations and relative sequencing depths (average and confidence
295 intervals) were computed and visualised using the R-package Gviz (Hahne and Ivanek
296 2016). Lastly, we scanned the literature for genes that are known to affect pelvic and hind
297 limb development, and searched for potential matches in the nine-spined stickleback

genome (Varadharajan *et al*. 2019) in regions where significant QTL were found. The custom R-code for these analyses is provided in DRYAD (DOI https://doi.org/10.5061/dryad.76hdr7str).

Simulations

The repeated local adaptation in independently colonised freshwater stickleback populations is widely considered to be due to selection on standing genetic variation available in the sea, which in turn is maintained by recurrent gene flow from previously colonised freshwater populations (the "transporter hypothesis"; Schluter and Conte 2009). The effects of standing genetic variation (assuming a panmictic ancestral population) on the probability of parallel local adaptation have already been studied (Ralph and Coop 2015a; MacPhearson and Nuisimer 2016; Lee and Coop 2017; Galloway and Cresko 2019), along with the effects of population structure on adaptation in continuous landscapes (Hermisson and Pennings 2005; Ralph and Coop 2010, 2015b). However, no simulations to date have explicitly considered the effects of heterogeneous patterns in the distribution of standing genetic variation on local adaptation/parallel evolution. Thus, we used forward-in-time simulations to explore the effects of the differences in population demographic parameters between nine- and three-spined sticklebacks on the probability of local adaptation, as described in Supplementary File 5. These simulations were parameterised using empirical estimates of population structure and isolation by distance in the sea for the two species (Supplementary File 4).

In short, two freshwater populations were connected to different marine populations at the ends of a stepping-stone chain of marine populations with the carrying capacity (K) and migration (M) both in the sea, and between the founding (marine) and founder (freshwater) population (Supplementary Fig. 3), thus generating comparable patterns of population structuring and IBD in the sea as in the empirical data (see Results). The two freshwater populations were colonised once and allowed to become locally adapted for 10,000 years. All freshwater adaptation was due to standing genetic variation in the sea that was continually replenished from a "refuge" freshwater population (situated at an equal distance from the two focal freshwater populations), as well as from the two focal freshwater populations once they were colonised. Three different genetic architectures were studied. One (architecture A) corresponds to a single large effect QTL (allowing 100% local adaptation, but with minor effect QTL also affecting the trait), another (architecture B) to a situation where at least two loci are needed to allow local adaptation (allowing 40% or 60% local adaptation), and the third (architecture C) corresponds to the same scenario as A, except that the large effect locus is recessive. Mutation rate was fixed at $1e^{-8}$ and all loci were polymorphic at the beginning of the simulations. More details are given in Supplementary File 4.

Results

Heterogeneity of pelvic reduction in the wild

Re-analysis of previously published phenotypic data from the wild confirmed a high degree of variation among different populations with respect to pelvic spine and girdle

341 lengths and their inter-correlations (Fig. 1, Supplementary Table 1 and Supplementary
342 Fig. 2). For instance, while spines were absent and pelvic girdles strongly reduced (but
343 not completely absent) in RYT, spines and girdles were completely lacking in the MAS
344 population (Fig. 1 and Supplementary Table 1). Furthermore, in the BYN population,
345 spines were absent but pelvic girdles were only partially reduced; in BOL (a population
346 adjacent to MAS), large variation in both spine (SD = 0.037) and girdle (SD = 0.047)
347 lengths was observed, although these two traits were strongly correlated (r^2 = 0.61, Fig. 1,
348 Supplementary Table 1 and Supplementary Fig. 2). This suggests that a large effect locus
349 affecting both pelvic spines and girdles segregated in this population at the time of
350 sampling. In six pond populations (*viz*. ABB, KAR, KRK, NAV, HAN, and PYÖ; Fig.
351 1), relative spine (mean = 0.079) and girdle (mean = 0.15) lengths were only slightly
352 smaller relative to the marine populations (0.11 and 0.16 for spine and girdle lengths,
353 respectively; Supplementary Table 1). A lack of pelvic reduction (relative to the marine
354 populations) was observed only in one pond population (KAR; Fig. 1, Supplementary
355 Table 1 and Supplementary Fig. 2).
356
357 *QTL mapping of pelvic reduction in the Helsinki × Rytilampi cross*

358 After LD-network based complexity reduction, all QTL mapping analyses were
359 performed on 241 PCs (six sex-linked PCs were removed). Re-analyses of the 279 F_2
360 progeny from the HEL × RYT cross confirmed a single QTL region on LG7 for both
361 pelvic spine and girdle lengths in the single-locus analyses (Fig. 2a, b and Table 1). In the
362 multi-locus analyses (absolute trait values), LG7 explained 74–86% of the PVE for both

363 spine and girdle lengths, while all other chromosomes individually explained less than
364 3% of the phenotypic variance (Table 2). An approximately equal amount of the
365 phenotypic variance was explained by alleles inherited from the F_0 male (pond
366 individual) and the F_0 female (marine individual; ~30% of the total variance for all traits;
367 Table 2), respectively, with 15–21% of the phenotypic variance also explained by
368 dominance effects. This is the expected outcome for a recessive QTL when the F_0
369 individuals are fixed for different large effect causal alleles, and when no additional
370 smaller effect loci affect the trait (Klug and Cummings 2018).

371

372 *QTL mapping of pelvic reduction in the Helsinki × Bynästjärnen cross*

373 Among the 308 F_2 progeny of the HEL × BYN cross, single-locus mapping analyses of
374 pelvic spine lengths detected two significant QTL on LG15 (PVE = 8.9%) and LG16
375 (PVE = 13.7%; Fig. 2c, Table 1 and 2) for alleles deriving from the F_0 male (pond
376 individual). Thus, the causal alleles for these QTL segregated in the F_0 pond male (as
377 explained in Supplementary File 1). No dominance effects were detected for these QTL,
378 suggesting that the allelic effects were additive. One QTL on LG6 with an allelic effect
379 deriving only from the F_0 female was also detected (Fig. 2c and Table 1). The QTL
380 significance profiles, in particular for LG15 and LG16, spanned large genomic regions
381 with no obvious peaks (in contrast to LG7 in the HEL × RYT cross; Fig. 2a, b). This
382 remained true when analysing all SNPs individually (fine-mapping; Supplementary Fig.
383 4). The individual and multi-locus phenotypic effects on spine lengths for the QTL on
384 LGs 6, 15 and 16 (using the most significant QTL for each QTL region) are detailed in

385 Figure 3. In the absence of any large effect loci, and when all QTL are additive and
386 independent (as is the case here), phenotypes in the F_2 generation are expected to be
387 normally distributed (Klug and Cummings 2018). For some multi-locus genotypes (of the
388 QTL on LGs 6, 15 and 16), the distribution of spine lengths was approximately normally
389 distributed, except the highly reduced spine lengths, which had long tails, implicating that
390 some of these loci could be involved in epistatic interactions (this was investigated
391 further in Supplementary File 6). One significant male QTL on LG4 for girdle length was
392 also detected (Fig. 2d and Table 1). Results for relative trait values were highly similar to
393 the absolute trait values, except for an additional significant male and female QTL on
394 LG1, as well as another significant female QTL on LG16 (Supplementary Fig. 5, Table 1
395 and Supplementary Table 3).

396 Multi-locus analyses (for absolute trait values) identified 11 LGs that accounted
397 for at least 3% of the phenotypic variance in pelvic spine or girdle lengths in the HEL ×
398 BYN cross (Table 2). The largest of these effects were found on LG15 and LG16, which
399 explained 9% and 12% of the variation in pelvic spine lengths, respectively (Table 2). For
400 pelvic spine and girdle lengths, 39% and 32% of the PVE, respectively, were accounted
401 for by all SNPs in the dataset, which equates to the narrow sense heritability (h^2) also
402 accounting for dominance (but not epistatic interaction) effects. For spine lengths, 24% of
403 the total PVE was attributed to alleles deriving from the F_0 male, and 17% was attributed
404 to alleles deriving from the F_0 female; only 1% was attributed to the dominance effect
405 (Table 2). For girdle lengths, the corresponding numbers were 16%, 6% and 13% (Table
406 2).

407

408 *QTL mapping of pelvic reduction in the Helsinki × Pyöreälampi cross*

409 In the HEL × PYÖ cross (283 F_2 individuals), one QTL for spine length was found on
410 LG9, which was explained by alleles inherited from the F_0 male. Two significant QTL for
411 girdle length were found, one on LG19 (explained by alleles inherited from the F_0 male)
412 and the other on LG4 (explained by alleles inherited from the F_0 female; Fig. 2f and
413 Table 1). In the multi-locus analyses, the PVE for different LGs mirrored these results;
414 the LGs that contain significant QTL explain most of the PVE, while PVE for all other
415 LGs were <4% (Table 2). Overall, the multi-locus analyses revealed that PVE for pelvic
416 traits was lower than in the HEL × BYN cross; 14% and 16% for spine length and girdle
417 length, respectively, with <2% PVE due to dominance effects (Table 2). When analysing
418 relative trait values, one additional female QTL peak was found for both girdle and spine
419 lengths on LG1 (Supplementary Fig. 5, Table 1 and Supplementary Table 3). Due to the
420 large size of the identified QTL regions, it is not possible to know whether this QTL and
421 that on LG1 detected in the HEL × BYN cross are due to the same or different underlying
422 causal mutations (we consider this as a single QTL region).

423

424 *Trait correlations in the QTL crosses*

425 There was a strong correlation between relative pelvic spine and girdle lengths ($r^2 = 0.85$,
426 Fig. 1, Supplementary Fig. 2 and Supplementary Table 1) in the HEL × RYT cross as
427 expected, since pelvic reduction in this cross is controlled by a single large effect QTL
428 affecting both traits. However, in the HEL × BYN cross, correlation between relative

429 pelvic spine and girdle lengths was much weaker (r^2 = 0.11, Fig. 1, Supplementary Fig. 2
430 and Supplementary Table 1). This finding is consistent with pelvic spine and girdle
431 reductions being independently controlled by different QTL. Furthermore, of the 306 F_2
432 individuals, only four displayed complete lack of spines, despite the fact that the BYN
433 population is fixed for complete spine reduction in the wild (and the spine was absent in
434 the F_0 male). This is consistent with spine length being controlled by multiple additive
435 loci in the HEL × BYN cross. Among the F_2 individuals from the HEL × BYN cross,
436 relative girdle lengths were normally distributed with much smaller variances (SD =
437 0.012) compared to the HEL × RYT cross (SD = 0.041), with only two individuals with
438 reduced girdles (Fig. 1, Supplementary Fig. 2 and Supplementary Table 1). This is
439 consistent with the lower heritability of pelvic girdle lengths in the HEL × BYN
440 compared to the HEL × RYT cross, with contributions from many small effect loci. In the
441 HEL × PYÖ cross, phenotypic variation was comparable to that in the HEL × BYN cross
442 (SD = 0.012), although the mean relative spine length was slightly smaller (0.078 *vs*
443 0.091; Supplementary Table 1) and the mean for relative girdle length was slightly larger
444 (0.175 *vs* 0.169; Supplementary Fig. 2; Supplementary Table 1).
445
446 *Scanning for Pel deletions in the full-genome sequence data*

447 In the full-genome re-sequencing data of wild-collected individuals, a large deletion
448 upstream of *Pitx1* was fixed for all 21 individuals from Rytilampi, where pelvic reduction
449 maps to this region (Fig. 4). The deletion is around 3.5 kb in size and fully encompasses
450 the Pel-2.5kb^{SALR} region of the three-spined stickleback (Chan *et al*. 2010). No

451 comparable deletion was found in any other individuals in the dataset (Fig. 4b). This
452 included the two White Sea populations, in which either complete reduction of both the
453 pelvic spines and girdles was observed (MAS), or a putative large effect locus affecting
454 both spine and girdle length was found to be segregating (BOL; Fig. 1; Supplementary
455 Fig. 2 and Supplementary Table 1).

456

457 *Candidate genes*

458 Seven candidate genes or regulatory elements for pelvic reduction identified from the
459 literature (Supplementary Table 4) were found in the LGs with significant QTL (Fig. 2).
460 *Hif1a* (Mudie *et al*. 2014), *Pel* (Chan *et al*. 2010) and *Pou1f1* (Kelberman *et al*. 2009) are
461 known to regulate the expression of *Pitx1,* whereas four genes (*Fgf8*, *Wnt8c*, *Wnt8b* and
462 *Hoxd9*) are involved in the pelvic fin/hind limb development downstream of *Pitx1* (Don
463 *et al.* 2012; Tanaka *et al*. 2005). However, aside from *Pel* (LG7), only three – *Wnt8c*
464 (LG6), *Hif1a* (LG15) and *Pouf1f1* (LG16) – were clearly within the significant QTL
465 regions (Fig. 2). One candidate locus, *Hif1a*, is also on LG1 where significant QTL peaks
466 were found when analysing relative (but not absolute) trait values (Table 1 and
467 Supplementary Fig. 5).

468

469 *Simulations*

470 In the empirical data from the marine populations, IBD (the slope of the regression line
471 between linearised F_{ST} and geographic distance point distance in km) was 2.13^{-7} (95% of

472 the $1e^6$ bootstrap replicates were between 2.05^{-7} and 2.2^{-7}) for nine-spined sticklebacks,
473 and 1.53^{-8} (95% of the bootstrap replicates were between 1.10^{-8} and 1.96^{-8}) for three-
474 spined sticklebacks. Thus, the slope of the IBD regression line in this dataset was 13.9
475 times higher for nine-spined sticklebacks compared to three-spined sticklebacks (Fig. 5a).
476 The geographic distance (which is on an arbitrary scale) in the simulated data was scaled
477 based on the observed levels of IBD in the empirical data. This resulted in the distance
478 between the two simulated marine populations furthest away from each other (i.e. the
479 marine populations from which the focal freshwater populations were founded)
480 corresponding to 264 km and 352 km for three- and nine-spined sticklebacks,
481 respectively (Fig. 5a). Thus, with comparable levels of IBD as in the empirical data, our
482 simulations mimic the levels of parallel evolution that can be expected in three- and nine-
483 spined sticklebacks at relatively short geographic scales (<400 km). The difference in
484 IBD between the two species is also close to that in the empirical data (264/352 km =
485 0.75). In the empirical data, genetic differentiation between freshwater habitats for
486 populations <400 km from each other (mean F_{ST} = 0.19 and 0.49 for three- and nine-
487 spined sticklebacks, respectively) was also on par with the simulations (mean [across
488 simulation replicates] F_{ST} = 0.21 and 0.58, respectively). Notably, F_{ST} was >0.8
489 (linearised F_{ST}>4) between Pyöreälampi (no pelvic reduction) and Rytilampi (pelvic
490 reduction controlled by *Pitx1*/*Pel*), although these ponds are situated only 15 km from
491 each other (Fig. 5a; max and median F_{ST} = 0.96 and 0.62, respectively, for all pairwise
492 freshwater-freshwater comparisons). This is higher than the F_{ST} between any pair of
493 three-spined stickleback populations in the data (max = 0.78 and median = 0.26 for all

494 pairwise freshwater-freshwater comparisons).

495

496 *Do population demographic parameters influence local adaptation?*

497 In the simulations (Supplementary File 5), the relationship between marine-freshwater
498 F_{ST} and freshwater-freshwater F_{ST} depended on both the species and genetic architecture
499 (Fig. 5b). For instance, when the genetic architecture included one additive large effect
500 locus (architecture A; heterozygotes for this locus were 100% locally adapted to
501 freshwater), this locus was often involved in parallel evolution when population structure
502 was parameterised based on three-spined sticklebacks (65% of replicates, n = 100), but
503 not when population structure was parameterised based on nine-spined sticklebacks (3%
504 of the replicates). In 20% of the replicates for both species, the freshwater allele for this
505 locus was fixed in only one of the focal freshwater populations (i.e. local adaptation, but
506 not parallel evolution). When the trait under selection was controlled by several medium
507 effect loci (architecture B), parallel evolution was more common in both stickleback
508 scenarios, particularly for the locus with the largest allelic value (20% and 80% for nine-
509 and three-spined sticklebacks, respectively; Fig. 5b, c; heterozygotes for this locus were
510 60% locally adapted to freshwater). Local adaptation also occurred in 57% of the
511 replicates in the nine-spine stickleback-like scenario, and in 18% in the 3-sp (Fig. 5b, c).
512 For the locus with the second highest allelic value in architecture B (heterozygotes for
513 this locus were 40% locally adapted to freshwater), the cases of both parallel evolution
514 and local adaptation collectively dropped to 38% and 43% three and nine-spined like
515 stickleback scenarios, respectively. With one non-additive large effect locus with

516 recessive alleles locally adapted to freshwater (architecture C), this locus was less likely
517 to be involved in parallel evolution in the three-spined stickleback-like scenario (48%),
518 compared to architecture A. However, results for the nine-spined stickleback-like
519 scenario was similar to architecture A (6% parallel evolution; Fig. 5b, c). Thus, at
520 relatively short (<400 km) geographical distances, parallel evolution is an expected
521 outcome in three- but not nine-spined stickleback-like scenarios, particularly when a
522 single additive large effect locus is responsible for freshwater adaptation. Results
523 addressing the questions of how local adaptation depend on ancestral allele frequency are
524 presented in Supplementary File 7.
525

526 **Discussion**

527 The results demonstrate that pelvic reduction in nine-spined sticklebacks is not nearly as
528 common as in three-spined sticklebacks, and when it does exist, the genetic basis is more
529 variable in nine-spined sticklebacks compared to three-spined sticklebacks. This is
530 consistent with the re-analyses of neutral population genomic data showing that both IBD
531 in the sea and the genetic structure among marine populations are >10 times higher in
532 nine-spined sticklebacks, causing heterogeneity in the distribution of standing genetic
533 variation that is not present to the same extent in the marine three-spined stickleback
534 populations. Thus, the level and distribution of ancestral variation available for local
535 adaptation (and also for parallel phenotypic evolution) is likely a function of population
536 demographic parameters, with local adaptation being less likely in poorly connected
537 species, as suggested by Merilä (2013, 2014) and corroborated by our simulations.

538 However, other non-mutually exclusive factors, such as the genetic architecture (i.e.
539 dominance effects, heritability, mutation rates and numbers of causal loci involved), as
540 well as non-parallelism in phenotypic selection optima are also likely to play roles. In the
541 following, we discuss the possible causes of the discrepancy in pelvic structure
542 development between nine- and three-spined sticklebacks, as well as their implications to
543 our understanding of adaptive evolution in the wild.
544
545 *Can the genetic architecture of pelvic reduction be explained by population demographic*
546 *parameters?*
547 Together with earlier QTL studies (Shapiro *et al.* 2006, 2009; Shikano *et al.* 2013), we
548 show that multiple genomic regions (11 QTL, ten of which are novel to this study) are
549 associated with pelvic reduction in nine-spined sticklebacks across their distribution
550 range. Only one small effect QTL region (LG1; Table 1 and Supplementary Fig. 5) was
551 shared between any two crosses (HEL × BYN and HEL × PYÖ) in our study, but even
552 here it is not certain that the underlying causal mutations are the same. Although the high
553 frequency of *Pel* deletions (disrupting pelvic armour development) means that most of
554 the deletions associated with pelvic reduction in three-spined sticklebacks are
555 independently derived (Xie *et al.* 2019), *Pitx1*/*Pel* is nevertheless predominantly
556 responsible for pelvic reduction in three-spined sticklebacks (Chan *et al.* 2010; Xie *et al.*
557 2019). In contrast, the major effect and recessive *Pitx1*/*Pel* allele in nine-spined
558 sticklebacks is known to be responsible for pelvic reduction in only one Canadian
559 population and Rytilampi. In both Bynästjärnen and Pyöreälampi, pelvic reduction is less

560 heritable, polygenic and additive. In addition, based on morphological data, it seems
561 likely that major effect loci not associated with *Pel*-deletions (Fig. 4) control pelvic
562 reduction in two Russian ponds, with one additional Alaskan population being controlled
563 by a large effect additive locus mapping to LG4 (Shapiro *et al.* 2009). We do not yet
564 know the mutation rate of the *Pel*-deletions in nine-spined sticklebacks. Regardless,
565 pelvic reduction in the nine-spined stickleback is much less common than in three-spined
566 sticklebacks in general (Klepaker *et al.* 2013), and when it occurs, the QTL that control
567 pelvic reduction are much more variable with respect to heritability, effect size
568 distribution and dominance relationships. Assuming that selection for pelvic reduction in
569 nine-spined stickleback freshwater populations is universal (see below), such a pattern is
570 consistent with poor connectivity, causing heterogeneity in the patterns of standing
571 genetic variation. This would restrict local adaptation, and hence, also parallel phenotypic
572 evolution. However, the focus of this study is on the role of population demographic
573 parameters on local adaptation from standing genetic variation, regardless of whether the
574 alleles that confer freshwater adaptations are identical by descent (e.g. EDA) or
575 independently derived (e.g. *Pel*-deletions in three-spined sticklebacks), and only
576 secondarily on parallelism (i.e. when the same trait is locally adapted in multiple
577 independently colonised populations, regardless of genetic architecture).
578 It has been shown that multiple independent alleles for the same adaptation can
579 temporarily co-exist in structured populations, causing heterogeneity in standing genetic
580 variation across the distribution range of a species. This can be true when considering the
581 dynamics of novel mutations (Ralph and Coop 2010) or neutral alleles segregating in the

population before they become adaptive (Ralph and Coop 2015a). However, these studies assume uniform selection across single continuous populations and thus do not consider the scenario where locally adapted freshwater alleles – which can be different in different parts of the distribution range – are continually introduced to the sea. Furthermore, simulations showing that the probability of local parallel evolution depends on the migration rate from the sea (Galloway and Cresko 2019) or from other populations of the same habitat (Ralph and Coop 2015b) implicitly assume a homogenous pool of ancestral standing genetic variation. Our simulations show that local adaptation is also a function of the allele frequency in the founding marine population; with stronger IBD in the sea, standing genetic variation in the ancestral marine population was a stronger bottleneck for parameters that resulted in nine-spined stickleback-like population structure (9-sp) than for parameters that resulted in three-spined stickleback-like population structure (3-sp; Supplementary Figures 7 and 8). Particularly for smaller effect loci, there was a stronger dependence between the allele frequency in the ancestral marine population and local adaptation in 9- compared to 3-sp, indicating that smaller effect QTL have a stronger influence on local adaptation in 9- compared to 3-sp (Supplementary Figures 7 and 8). This resulted in more polygenic and/or less complete local adaptation in 9- compared to 3-sp (Supplementary Figures 7 and 8).

While our simulation results match closely with estimates of population structure and IBD in the empirical data, these simulations explore only a small proportion of the possible parameter space with respect to selection intensity, effective population sizes and migration rates (detailed in Supplementary File 5). We also assume that all parallel

604 evolution is due to standing genetic variation of alleles that are identical by descent (since
605 mutation rate is low), whereas it is known that pelvic reduction in three-spined
606 sticklebacks is, to a large extent, due to independently derived *Pel*-deletions attributable
607 to high mutation rates (Xie *et al.* 2019). Nevertheless, most three-spined stickleback
608 studies indicate that freshwater adaptation is indeed chiefly due to alleles identical by
609 descent that have segregated in the population for millions of years (e.g. Jones *et al.*
610 2012; Nelson and Cresko 2018), making *Pitx1* an exception. In the case of nine-spined
611 sticklebacks, *Pitx1* is with certainty associated with pelvic reduction in only two of all
612 studied populations, suggesting a minor role for recurrent mutation in determining pelvic
613 reduction in this species. While our simulations demonstrate that heterogeneous patterns
614 of standing genetic variation can limit local adaptation in species characterized by low
615 connectivity, further simulations with a wider parameter space could be helpful in
616 advancing our understanding of the limits of adaptation in small and structured
617 populations/species.
618
619 *Geographic heterogeneity in selection optima*
620 The high heterogeneity in the genetic architecture of pelvic reduction in the nine-spined
621 stickleback can alternatively be attributed to within habitat environmental variation
622 resulting in different selection optima in different pond populations (cf. Stuart *et al.* 2017;
623 Thompson *et al.* 2019). For example, the small differences between pelvic reduced
624 phenotypes in our study (e.g. in BYN and MAS/BOL spines are completely absent, while
625 in RYT they are only strongly reduced; Fig. 1, Supplementary Table 1 and

626 Supplementary Fig. 2) could in fact indicate different selection optima in the different
627 populations (Stuart *et al*. 2017; Thompson *et al*. 2019). It is possible to use available
628 phenotypic data to estimate the phenotypic optima of pelvic morphology (hypersphere) in
629 each of the populations using Fisher's geometric model (Stuart *et al.* 2017; Thompson *et*
630 *al*. 2019), where a strong overlap would suggest a higher probability of genetic
631 parallelism (Thompson *et al.* 2019). However, this assumes that the populations have
632 access to exactly the same ancestral variation and are free to evolve and reach their
633 optima, which is at odds with the results presented here. Without detailed environmental
634 data or direct estimates of strength of selection on pelvic phenotypes, disentangling the
635 effects of gene flow and within habitat environmental variation (assuming this leads to
636 non-parallel angles of selection) is not possible (Stuart *et al.* 2017).
637 In a recent simulation study, Thompson *et al.* (2019) showed that genetic
638 parallelism from standing genetic variation rapidly declines as selection changes from
639 fully parallel (optima angle of 0°) to divergent (optima angle of 180°), especially when
640 the trait is polygenic. However, although selection was fully parallel in our simulations,
641 we did not observe strong genetic parallelism for smaller effect loci (with allelic effects <
642 6) in both the three- and nine-spined stickleback-like scenarios (Fig. 5b, c). This suggests
643 that the effects of the underlying genetic architecture on parallelism (in conjunction with
644 some IBD and population structuring) can be independent of the angle of optimal
645 phenotypes between two habitats. In other words, non-parallelism in selection optima is
646 not necessary to explain non-parallelism at the phenotypic and genetic levels. Thus, it is
647 important not to disregard the population demographic setting as a factor that could

severely restrict heritability for adaptation and/or constrain adaptation to less optimal solutions. Evolutionary studies of species with population demographic parameters comparable to those typical for vulnerable or endangered species/populations, such as the nine-spined stickleback, would be valuable to gain a better understanding of how such species may respond to environmental changes and urbanisation (Thompson *et al.* 2018).

Pelvic reduction outside marine-freshwater study systems

While the evidence for genetic parallelism on large geographical scales in the marine-freshwater stickleback model system is extensive (Colosimo *et al.* 2005; Jones *et al.* 2012; Terekhanova *et al.* 2014, 2019; Nelson and Cresko 2018; Fang *et al*. 2019), the level of parallelism in lake-stream and pelagic-benthic ecotype pairs of three-spined sticklebacks is much more diverse (Peichel *et al*. 2001, 2017; Conte *et al*. 2012, 2015; Stuart *et al*. 2017). For instance, Conte *et al*. (2015) found that among benthic-limnetic three-spined stickleback pairs from Paxton and Priest lakes (Vancouver Island, BC, Canada), 76% of 42 phenotypic traits diverged in the same direction, whereas only 49% of the underlying QTL evolved in parallel in both lakes. For highly parallel traits in two other pairs of benthic-limnetic sticklebacks, only 32% of the underlying QTL were shared (Conte *et al*. 2012). Similarly, Stuart *et al*. (2017) found that among 11 evolutionary independent replicate pairs of lake-stream three-spined stickleback populations (Vancouver Island, BC, Canada), both within habitat variation and constraints to gene flow contributed to the observed variation in levels of phenotypic parallelism. Different lakes and streams likely do not have similar access to the same global pool of ancestral

670 variation as pairs of marine-freshwater three-spined sticklebacks, where gene flow in the
671 sea is high. This is consistent with the notion that more heterogeneous access to ancestral
672 variation can indeed limit genetic parallelism. This is also true for one example of
673 marine-freshwater three-spined stickleback divergence among isolated insular freshwater
674 populations in the Haida Gwaii archipelago off the northern Pacific coast of Canada
675 (Deagle *et al.* 2013). Here, similar to the nine-spined sticklebacks in this study, several
676 freshwater populations did not display any reduction in pelvic armour. However, those
677 populations that were fully plated were also genetically more similar to adjacent marine
678 individuals, suggesting that recent marine-freshwater admixture and/or selection
679 favouring plated freshwater individuals could explain this pattern. Thus, with respect to
680 access to ancestral variation available for freshwater adaptation, nine-spined sticklebacks
681 are likely closer to the three-spined stickleback lake-stream and benthic-limnetic study
682 systems than to the three-spined stickleback marine-freshwater study system. The only
683 notable exception is the lake-stream three-spined sticklebacks mentioned above, where
684 genetic structuring also is high.
685
686 *Epistatic control of pelvic reduction?*
687 For traits with more than one additive QTL of equal effect sizes, the F_2 phenotypes are
688 expected to be normally distributed with a mean close to that of the mean for the parents
689 (Klug and Cummings 2018). This was not the case for spine length in the HEL × BYN
690 cross, which was controlled by three QTL with similar effect sizes. In this case, the bulk
691 of the phenotypic values was centred around the mean, but also had a long tail of

692 individuals with strongly reduced pelvic spines (Fig. 1, Fig. 3 and Supplementary Fig. 2).
693 This skew in the phenotype distribution could be caused by epistatic interactions among
694 loci controlling pelvic spine length (Wolf *et al*. 2000). Consistent with this hypothesis,
695 complete spine reduction most likely occurred when the allele responsible for spine
696 reduction for the LG6 QTL was combined with at least one allele causing spine reduction
697 from the LG15 and LG16 QTL (Fig. 3b and Supplementary File 3). If a threshold number
698 of alleles are needed for complete pelvic reduction, this could also explain how standing
699 genetic variation in the sea is maintained, as the necessary multi-locus genotypes that
700 cause sub-optimal phenotypes in the sea are rarely formed, due to overall lower
701 frequencies of spine-reducing alleles in the sea. This is analogous to "epistatic shielding"
702 that can contribute to the persistence of disease alleles in populations (Phillips and
703 Johnsson 1998; Phillips 2008). Consistent with this possibility, LG6 of the F_0 female of
704 the HEL × BYN cross (from the sea) was polymorphic for the pelvic spine QTL effect
705 (Fig. 2) – evidently, a single pelvic-reducing allele alone in this female was not enough to
706 cause any pelvic reduction at all (this female had a complete pelvis).
707
708 *Candidate genes*
709 While the QTL peak for the *Pitx1*/*Pel* region in the HEL × RYT cross was narrow, this
710 was not the case for the other QTL we detected. Hence, due to the large QTL regions
711 detected by four-way single-mapping analyses, it was not meaningful to perform gene
712 ontology enrichment (GO) analyses – the QTL regions would have contained possibly
713 thousands of genes. Instead, we searched the literature for known candidate genes related

714 to pelvic reduction, and found three (excluding *Pitx1*/*Pel*) that were clearly contained
715 within the identified QTL regions (Fig. 2 and Supplementary Table 4). Due to the
716 aforementioned large QTL regions, these can only be considered as highly putative
717 candidate genes for pelvic reduction and will not be discussed further. However, further
718 studies of pelvic reduction might find these candidates worthy of attention.
719
720 *Conclusions*
721 Our results show that the repeated parallel reduction in pelvic structures in freshwater
722 populations of nine-spined sticklebacks is due to a diverse set of genetic changes: only
723 one small effect QTL for pelvic reduction was shared between the three experimental
724 crosses in this study. In one cross, pelvic reduction was mapped to the previously
725 identified *Pitx1*/*Pel* regions, but in the other two crosses, the genetic basis of pelvic
726 reduction was polygenic, and mapped to many different chromosomes. In addition to
727 these, yet another large effect QTL different from the *Pitx1*/*Pel* locus likely segregates in
728 one nine-spined stickleback population, which is yet to be identified. The results also
729 shed light on the possible drivers of the observed genetic heterogeneity underlying pelvic
730 reduction; as shown by simulations, heterogeneous genetic architectures are more likely
731 to emerge when access to ancestral variation is limited by strong isolation by distance and
732 population structuring. This reinforces the role of the nine-spined sticklebacks as a useful
733 model system, alongside the three-spined stickleback, to study adaptive evolution in the
734 wild. Furthermore, since the population demographic characteristics of nine-spined
735 sticklebacks are similar to small and endangered species/populations, it is also likely to

be a well-suited model to study the genetics of adaptation in populations of conservation concern.

Acknowledgments

We thank Alexandre Budria, Chris Eberlein, Sami Karja, Heini Natri and Ismo Rautiainen for help in fish collection and/or rearing. Thanks are also due to Kirsi Kähkönen for help in laboratory, and Oulanka Biological Station (University of Oulu) for logistic support. We are grateful for the computing resource support from CSC – the Finnish IT Center for Science Ltd administered by the Ministry of Education and Culture, Finland. B.G. thanks support from CAS Pioneer Hundred Talents Program and the National Natural Science Foundation of China (31672273). This work is supported by the Academy of Finland (grant numbers, 129662, 134728 and 218343 to J.M.), a grant from Helsinki Institute of Life Science (HiLIFE to JM) and a personal grant to Petri Kemppainen from the Finnish Cultural Foundation (00190489).

Literature cited

Arendt, J., and D. Reznick. 2007. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends Ecol. Evol. 23: 26–32. https://doi.org/10.1016/j.tree.2007.09.011

Barrett, R. D. H., and D. Schluter. 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. *23*: 38–44. https://doi.org/10.1016/j.tree.2007.09.008

Becker, W. A. 1984. Manual of quantitative genetics (Fourth ed.). Academic enterprises, Washington.

Bell, M. A., and S. A. Foster (Eds). 1994. The evolutionary biology of the threespine Stickleback. Oxford University Press, Oxford.

Bell, M. A., Khalef, V., and M. Travis. 2007. Directional asymmetry of pelvic vestiges in threespine stickleback. J. Exp. Zool. Part B, Mol. Dev. Evol. 308B: 189–199. https://doi.org/10.1002/jez.b

Bell, M. A., Orti, G., Walker, J. A., and J.P. Koenings. 1993. Evolution of pelvic reduction in threespine stickleback fish: a test of competing hypotheses. Evolution 47: 906–914. https://doi.org/10.2307/2410193

Bell, M. A., Windsor, E. A., and Buck, N. J. 2004. Twelve years of contemporary evolution in a threespine stickleback population. Evolution 58: 814–24. 10.1554/03-373.

Blouw, D. M., and G.J. Boyd. 1992. Inheritance of reduction, loss, and asymmetry of the pelvis in *Pungitius pungitius* (ninespine stickleback). Heredity, 68: 33–42. https://doi.org/10.1038/hdy.1992.4

Bolnick, D. I., Barrett, R. D. H., Oke, K. B., Rennison, D. J., and Y.E. Stuart. 2018. (Non) Parallel Evolution. Annu. Rev. Ecol. Evol. Syst. 12: 303–330.

Chan, Y. F., Marks, M. E., Jones, F. C., Jr, G. V., Shapiro, M. D., Brady, S. D., … D.M. Kingsley. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science 327: 302–305.

Cohan, F. M. 1984. Can uniform selection retard random genetic divergence between isolated conspecific populations? Evolution 38: 495–504. https://doi.org/10.2307/2408699

Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Dickson, H., Grimwood, J., … D.M. Kingsley. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science 307: 1928–1933. https://doi.org/10.1126/science.1107239

Conte, G. L., Arnegard, M. E., Best, J., Chan, Y. F., Jones, F. C., Kingsley, D. M., … C.L. Peichel. 2015. Extent of QTL reuse during repeated phenotypic divergence of sympatric threespine stickleback. Genetics 201: 1189–1200. https://doi.org/10.1534/genetics.115.182550

Conte, G. L., Arnegard, M. E., Peichel, C. L., and D. Schluter. 2012. The probability of genetic parallelism and convergence in natural populations. Proc. R. Soc B. 279: 5039–5047. https://doi.org/10.1098/rspb.2012.2146

Coyle, S. M., Huntingford, F. A., and C.L. Peichel. 2007. Parallel evolution of Pitx1 underlies pelvic reduction in Scottish Threespine Stickleback (Gasterosteus aculeatus). Genetics 98: 581–586. https://doi.org/10.1093/jhered/esm066

Cresko, W. A., Amores, A., Wilson, C., Murphy, J., Currey, M., Phillips, P., … J.H. Postlethwait. 2004. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. Proc. Natl. Acad. Sci. USA 101: 6050–6055.

Deagle, B. E., Jones, F., Chan, Y. F., Absher, D. M., Kingsley, D. M., and T. E. Reimchen. 2012. Population genomics of parallel phenotypic evolution in stickleback across stream–lake ecological transitions. Proc. R. Soc. B 279: 1277–1286 doi:10.1098/rspb.2011.1552

Deagle, B. E., Jones, F. C., Absher, D. M., Kingsley, D. M., and T. E. Reimchen. 2013. Phylogeography and adaptation genetics of stickleback from the Haida Gwaii archipelago revealed using genome-wide single nucleotide polymorphism genotyping. Mol. Ecol. 22: 1917–1932. https://doi.org/10.1111/mec.12215

DeFaveri, J., Shikano, T., Ghani, N. I. A., and J. Merilä. 2012. Contrasting population structures in two sympatric fishes in the Baltic Sea basin. Mar. Biol. 159: 1659–1672. https://doi.org/10.1007/s00227-012-1951-4

Don, E. K., Currie, P. D., and N. J. Cole. 2012. The evolutionary history of the development of the pelvic fin/hindlimb. J. Anatomy 222: 114–133. https://doi.org/10.1111/j.1469-7580.2012.01557.x

Elmer, K. R., and A. Meyer. 2011. Adaptation in the age of ecological genomics: Insights from parallelism and convergence. Trends Ecol. Evol. 26: 298–306. https://doi.org/10.1016/j.tree.2011.02.008

Fang, F., Kemppainen, P., Momigliano, P., and J. Merilä. 2019. Oceans apart: Heterogeneous patterns of parallel evolution in sticklebacks. BioRxiv, doi: https://doi.org/10.1101/826412

Galloway, J., Cresko, W., and Ralph, P. 2019. A few stickleback suffice for the transport of alleles to new lakes G3: Genes Genomes Genetics 10: 505–514. https://dx.doi.org/10.1534/g3.119.400564

Gibson, G. 2005. The synthesis and evolution of a super-model. Science 307: 1890–1891

Giles, N. 1983. The possible role of environmental calcium levels during the evolution of phenotypic diversity in outer Hebridean populations of the Three-spined stickleback,

Gasterosteus aculeatus. J. Zool. 199: 535–544. https://doi.org/10.1111/j.1469-7998.1983.tb05104.x

Hahne, F., and R. Ivanek. 2016. Statistical Genomics: Methods and Protocols. In Mathé E, Davis S (eds.), chapter Visualizing Genomic Data Using Gviz and Bioconductor, 335–351. Springer, New York.

Herczeg, G., Turtiainen, M., and J. Merilä. 2010. Morphological divergence of North-European nine-spined sticklebacks (*Pungitius pungitius*): signatures of parallel evolution. Biol. J. Linn. Soc. 101: 403–416.

Hermisson, J., and Pennings, P. (2005). Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics 169: 2335–2352. https://dx.doi.org/10.1534/genetics.104.036947

Hermisson, J., and P.S. Pennings. 2017. Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. Methods Ecol. Evol. 8: 700–716. https://doi.org/10.1111/2041-210X.12808

Hijmans, R.J., and J. van Etten. 2012. raster: Geographic analysis and modeling with raster data. R package version 2. 0-12. http://CRAN.R-project.org/package=raster

Karhunen, M., Merilä, J., Leinonen, T., Cano, J. M., and O. Ovaskainen. 2013. driftsel: An R package for detecting signals of natural selection in quantitative traits. Mol. Ecol. Res. 13: 746–754. https://doi.org/10.1111/1755-0998.12111

Karhunen, M., Ovaskainen, O., Herczeg, G., and Merilä, J. 2014. Bringing habitat information into statistical tests of local adaptation in quantitative traits: A case study of nine-spined sticklebacks. Evolution 68: 559–568.

Kelberman, D., Rizzoti, K., Lovell-Badge, R., Robinson I. C. A. F., and M. T. Dattani. 2009. Genetic regulation of pituitary gland development in human and mouse. Endocrine Reviews 30: 790–829, https://doi.org/10.1210/er.2009-0008

[dataset] Kemppainen, P., Li, Z., Rastas, P., Löytynoja, A., Fang, F., Yang, J., Guo, B., Shikano, T., and Merilä, M.; 2020; Genetic population structure constrains local adaptation in sticklebacks; Dryad; DOI https://doi.org/10.5061/dryad.76hdr7str.

Kemppainen, P., Knight, C. G., Sarma, D. K., Hlaing, T., Prakash, A., Maung Maung, Y. N., … C. Walton. 2015. Linkage disequilibrium network analysis (LDna) gives a global view of chromosomal inversions, local adaptation and geographic structure. Mol. Ecol. 15: 1031–1045. https://doi.org/10.1111/1755-0998.12369

856 Klepaker, T., Østbye, K., and M. A. Bell. 2013. Regressive evolution of the pelvic
857 complex in stickleback fishes: a study of convergent evolution. Evol. Ecol. Res. 15:
858 413–435.

859 Klug, W. S., and M.R. Cummings. 2018. Concepts of genetics. Pearson Education, Inc,
860 New Jersey.

861 Laine, V. N., Shikano, T., Herczeg, G., Vilkki, J., and J. Merilä. 2013. Quantitative trait
862 loci for growth and body size in the nine-spined stickleback *Pungitius pungitius* L.
863 Mol. Ecol. 22: 5861–5876. https://doi.org/10.1111/mec.12526

864 Leinonen, T., Cano, J. M., and J. Merilä. 2011. Genetics of body shape and armour
865 variation in threespine sticklebacks. J. Evol. Biol. 24: 206–218.
866 https://doi.org/10.1111/j.1420-9101.2010.02161.x

867 Lee, K., and Coop, G. 2017. Distinguishing among modes of convergent adaptation using
868 population genomic data. Genetics https://dx.doi.org/10.1534/genetics.117.300417

869 Li, H. 2018. Sequence analysis Minimap2 : pairwise alignment for nucleotide sequences.
870 Bioinformatics 34: 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

871 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … T. Sam. 2009.
872 The Sequence Alignment / Map format and SAMtools. Bioinformatics 25: 2078–
873 2079. https://doi.org/10.1093/bioinformatics/btp352

874 Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-
875 MEM. arXiv:1303.3997v2 [q-bio.GN]

876 Li, Z., Guo, B., Yang, J., Herczeg, G., Gonda, A., Balázs, G., … J. Merilä. 2017.
877 Deciphering the genomic architecture of the stickleback brain with a novel
878 multilocus gene-mapping approach. Mol. Ecol. 26: 1557–1575.
879 https://doi.org/10.1111/mec.14005

880 Li, Z., Kemppainen, P., Rastas, P., and J. Merilä. 2018. Linkage disequilibrium
881 clustering-based approach for association mapping with tightly linked genomewide
882 data. Mol. Ecol. Res. 18: 809–824. https://doi.org/10.1111/1755-0998.12893

883 MacPherson, A., and Nuismer, S. 2016. The probability of parallel genetic evolution
884 from standing genetic variation. J. Evol. Biol. 30: 326–337.
885 https://dx.doi.org/10.1111/jeb.13006

886 Merilä, J. 2013. Nine-spined stickleback (*Pungitius pungitius*): An emerging model for
887 evolutionary biology research. Ann. NY. Acad. Sci. 1289: 18–35.
888 https://doi.org/10.1111/nyas.12089

889 Merilä, J. 2014. Lakes and ponds as model systems to study parallel evolution. J. Limnol.
890 73: 33–45. https://doi.org/10.4081/jlimnol.2014.805

891 Mudie, S., Bandarra, D., Batie, M., Biddlestone, J., Ortmann, B., Shmakova, A., … S.
892 Rocha. 2014. PITX1, a specificity determinant in the HIF-1α- mediated
893 transcriptional response to hypoxia Sharon. Cell Cycle 13: 3878–3891.
894 https://doi.org/10.4161/15384101.2014.972889

895 Nelson, T. C., and W.A. Cresko. 2018. Ancient genomic variation underlies repeated
896 ecological adaptation in young stickleback populations. Evol. Lett. 2: 9–21.
897 https://doi.org/10.1002/evl3.37

898 Neuenschwander, S., Michaud, F., and J. Goudet. 2018. QuantiNemo 2: a Swiss knife to
899 simulate complex demographic and genetic scenarios, forward and backward in
900 time. Bioinformatics 35: 886–888. https://doi.org/10.1093/bioinformatics/bty737

901 Orr, H. A., and L. R. Unckless. 2008. Population extinction and the genetics of
902 adaptation. Am. Nat. 172:2, 160–169

903 Peichel, C. L., Nereng, K. S., Ohgi, K. A., Cole, B. L. E., Colosimo, P. F., Buerkle, C. A.,
904 …, and D.M. Kingsley. 2001. The genetic architecture of divergence between
905 threespine stickleback species. Nature 414: 901–905.

906 Phillips, P. C. 2008. Epistasis — the essential role of gene interactions in the structure
907 and evolution of genetic systems. Fundamental Concepts in Genetics 9: 855–867.
908 https://doi.org/10.1038/nrg2452

909 Phillips, P. C., and N. A. Johnson. 1998. The population genetics of synthetic lethals.
910 Genetics 150: 449–458.

911 Pritchard, J. R., and D. Schluter. 2001. Declining interspecific competition during
912 character displacement: Summoning the ghost of competition past. Evol. Ecol. Res.
913 3: 209–220.

914 Ralph, P., and Coop, G. 2010. Parallel adaptation: one or many waves of advance of an
915 advantageous allele? Genetics 186: 647–68.
916 https://dx.doi.org/10.1534/genetics.110.119594

917 Ralph, P., and G. Coop 2015a. Convergent evolution during local adaptation to patchy
918 landscapes. PLoS Genetics 11: e1005630.
919 https://dx.doi.org/10.1371/journal.pgen.1005630

920 Ralph, P., and G. Coop. 2015b. The role of standing variation in geographic convergent
921 adaptation. Am Nat 186 Suppl 1(S1), S5–23. https://dx.doi.org/10.1086/682948

922 Rastas, P. 2017. Lep-MAP3: Robust linkage mapping even for low-coverage whole
923 genome sequencing data. Bioinformatics 33: 3726–3732.
924 https://doi.org/10.1093/bioinformatics/btx494

925 Reid, S., Tibshirani, R., and J. Friedman. 2016. A study of error variance estimation in
926 lasso regression. Stat. Sinica 26: 35–67. https://www.jstor.org/stable/24721190

927 Reimchen, T. E. 1983. Structural relationships between spines and lateral plates in
928 threespine Stickleback (*Gasterosteus aculeatus*). Evolution 37: 931–946.
929 https://doi.org/10.2307/2408408

930 Reist, J. D. 1980. Predation upon pelvic phenotypes of brook stickleback, *Culaea*
931 *inconstancs*, by selected invertebrates. Can. J. Zool. 58: 1253–1258.

932 Rogers, S. M., Tamkee, P., Summers, B., Balabahadra, S., Marks, M., Kingsley, D. M.,
933 and D. Schluter. 2012. Genetic signatures of adaptive peakshift in threespine
934 sticklebacks. Evolution 66: 2439–2450.

935 Rosenblum, E. B., Parent, C. E., and E. E. Brandt. 2014. The molecular basis of
936 phenotypic convergence. Annu. Rev. Ecol. Syst. 45: 203–226. doi:10.1146/annurev-
937 ecolsys-120213-091851

938 Schluter, D., and G.L. Conte. 2009. Genetics and ecological speciation. Proc. Natl. Acad.
939 Sci. USA, 106: 9955–9962. https://doi.org/10.1073/pnas.0901264106

940 Schluter, D., Marchinko, K. B., Barrett, R. D., and S. M. Rogers. 2010. Natural selection
941 and the genetics of adaptation in threespine stickleback. Phil. Trans. R. Soc. Lond. B
942 Biol Sci. 365(1552): 2479–2486. doi:10.1098/rstb.2010.0036

943 Shapiro, M. D., Bell, M. A., and D.M. Kingsley. 2006. Parallel genetic origins of pelvic
944 reduction in vertebrates. Proc. Natl. Acad. Sci USA 103: 13753–13758.
945 https://doi.org/10.1073/pnas.0604706103

946 Shapiro, M. D, Marks, M. E., Peichel, C. L., Blackman, B. K., Nereng, K. S., Jo, B., …
947 and D. M. Kingsley. 2004. Genetic and developmental basis of evolutionary pelvic
948 reduction in threespine sticklebacks. Nature 428: 717–724.

Shapiro, Michael D, Summers, B. R., Balabhadra, S., Aldenhoven, J. T., Miller, A. L., Cunningham, C. B., … and D.M. Kingsley. 2009. The genetic architecture of skeletal convergence and sex determination in ninespine sticklebacks. Curr. Biol. 19: 1140–1145. https://doi.org/10.1016/j.cub.2009.05.029.The

Shikano, T., Laine, V. N., Herczeg, G., Vilkki, J., and J. Merilä. 2013. Genetic architecture of parallel pelvic reduction in ninespine sticklebacks. 3G: Genes, Genomes, Genetics, 3: 1833–1842. https://doi.org/10.1534/g3.113.007237

Shikano, T., Shimada, Y., Herczeg, G., and J. Merilä. 2010a. History vs. habitat type: Explaining the genetic structure of European nine-spined stickleback (*Pungitius pungitius*) populations. Mol. Ecol. 19: 1147–1161. https://doi.org/10.1111/j.1365-294X.2010.04553.x

Shikano, T., Ramadevi, J., & Merilä, J. 2010b. Identification of local-and habitat-dependent selection: scanning functionally important genes in nine-spined sticklebacks (*Pungitius pungitius*). Mol. Biol. Evol. 27: 2775–2789.

Sillanpää, M. J. 2011. On statistical methods for estimating heritability in wild populations. Mol. Ecol. 20: 1324–1332. https://doi.org/10.1111/j.1365-294X.2011.05021.x

Stern, D. L. 2013. The genetic causes of convergent evolution. Genetics 14: 751–764. https://doi.org/10.1038/nrg3483

Stuart, Y. E., Veen, T., Weber, J. N., Hanson, D., Ravinet, M., Lohman, B. K., … and D.I. Bolnick. 2017. Contrasting effects of environment and genetics generate a continuum of parallel evolution. Nature Ecol. Evol. 1: 0158. https://doi.org/10.1038/s41559-017-0158

Taggart, J. B., Hynes, R. A., and A. Ferguson. 1992. A simplified protocol for routine total DNA isolation from salmonid fishes. J. Fish. Biol 40: 963–965.

Tanaka, M., Hale, L. A., Amores, A., Yan, Y., Cresko, W. A., Suzuki, T., and J. H. Postlethwait. 2005. Developmental genetic basis for the evolution of pelvic fin loss in the pufferfish *Takifugu rubripes*. Dev. Biol. 281: 227–239. https://doi.org/10.1016/j.ydbio.2005.02.016

Terekhanova, N. V, Barmintseva, A. E., Kondrashov, A. S., Bazykin, G. A., and N. S. Mugue. 2019. Architecture of parallel adaptation in ten lacustrine threespine stickleback populations from the White Sea area. Genome Biol. Evol. 11: 2605–2618. https://doi.org/10.1093/gbe/evz175

982 Terekhanova, N. V, Logacheva, M. D., Penin, A. A., Neretina, T. V, Barmintseva, A. E.,
983 Bazykin, G. A., … N.S. Mugue. 2014. Fast evolution from precast bricks :
984 Genomics of young freshwater populations of threespine stickleback *Gasterosteus*
985 *aculeatus*. PLoS Genet. 10: e1004696. https://doi.org/10.1371/journal.pgen.1004696

986 Thompson, K. A., Osmond, M. M., and D. Schluter. 2019. Parallel genetic evolution and
987 speciation from standing variation. Evol. Lett. 3:129–141.
988 https://doi.org/10.1002/evl3.106

989 Thompson, K. A., Rieseberg, L. H., and D. Schluter. 2018. Speciation and the city.
990 Trends Ecol. Evol 33: 815–826.

991 Tibshirani, R. 1996. Regression shrinkage and selection via the Lasso. J. R. Stat. Soc.
992 Series. B 58: 267–288.

993 Varadharajan, S., Rastas, P., Löytynoja, A., Matschiner, M., Calboli, F. C. F., Guo, B., …
994 and J. Merilä. 2019. A high-quality assembly of the nine-spined stickleback
995 (Pungitius pungitius) genome. Genome Biol. Evol. 11: 3291–3308. evz240,
996 https://doi.org/10.1093/gbe/evz240

997 Van Ooijen, J. W. 2009. MapQTL v. 6.0: Software for the mapping of quantitative trait
998 loci in experimental populations of diploid species. Kyazma BV, Wageningen, The
999 Netherlands.

1000 Wang, Y., Zhao, Y., Wang, Y., Li, Z., Guo, B., and J. Merilä. 2020. Population
1001 transcriptomics reveals weak parallel genetic basis in repeated marine and
1002 freshwater divergence in nine-spined sticklebacks. Mol. Ecol.
1003 https://doi.org/10.1111/mec.15435

1004 Wolf, J. B., Brodie, E. D., and M. J. Wade. 2000. Epistasis and the Evolutionary Process.
1005 Oxford University Press, Oxford.

1006 Xie, K. T., Wang, G., Thompson, A. C., Wucherpfennig, J. I., Reimchen, T. E., Maccoll,
1007 A. D. C., … and P. Lake. 2019. DNA fragility in the parallel evolution of pelvic
1008 reduction in stickleback fish. Science 84: 81–84.

1009 Xu, S. 1996. Mapping quantitative trait loci using four-way crosses. Genet. Res. 68: 175–
1010 181.

1011 Xu, S. 2013a. Genetic mapping and genomic selection using recombination breakpoint
1012 data. Genetics 195: 1103–1115. https://doi.org/10.1534/genetics.113.155309

1013 Xu, S. 2013b. Principles of Statistical Genomics. Springer, New York.
1014 https://doi.org/DOI 10.1007/978-0-387-70807-2

1015 Yang, J., Guo, B., Shikano, T., Liu, X., and J. Merilä. 2016. Quantitative trait locus
1016 analysis of body shape divergence in nine-spined sticklebacks based on high-density
1017 SNP-panel. Sci. Rep. 6: 26632. https://doi.org/10.1038/srep26632

1018

1019 **Data Accessibility Statement**

1020 All data and codes used in analyses can be found in Dryad (DOI
1021 https://doi.org/10.5061/dryad.76hdr7str).

1022 **Table 1 | Summary of QTL-mapping analyses.** Each row corresponds to a significant
1023 (arbitrarily numbered) QTL region, with the proportion variance explained (PVE), jointly
1024 estimated for all PCs (one for each significant LD-cluster) extracted from such regions.
1025 Coding indicates whether the QTL was significant for the alleles inherited from the F_1
1026 female (♀), the F_1 male (♂) or the dominance effect (dom). Effect size (β) is based on the
1027 first PC extracted from all SNPs from each significant QTL region. "Std." indicates
1028 whether the trait was standardised (Yes) or not (No). P and P_{COR} represent nominal and
1029 corrected P-values from single-mapping four-way analyses, respectively. Results for
1030 standardised trait values are only shown for QTL that were not also significant for
1031 absolute trait values.

Cross	Trait	Coding	LG	QTL	Std.	P	P_{COR}	PVE_{TOT}	β
HEL × RYT	Spine	male	7	1	No	1.31e-04	0.019	0.285	1.654
HEL × RYT	Spine	female	7	1	No	1.26e-04	0.026	0.311	1.573
HEL × RYT	Girdle	male	7	1	No	3.66e-04	0.04	0.311	2.041
HEL × RYT	Girdle	female	7	1	No	1.08e-04	0.024	0.263	1.711
HEL × RYT	Spine	dom	7	1	No	1.08e-07	<0.001	0.213	1.432
HEL × RYT	Girdle	dom	7	1	No	1.35e-04	0.028	0.154	1.463
HEL × BYN	Spine	male	15	2	No	6.12e-06	0.001	0.095	0.54
HEL × BYN	Spine	male	16	3	No	1.16e-04	0.01	0.086	0.493
HEL × BYN	Spine	male	21	4	No	3.36e-04	0.031	0.037	0.36
HEL × BYN	Spine	female	6	5	No	1.37e-04	0.018	0.07	0.5
HEL × BYN	Girdle	male	14	6	No	2.49e-04	0.024	0.039	0.307
HEL × PYÖ	Spine	male	9	7	No	8.96e-11	<0.001	0.108	0.352
HEL × PYÖ	Girdle	male	19	8	No	1.56e-05	0.001	0.056	0.325
HEL × PYÖ	Girdle	female	4	9	No	1.32e-04	0.02	0.045	0.301
HEL × BYN	Spine	female	16	10	Yes	2.53e-04	0.03	0.047	0.007
HEL × BYN	Girdle	female	1	11	Yes	2.13e-05	0.002	0.077	0.006
HEL × BYN	Girdle	male	1	11	Yes	1.41e-04	0.016	0.026	0.005
HEL × PYÖ	Spine	male	1	13	Yes	4.75e-06	<0.001	0.075	0.005
HEL × PYÖ	Girdle	male	1	13	Yes	9.69e-05	0.016	0.044	0.005

1032

1033 **Table 2 | Proportion phenotypic variance explained (PVE) in pelvic traits.**
1034 Percentages of total phenotypic variance explained by different linkage groups (LG), by
1035 all SNPs (Tot), by loci inherited from females (♀) and males (♂), as well as the
1036 dominance effect (Dom). Results are shown for each cross and trait separately, and for
1037 absolute trait values.

LG	**HEL x RYT**		**HEL x BYN**		**HEL x PYÖ**	
	Spine	Girdle	Spine	Girdle	Spine	Girdle
1	-	0.01	0.03	0.05	-	-
2	-	-	0.01	0.03	-	0.02
3	-	-	-	0.01	-	-
4	-	0.02	0.01	0.01	-	0.03
5	-	-	-	-	-	-
6	-	-	0.03	-	0.02	-
7	0.85	0.82	0.02	0.02	-	-
8	-	-	0.01	0.01	-	-
9	-	-	-	0.02	0.11	-
10	-	-	-	0.06	-	-
11	-	-	0.02	-	-	-
12	0.02	0.02	0.03	0.01	0.01	-
13	-	-	-	0.01	-	0.02
14	-	-	-	0.05	0.01	0.02
15	-	-	0.09	0.01	-	-
16	-	-	0.12	0.01	-	-
17	-	-	0.01	0.01	-	-
18	-	-	-	0.04	-	0.01
19	-	-	0.02	0.06	-	0.05
20	-	-	-	0.01	-	-
21	-	-	0.03	-	0.01	-
Tot	0.8	0.76	0.39	0.32	0.14	0.16
♂	0.3	0.38	0.24	0.16	0.14	0.08
♀	0.29	0.27	0.17	0.06	-	0.09
Dom	0.23	0.17	0.01	0.13	0.01	-

1038

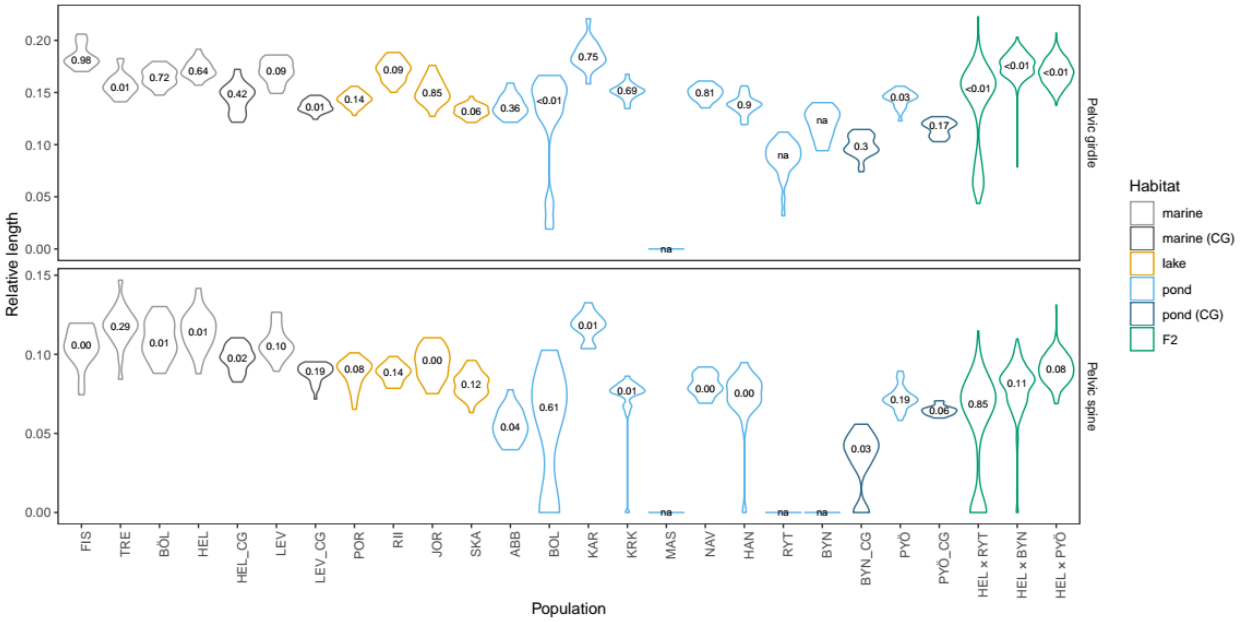

Figure 1 | Summary of previously published nine-spined stickleback pelvic phenotypes from the wild and common garden experiments. Violin plots depict relative (to standard body length) pelvic spine and girdle lengths according to population and habitat type. Numbers in the top panel show *P*-values for Pearson's product moment correlation between relative pelvic spine and relative pelvic girdle lengths, and numbers in the bottom panel show the respective squared correlation coefficients for each population/cross. "na" indicates data with no variation in spine or girdle lengths. Further details can be found in Supplementary Fig. 2 and Supplementary Table 1. Phenotypic data for F_2 individuals from the current study are also presented (HEL × RYT, HEL × BYN, and HEL × PYÖ) for comparison. Details of sample locations and data collection can be found from Herczeg *et al*. (2010) and Karhunen *et al*. (2013). Data from common garden experiments is indicated by the suffix "_CG" in the x-axis labels.

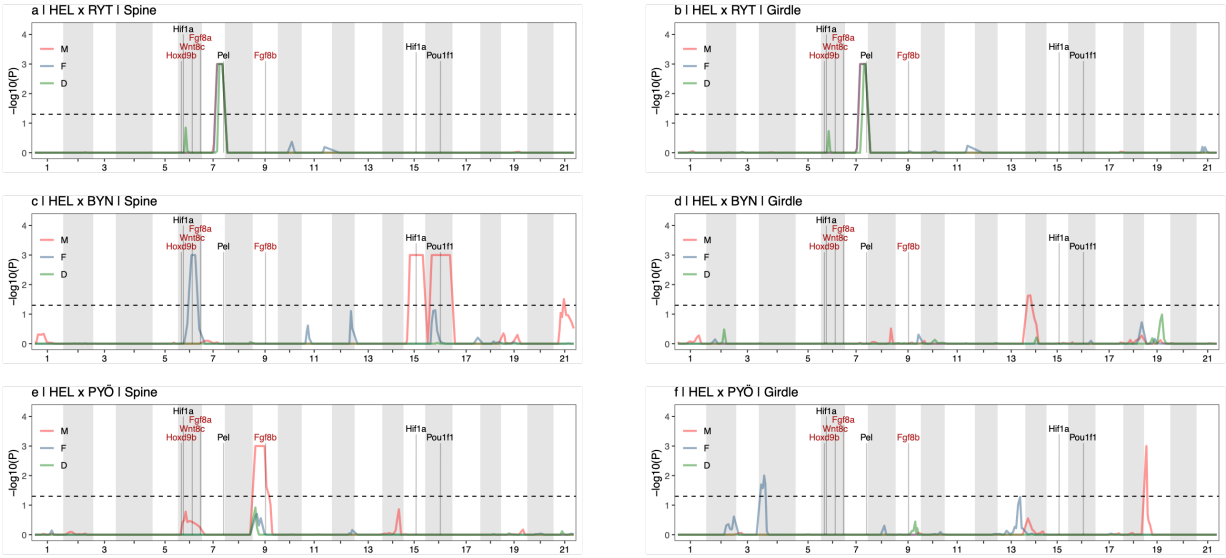

Figure 2 | Quantitative trait locus mapping of pelvic reduction in three independent stickleback crosses. Single-mapping four-way analyses of four morphological traits associated with pelvic reduction in (a-b) HEL × RYT cross, (c-d) HEL × BYN cross, and (e-f) HEL × PYÖ cross. QTL for pelvic spine length and girdle length are shown, with the x-axis indicating position in centi Morgans (cM). Results are based on permutations, and the dotted vertical line indicates genome wide significance at $\alpha = 0.05$. Results are shown separately for alleles inherited from the male F_1 (M), the female F_1 (F), together with the dominance effect (D) according to the legend. Absence of a dominance effect indicates that the trait inheritance is additive, whereas a peak only for M or F indicates that the allelic effect was segregating in the F_0 male or the F_0 female, respectively (see Supplementary File 1 for details). Candidate genes involved in pelvic development are indicated with black text representing genes that affect expression of *Pitx1*, and red text indicates genes that affect pelvic development downstream of *Pitx1* expression. Results for analyses based on relative trait values can be found in Supplementary Figure 5.

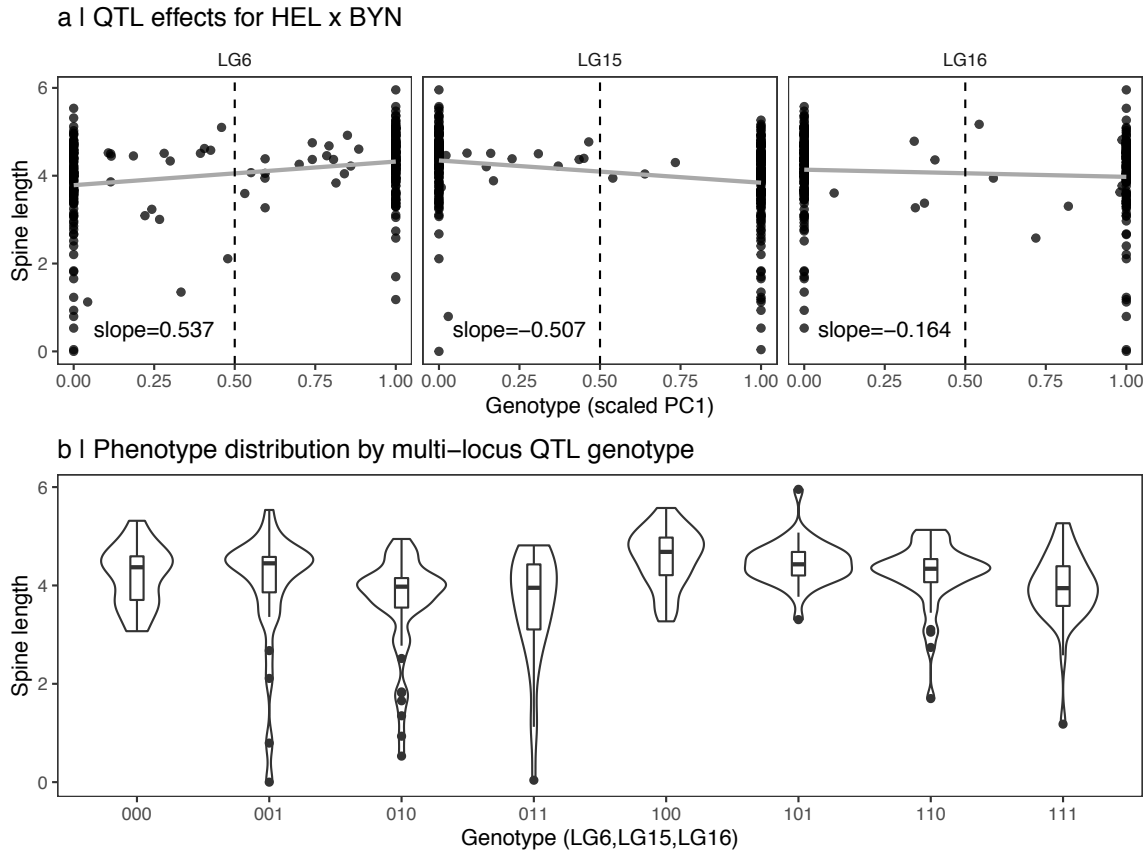

Figure 3 | Epistatic interactions in pelvic spine development for HEL × BYN cross. (a) The effects of individual genotype, where the genotype is given by the first PC (scaled between 0 and 1) from the cluster of SNPs that was the most significant for spine length on LG6, LG15 and LG16 (see Fig. 2), respectively. Genotypes were further based on the genotypes [x_{dij}, x_{sij}] depending on which of these were significant for the QTL effects (x_{dij} for LG6 and x_{sij} for LG15 and LG16). Some individuals have genotypes between 0 and 1 only because the genotype is based on the first PC of large sets of highly but not perfectly correlated SNPs. Slope of the regression line (grey) is shown. (b) distribution of spine lengths for all multi-locus genotypes from (a). The multi-locus genotype was based on rounding PC1 coordinates from (a) where values below 0.5 (left of vertical dashed line) were considered as "Allele 0", and those above or equal to 0.5 were considered as "Allele 1". The first digit for genotypes in (b) thus represents the genotype of the QTL on LG6, followed by LG15 and LG16, respectively.

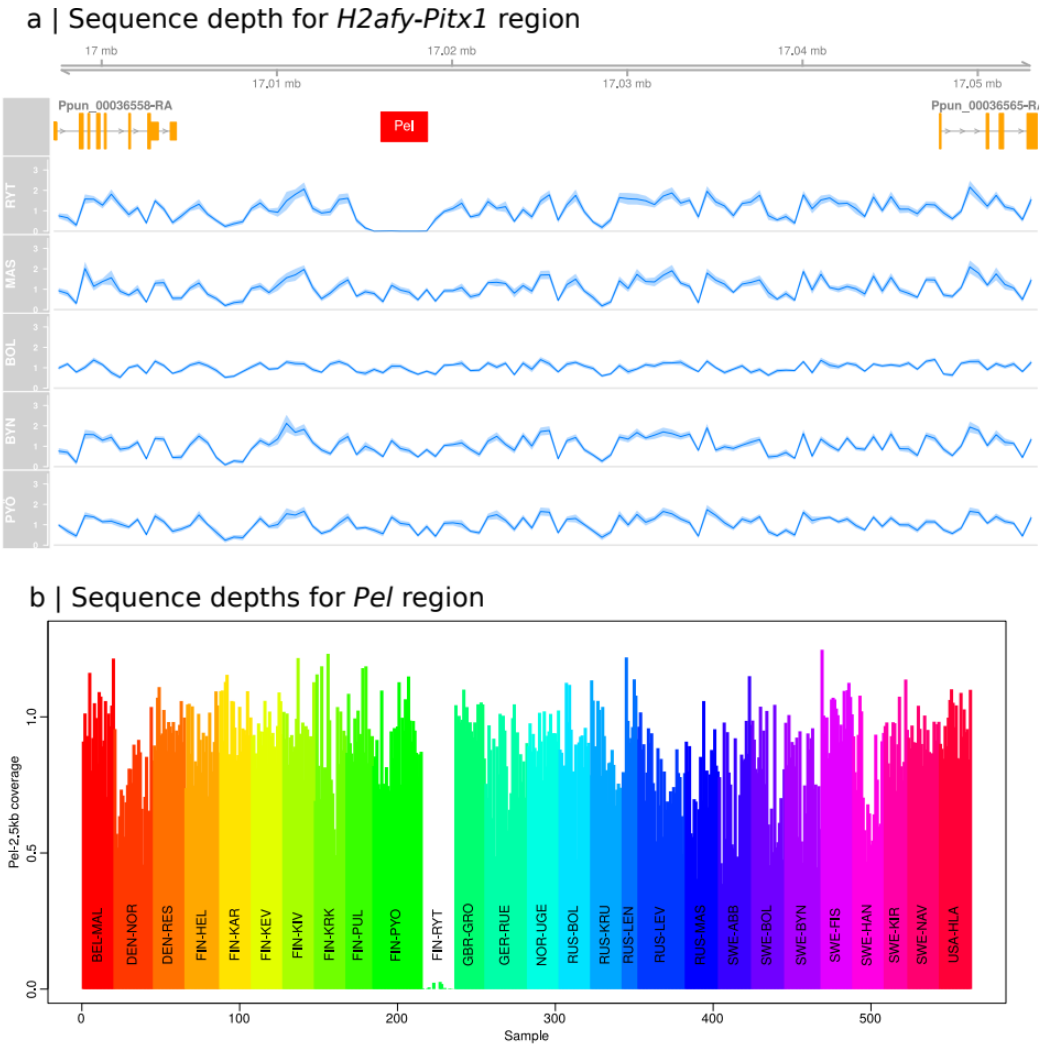

Figure 4 | Sequence depths for *H2afy-Pitx1* region. (a) Relative sequencing depth across the *H2afy-Pitx1* intergenic region for 10 individuals from five different populations. Blue line and shading indicate population average and 95% confidence intervals, respectively. The 3.5kb deletion in RYT, seen as a dip in sequencing depth, fully encloses the Pel-2.5kb^{SALR} of three-spined stickleback indicated by the red box. (b) Normalised sequencing depths for the Pel-2.5kb^{SALR} region for 563 samples from 27 populations, showing that the complete deletion of the *Pel* region is unique to RYT. Phenotypic data in the Russian populations MAS and BOL suggest that a large effect locus is responsible for pelvic reductions, and BYN and PYÖ (analysed here) are populations in which pelvic reduction does not map to LG7.

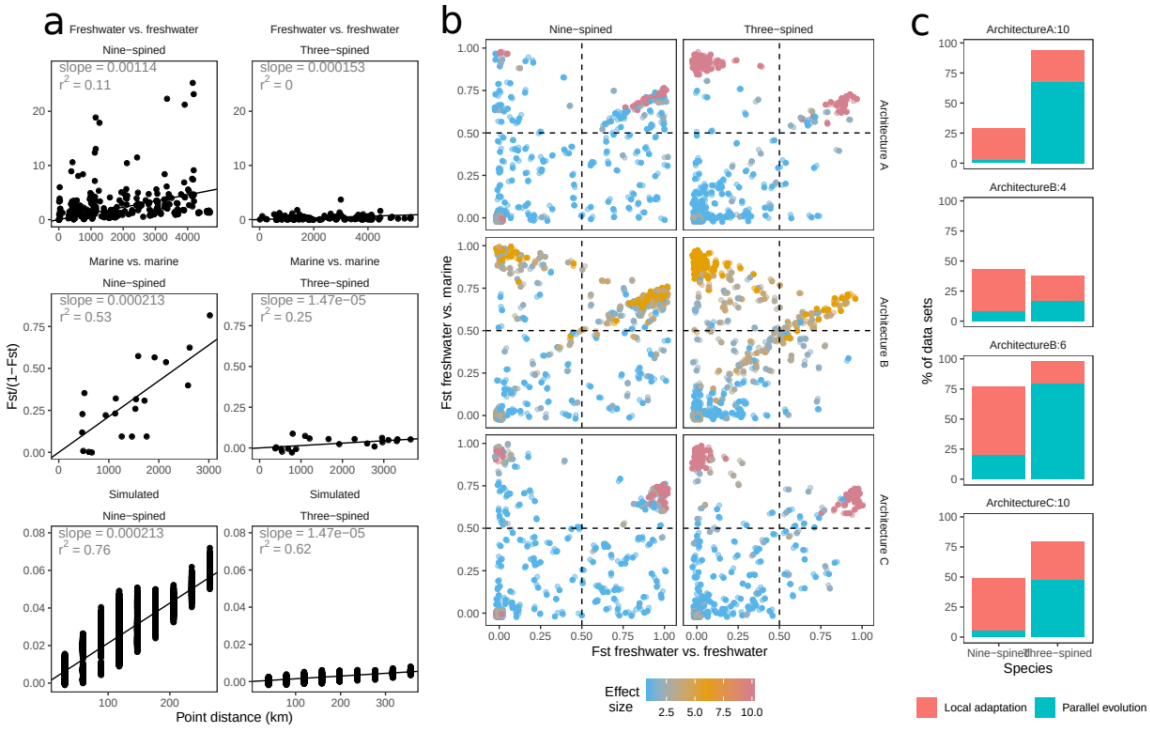

Figure 5 | Simulation results. Linearised F_{ST} against geographic distance (IBD) for empirical and simulated data (a), with slope and squared Pearson's product moment correlation coefficient indicated. Geographic distance for simulated data (IBD in the sea) is scaled to match the slope for the IBD-plot in the sea in the empirical data. Freshwater-freshwater F_{ST} against marine-freshwater F_{ST} from all QTL from all simulated data (n = 100), with effect sizes indicated as shown in legend (b). Loci in the upper left quadrant are classified as being involved in parallel evolution, and loci in the upper right quadrant are loci that are involved in local adaptation in only one freshwater population. This data is summarised in (c) focusing on the four largest effect loci, with genetic architecture and effect sizes indicated by the figure titles.