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[1,6]

[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

[6]Division of Ecology and Biodiversity, Kadoorie Building, The University of Hong Kong, Pokfulam Road, Hong Kong, SAR

*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 >2x stronger population structuring among freshwater populations and >10x stronger genetic isolation by distance in the sea in nine-spined sticklebacks as compared to three-spined sticklebacks.

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). Epistatic interactions and the order at which beneficial alleles enter and become established in a population also play important roles in determining the genetic architecture of the trait in question (Orr 1998). Thus, the demographic history of populations likely plays a central role in determining the

64 likelihood of local adaptation, and hence also parallel phenotypic evolution.

65 In recent years, the nine-spined stickleback (*Pungitius pungitius*) has been
66 emerging as a model system for studying adaptive evolution in the wild (Merilä 2013)
67 alongside the more widely studied three-spined stickleback (*Gasterosteus aculeatus*; Bell
68 and Foster 1994; Gibson 2005). The latter species' ability to rapidly adapt to local
69 environmental conditions has often been shown to stem from a global pool of ancestral
70 standing genetic variation (Schluter and Conte 2009; Jones *et al*. 2012; Terekhanova *et*
71 *al.* 2014, 2019). The nine-spined stickleback differs from the three-spined stickleback by
72 having smaller effective population sizes (N_e), reduced gene flow in the sea, and a
73 tendency to occur in small landlocked ponds in complete isolation from other populations
74 (Shikano *et al.* 2010a; DeFaveri *et al*. 2012; Merilä 2013; this study). Given their
75 contrasting population demographic characteristics, three- and nine-spined sticklebacks
76 can thus be expected to respond differently with respect to local adaption to newly
77 colonised freshwater habitats.

78 Regressive evolution of the pelvic complex has occurred in freshwater
79 populations of three (*viz*., *Gasterosteus*, *Pungitius*, and *Culaea*) of the five recognised
80 stickleback genera since the last glacial period (reviewed in Klepaker *et al*. 2013). While
81 marine populations of three- and nine-spined sticklebacks usually have complete pelvic
82 structures with fully developed pelvic girdles and lateral pelvic spines, partial or even
83 complete pelvic reduction is common in freshwater populations (Blouw and Boyd 1992;
84 Shapiro *et al*. 2004, 2006; Herczeg *et al*. 2010; Klepaker *et al*. 2013). Several factors may
85 contribute to this, including the absence of gape-limited predatory fish and limited

86 calcium availability, as well as the presence of certain insect predators (Reist *et al*. 1980;
87 Reimchen *et al*. 1983; Giles 1983; Bell *et al*. 1993; Karhunen *et al*. 2013; Chan *et al*.
88 2010) in freshwater environments. Collectively, sticklebacks provide an important model
89 system to study the genetic mechanisms underlying the adaptive parallel pelvic reduction
90 at both intra- and inter-specific levels, under a wide range of population demographic
91 settings. However, studies of parallel patterns of marine-freshwater divergence in nine-
92 spined sticklebacks are still scarce (Herczeg *et al*. 2010; Shikano *et al*. 2010b; Wang *et*
93 *al*. 2020), especially at the genetic level, precluding any comprehensive and conclusive
94 comparison of the two species.

95 The genetic basis of pelvic reduction in the three-spined stickleback is well
96 understood. Quantitative trait locus (QTL) mapping studies have identified a single
97 chromosomal region containing the gene *Pituitary homeobox transcription factor 1*
98 (*Pitx1*) that explains more than two thirds of the variance in pelvic size in crosses
99 between individuals with complete pelvic girdles and spines, and pelvic-reduced
100 individuals (Cresko *et al*. 2004; Shapiro *et al*. 2004; Coyle *et al*. 2007). Pelvic loss in the
101 marine-freshwater three-spined stickleback model system (as well as in benthic-limnetic
102 pairs of three-spined sticklebacks; Peichel *et al.* 2001) is predominantly caused by
103 expression changes of the *Pitx1* gene due to recurrent deletions of the pelvic enhancer
104 (*Pel*) upstream of *Pitx1* (Chan *et al*. 2010; Xie *et al.* 2019; in lake-stream pairs of three-
105 spined sticklebacks, the genetic architecture of pelvic reduction is more diversified;
106 Peichel and Marques 2017; Deagle *et al*. 2012; Stuart *et al.* 2017; Rennison *et al.* 2019).
107 While pelvic reduction in freshwater is much less common in nine- than three-spined

108 sticklebacks (Klepaker *et al*. 2013; Fig. 1), two independent QTL studies also identified
109 *Pitx1* in linkage group (LG) 7 as a major cause for pelvic reduction in a Canadian
110 (Shapiro *et al.* 2006) and a Finnish (Shikano *et al.* 2013) population. Another large effect
111 region in LG4 was found to be associated with pelvic reduction in an Alaskan population
112 (Shapiro *et al.* 2009). Similar to the three-spined stickleback, pelvic spine and pelvic
113 girdle sizes are strongly correlated in the population from the Finnish pond Rytilampi,
114 since *Pitx1* controls both phenotypes (Shikano *et al.* 2013; Fig. 1 and Supplementary
115 Table 1). In contrast, a considerable amount of phenotypic variation with respect to these
116 traits and their inter-correlations exists among different freshwater pond populations in
117 northern Europe (Herczeg *et al.* 2010; Karhunen *et al.* 2013, 2014; Fig. 1). Given their
118 high heritability (Blouw and Boyd 1992; Leinonen *et al.* 2011), the lack of correlation
119 between spine and girdle lengths suggests that they can be independently controlled by
120 different QTL. Thus, the genetic underpinnings of pelvic reduction in nine-spined
121 sticklebacks (when it occurs) appear to be more diversified than those in the marine-
122 freshwater three-spined stickleback model system (Merilä 2013, 2014).
123 Starting with a survey of previously published phenotypic data on pelvic
124 reduction in the wild, we aimed to investigate the possible genetic heterogeneity
125 underlying pelvic reduction in different nine-spined stickleback populations by mapping
126 QTL for pelvic reduction in three independently colonised ponds. One was the previously
127 studied Rytilampi (earlier analysed only with microsatellites, Shikano *et al.* 2013), now
128 re-analysed along with two new populations (Bynästjärnen and Pyöreälampi) using >75
129 000 SNPs. These data were subjected to a cutting-edge mapping approach (Li Z. *et al.*

130 2017, 2018) that can provide more information on the source of the QTL effects than has
131 been previously possible. To further assess the extent to which *Pel* could be responsible
132 for pelvic reduction in the nine-spined stickleback, we screened the whole genomes of
133 individuals from 27 wild populations for deletions in the genomic region spanning the *Pel*
134 element and the *Pitx1* gene. Finally, utilising more comprehensive geographic sampling
135 than in previous studies (Shikano *et al.* 2010a; DeFaveri *et al*. 2012; Merilä 2013), along
136 with high-quality SNP data, we re-assessed the differences in population structuring and
137 genetic isolation by distance (IBD) between nine-and three-spined sticklebacks. We
138 hypothesised that, in contrast to three-spined sticklebacks, both the scarcity and
139 variability in the genetic architecture of pelvic reduction in nine-spined sticklebacks is
140 due to heterogeneity in the spatial patterns of standing genetic variation caused by strong
141 population structuring – both among adjacent pond populations as well as in the ancestral
142 sea population(s). Building on previous theoretical work, this hypothesis was tested using
143 simulated data parameterised on population demographic parameters obtained from the
144 empirical data.
145
146 **Materials and Methods**
147 *Fish collection, crossing, and rearing*
148 For the QTL crosses, three different marine F_0 generation females from the Baltic Sea
149 (Helsinki, Finland; 60°13′N, 25°11′E) were crossed with a freshwater F_0 generation male
150 from either Bynästjärnen (64°27′N, 19°26′E), Pyöreälampi (66°15′N, 29°26′E) or
151 Rytilampi (66°23′N, 29°19′E) ponds. Fish crossing, rearing, and sampling followed the

experimental protocol used in the earlier study of the Rytilampi population (Shikano *et al.* 2013; Laine *et al.* 2013). For Rytilampi, the F_0 generation fish were artificially mated in the lab during the early breeding season of 2006 (Shikano *et al.* 2013), and the resulting full-sib F_1-offspring were group-reared in aquaria until mature, as explained in Shikano *et al.* (2013). Two randomly chosen F_1 individuals were mated repeatedly (seven different clutches) to produce the F_2 generation between September and October 2008. The offspring were reared in separate aquaria. The same procedure was followed for Pyöreälampi (F_0 mating: Jun 2011; F_1 mating: Jul–Sep 2012; F_2 rearing: Jul 2012–Apr 2013; 8 different clutches) and Bynästjärnen (F_0 mating: Jun 2011; F_1 mating: Nov 2013–Jan 2014, F_2 rearing: Nov 2013–Aug 2014; 6 different clutches). The F_2 fish were euthanized at 187, 238, and 238 days post-hatch for Rytilampi, Pyöreälampi and Bynästjärnen, respectively. At this point, the fish were on average 52.3 mm in standard length (Rytilampi = 48.6 mm; Pyöreälampi = 52.3 mm; Bynästjärnen = 53.6 mm), and weighed on average 1.34 g (wet weight; Rytilampi = 1.07 g; Pyöreälampi = 1.49 g; Bynästjärnen = 1.44 g). In total, 308, 283 and 279 F_2 offspring were available for analyses from Helsinki × Bynästjärnen (HEL × BYN), Helsinki × Pyöreälampi (HEL × PYÖ) and Helsinki × Rytilampi (HEL × RYT) crosses, respectively.

The experiments were conducted under licenses from the Finnish National Animal Experiment Board (#STH379A and #STH223A). Experimental protocols were approved by the ethics committee of the University of Helsinki, and all experiments were performed in accordance with relevant guidelines and regulations.

174 *Morphological data*

175 To visualise bony elements, all of the F_2-progeny were stained with Alizarin Red S
176 following Pritchard and Schluter (2001). Spine lengths was measured from its distal point
177 to the basal point and the length of branch and girdle was measured at its maximum linear
178 length with digital callipers to the nearest 0.01 mm (both sides of the body). Landmark
179 definitions with illustrations depicting them can be found from Herczeg *et al*. (2010) and
180 Shapiro et al. (2004). Although it is known that the left-right asymmetry of the pelvic
181 girdle is heritable in sticklebacks (Blouw and Boyd 1992; Bell *et al.* 2007; Coyle *et al.*
182 2007), we did not specifically study this. To reduce the number of tests, the mean of the
183 left and right-side measurements was used (analyses conducted for the two sides
184 separately always yielded qualitatively similar results as the tests conducted with the
185 averages; results not shown). All measurements were made by the same person twice
186 with repeatability >0.97 for all traits (R; Becker 1984). The QTL analyses were
187 performed on both absolute and relative (scaled to the total body length) trait values, but
188 for all of the analyses that compared phenotypic data between populations (which also
189 differ in body sizes), only relative trait values were used. Previously published
190 phenotypic data from 19 wild populations were obtained from Herczeg *et al*. (2010) and
191 Karhunen *et al.* (2013). These included data on pelvic spine and girdle lengths of wild-
192 collected individuals from ten pond populations (Abbortjärn = ABB, Bolotjone = BOL,
193 Karilampi = KAR, Kirkasvetinen lampi = KRK, Mashinnoje = MAS, Lil-Navartjärn =
194 NAV, Hansmytjärn = HAN, Rytilampi = RYT, Bynästjärnen = BYN), four lake
195 populations (Iso Porontima = POR, Riikojärvi = RII, Joortilojärvi = JOR, Västre-

196 Skavträsket = SKA) and five marine populations (Fiskebäckskil = FIS, Trelleborg = TRE,
197 Bölesviken = BÖL, Helsinki = HEL LEV = Levin Navolak), as well data on common
198 garden-reared F_1 generation individuals from two marine (HEL and LEV) and two pond
199 (BYN and PYÖ) populations. Visibly broken spines were treated as missing data. A map
200 showing the geographic location of these populations is provided in Supplementary
201 Figure 1.
202
203 *Genotyping and linkage map construction*
204 The RAD sequencing protocol used to obtain the SNP data was the same as in Yang *et al.*
205 (2016) and Li Z. *et al*. (2017). In short, genomic DNA was extracted from ethanol
206 preserved fin clips using the phenol-chloroform method (Taggart *et al*. 1992). DNA was
207 fragmented with PstI restriction enzyme and the resulting 300–500 bp long fragments
208 were gel purified. Illumina sequencing adaptors and library specific barcodes were
209 ligated to the digested DNA fragments, and the barcoded RAD samples were pooled and
210 sequenced on 24 lanes of the Illumina HiSeq2000 platform with 45 bp single-end
211 strategy. RAD library construction and sequencing were conducted by BGI
212 HONGKONG CO., LIMITED. After eliminating adapters and barcodes from reads, a
213 quality check was done using FastQC
214 (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/).

215 A linkage map based on the three crosses was constructed using Lep-MAP3
216 (LM3; Rastas 2017), as described in detail in Li H. *et al.* (2018). LM3 can infer the
217 parental/grandparental phase based on dense SNP data, which allowed us to utilise the

218 four-way cross QTL mapping method described below. Input data were generated by first
219 mapping individual fastq files to the nine-spined stickleback reference genome using
220 BWA mem (Li H. 2013) and SAMtools (mpileup; Li H. *et al.* 2009), followed by
221 pileupParser.awk and pileup2posterior.awk scripts from the LM3 pipeline using default
222 parameters (see Supplementary File 1 for details).
223
224 *Dimensionality reduction by linkage disequilibrium network clustering*
225 It is essential in QTL mapping to correct for multiple testing in order to reduce the rate of
226 false positives. Moreover, in large genomic datasets, physically adjacent SNPs –
227 particularly those from experimental crosses – are often in linkage disequilibrium (LD),
228 i.e. correlated. Since a group of SNPs in high LD explains similar amounts of genetic
229 variation in a given trait, it is reasonable to apply a dimensional reduction procedure
230 before QTL-mapping to exclude the redundant information from the data. Here we used
231 linkage disequilibrium (LD) network clustering (LDn-clustering) and PC regression as
232 dimensionality reduction tools prior to single-locus QTL mapping (Li Z. *et al.* 2018). The
233 first step of this approach involves an extension of a method developed by Kemppainen
234 *et al.* (2015), which uses LD network analysis for grouping loci connected by high LD.
235 The second step involves principal component analysis (PCA) as a method for
236 dimensionality reduction in each cluster of loci connected by high LD (LD-clusters). This
237 was achieved by the function LDnClustering in the R-package LDna (v.0.64; Li Z. *et al.*
238 2018). The method used here differs slightly from the original method described in Li Z.
239 *et al.* (2018) to increase efficiency in both computational speed and complexity reduction

(see Supplementary File 2 for details).

QTL mapping: four-way crosses model

In some circumstances, such as a four-way cross (Xu 1996), F_1 hybrids of two heterozygous parents (Van Ooijen 2009), and an outbred F_2 design (Xu 2013a), it is possible that up to four different alleles – A and B from the dam, and C and D from the sire – segregate in the population. In such cases, a linear regression model for QTL analysis of the outbred F_2 data (Li Z. *et al.* 2018) is defined by:

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

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

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

(Xu 2013b), β_0 is the parameter of the population mean, β_{dj} is the substitution effect of alleles A and B of the dam (i.e. the grandfather in F_0) at the locus j (j=1,…,p; p is the total number of SNPs), β_{dj} is the substitution effect of alleles C and D of the sire (i.e. the grandmother in F_0), γ_j is the dominance effect, and ε_i is the residual error term mutually following an independent normal distribution.

The model (1) requires the knowledge of the grandparental phase (produced by LM3) with the benefit that the source (*viz*. F_0 female or F_0 male) of the observed QTL

effect can be inferred, as described in more detail in Supplementary File 3. A multiple correction on the basis of permutation tests was further conducted to control for false positives due to multiplicity (Li Z. *et al.* 2017) with 1×10^5replicates.

Estimating the proportion of phenotypic variance explained by SNPs

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

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

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

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

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

where $\hat{\beta}_j$ is the effects of the SNPs, and σ_0^2 is the LASSO residual variance estimated by a cross-validation-based estimator introduced by Reid *et al*. (2016). The PVE explained by

279 each linkage group was estimated on the basis of the LASSO estimates using the

280 following formula:

281 $$PVE_{LG} \approx PVE_{total} - \frac{var\left(\sum x_j \hat{\beta}_j\right)}{var(y)}, (5)$$

282 where $\hat{\beta}_j (j \notin G)$ represents all the effects estimated by the LASSO of the SNPs that do not

283 belong to a set of SNPs (e.g. to a chromosome/linkage group). A similar approach was

284 used to estimate the contribution of grandparental alleles and to evaluate the dominance

285 component by treating the coding systems [x_{dij}, x_{sij}, z_{ij}] (2) as different groups of SNPs.

286

287 *Scanning for Pel deletions in full-genome sequence data*

288 A minimum of 20 samples from populations RYT, MAS, BOL, BYN and PYÖ, and 10–

289 31 samples from an additional 22 populations (from Northern Europe and USA;

290 Supplementary Fig. 1 and Supplementary Table 2) were sequenced to 10× coverage by

291 BGI HONGKONG CO., LIMITED. Reads were mapped to the nine-spined stickleback

292 reference genome (Varadharajan *et al*. 2019) with BWA mem (Li H. 2013), and site-wise

293 sequencing coverages were computed with SAMtools (depth; Li H. *et al*. 2009). Relative

294 sequence depths across the *H2afy-Pitx1* intergenic region were estimated for the five

295 focal populations by first computing the median depths for 1000 bp sliding windows, and

296 then normalising these by the median depth for the full intergenic region. The Pel-

297 2.5kb^{SALR} region was extracted from the original BAC contig (GenBank accession

298 number GU130433.1) and mapped to the nine-spined stickleback intergenic region with

299 minimap2 (Li H. 2018). The sequencing depths for the Pel-2.5kb^{SALR} were normalised by

300 dividing the mean depths of the *Pel* region by the mean depths of the full intergenic
301 region. Gene annotations and relative sequencing depths (average and confidence
302 intervals) were computed and visualised using the R-package Gviz (Hahne and Ivanek
303 2016). Lastly, we scanned the literature for genes that are known to affect pelvic and hind
304 limb development, and searched for potential matches in the nine-spined stickleback
305 genome (Varadharajan *et al*. 2019) in regions where significant QTL were found.
306
307 *Simulations*
308 The repeated local adaptation in independently colonised freshwater stickleback
309 populations is widely considered to be due to selection on standing genetic variation
310 available in the sea, which in turn is maintained by recurrent gene flow from previously
311 colonised freshwater populations (the "transporter hypothesis"; Schluter and Conte
312 2009). The effects of standing genetic variation (assuming a panmictic ancestral
313 population) on the probability of parallel local adaptation have already been studied
314 (Ralph and Coop 2015a; MacPhearson and Nuisimer 2016; Lee and Coop 2017;
315 Galloway and Cresko 2019), along with the effects of population structure on adaptation
316 in continuous landscapes (Hermisson and Pennings 2005; Ralph and Coop 2010, 2015b).
317 However, no simulations to date have explicitly considered the effects of heterogeneous
318 patterns in the distribution of standing genetic variation on local adaptation/parallel
319 evolution. Thus, we used forward-in-time simulations to explore the effects of the
320 differences in population demographic parameters between nine- and three-spined
321 sticklebacks on the probability of local adaptation, as described in Supplementary File 4.

322 These simulations were parameterised using empirical estimates of population structure
323 and isolation by distance in the sea for the two species (Supplementary File 5).
324 In short, two freshwater populations were connected to different marine
325 populations at the ends of a stepping-stone chain of marine populations with the
326 carrying capacity (K) and migration (M) both in the sea, and between the founding
327 (marine) and founder (freshwater) population (Supplementary Fig. 2), thus
328 generating comparable patterns of population structuring and IBD in the sea as in
329 the empirical data (see Results). The two freshwater populations were colonised
330 once and allowed to become locally adapted for 10,000 years (20,000 generations).
331 All freshwater adaptation was due to standing genetic variation in the sea that was
332 continually replenished from a "refuge" freshwater population (situated at an equal
333 distance from the two focal freshwater populations), as well as from the two focal
334 freshwater populations once they were colonised. Three different genetic
335 architectures were studied. One (architecture A) corresponds to a single large effect
336 QTL (allowing 100% local adaptation, but with minor effect QTL also affecting the
337 trait), another (architecture B) to a situation where at least two loci are needed to
338 allow local adaptation (allowing 40% or 60% local adaptation), and the third
339 (architecture C) corresponds to the same scenario as A, except that the large effect
340 locus is recessive. Mutation rate was fixed at $1e^{-8}$ and all loci were polymorphic at
341 the beginning of the simulations. More details are given in Supplementary File 4.
342
343

344 **Results**

345 *Heterogeneity of pelvic reduction in the wild*

346 Re-analysis of previously published phenotypic data from the wild confirmed a high
347 degree of variation among different populations with respect to pelvic spine and girdle
348 lengths and their inter-correlations (Fig. 1, Supplementary Table 1 and Supplementary
349 Fig. 3). For instance, while spines were absent and pelvic girdles strongly reduced (but
350 not completely absent) in RYT, spines and girdles were completely lacking in the MAS
351 population (Fig. 1 and Supplementary Table 1). Furthermore, in the BYN population,
352 spines were absent but pelvic girdles were only partially reduced; in BOL (a population
353 adjacent to MAS), large variation in both spine (SD = 0.037) and girdle (SD = 0.047)
354 lengths was observed, although these two traits were strongly correlated (r^2 = 0.61, Fig. 1,
355 Supplementary Table 1 and Supplementary Fig. 3). This suggests that a large effect locus
356 affecting both pelvic spines and girdles segregated in this population at the time of
357 sampling. In six pond populations (*viz*. ABB, KAR, KRK, NAV, HAN, and PYÖ; Fig.
358 1), relative spine (mean = 0.079) and girdle (mean = 0.15) lengths were only slightly
359 smaller relative to the marine populations (0.11 and 0.16 for spine and girdle lengths,
360 respectively; Supplementary Table 1). A lack of pelvic reduction (relative to the marine
361 populations) was observed only in one pond population (KAR; Fig. 1, Supplementary
362 Table 1 and Supplementary Fig. 3).

363

364 *QTL mapping of pelvic reduction in the Helsinki × Rytilampi cross*

365 After LD-network based complexity reduction, all QTL mapping analyses were

366 performed on 241 PCs (six sex-linked PCs were removed). Re-analyses of the 279 F_2
367 progeny from the HEL × RYT cross confirmed a single QTL region on LG7 for both
368 pelvic spine and girdle lengths in the single-locus analyses (Fig. 2a, b and Table 1). In the
369 multi-locus analyses (absolute trait values), LG7 explained 74–86% of the PVE for both
370 spine and girdle lengths, while all other chromosomes individually explained less than
371 3% of the phenotypic variance (Table 2). An approximately equal amount of the
372 phenotypic variance was explained by alleles inherited from the F_0 male (pond
373 individual) and the F_0 female (marine individual; ~30% of the total variance for all traits;
374 Table 2), respectively, with 15–21% of the phenotypic variance also explained by
375 dominance effects. This is the expected outcome for a recessive QTL when the F_0
376 individuals are fixed for different large effect causal alleles, and when no additional
377 smaller effect loci affect the trait (Klug and Cummings 2018). This is also consistent with
378 a strong correlation between relative pelvic spine and girdle lengths ($r^2 = 0.85$, Fig. 1,
379 Supplementary Fig. 3 and Supplementary Table 1).

380

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

382 Among the 308 F_2 progeny of the HEL × BYN cross, single-locus mapping analyses of
383 pelvic spine lengths detected two significant QTL on LG15 (PVE = 8.9%) and LG16
384 (PVE = 13.7%; Fig. 2c, Table 1 and 2) for alleles deriving from the F_0 male (pond
385 individual). Thus, the causal alleles for these QTL segregated in the F_0 pond male (as
386 explained in Supplementary File 3). No dominance effects were detected for these QTL,
387 suggesting that the allelic effects were additive. One QTL on LG6 with an allelic effect

388 deriving only from the F_0 female was also detected (Fig. 2c and Table 1). The QTL
389 significance profiles, in particular for LG15 and LG16, spanned large genomic regions
390 with no obvious peaks (in contrast to LG7 in the HEL × RYT cross; Fig. 2a, b). This
391 remained true when analysing all SNPs individually (fine-mapping; Supplementary Fig.
392 4). In addition, phenotypes of the F_2 individuals from the HEL × RYT cross were not
393 normally distributed and the individual and multi-locus phenotypic effects on spine
394 lengths for the QTL on LGs 6, 15 and 16 indicated epistatic interactions (Supplementary
395 File 6). One significant male QTL on LG4 for girdle length was also detected (Fig. 2d
396 and Table 1). Results for relative trait values were highly similar to the absolute trait
397 values, except for an additional significant male and female QTL on LG1, as well as
398 another significant male QTL on LG14 (Supplementary Fig. 5, Table 1 and
399 Supplementary Table 3).

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

410 2).

411 Compared to the HEL × RYT cross, the correlation between relative pelvic spine
412 and girdle lengths was much weaker in HEL × BYN cross ($r^2 = 0.11$, Fig. 1,
413 Supplementary Fig. 3 and Supplementary Table 1) consistent with pelvic spine and girdle
414 reductions being independently controlled by different QTL. Furthermore, of the 306 F_2
415 individuals, only four displayed complete lack of spines (and the spine was absent in the
416 F_0 male), consistent with spine length being controlled by multiple additive loci.

417

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

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

432 causal mutations (we consider this as a single QTL region).

433

434 *Scanning for Pel deletions in the full-genome sequence data*

435 In the full-genome re-sequencing data of wild-collected individuals, a large deletion
436 upstream of *Pitx1* was fixed for all 21 individuals from Rytilampi, where pelvic reduction
437 maps to this region (Fig. 3). The deletion is around 3.5 kb (LG7:17015109-17018744;
438 Supplementary Fig. 6) in size and fully encompasses the Pel-2.5kb^{SALR} region of the
439 three-spined stickleback (Chan *et al*. 2010). No comparable deletion was found in any
440 other individuals in the dataset (Fig. 3b). This included the two White Sea populations, in
441 which either complete reduction of both the pelvic spines and girdles was observed
442 (MAS), or a putative large effect locus affecting both spine and girdle length was found
443 to be segregating (BOL; Fig. 1; Supplementary Fig. 3 and Supplementary Table 1).

444

445

446 *Candidate genes*

447 Seven candidate genes or regulatory elements for pelvic reduction identified from the
448 literature (Supplementary Table 4) were found in the LGs with significant QTL (Fig. 2).
449 *Hif1a* (Mudie *et al*. 2014), *Pel* (Chan *et al*. 2010) and *Pou1f1* (Kelberman *et al*. 2009) are
450 known to regulate the expression of *Pitx1,* whereas four genes (*Fgf8*, *Wnt8c*, *Wnt8b* and
451 *Hoxd9*) are involved in the pelvic fin/hind limb development downstream of *Pitx1* (Don
452 *et al.* 2012; Tanaka *et al*. 2005). However, aside from *Pel* (LG7), only three – *Wnt8c*

453 (LG6), *Hif1a* (LG15) and *Pouf1f1* (LG16) – were clearly within the significant QTL
454 regions (Fig. 2). One candidate locus, *Hif1a*, is also on LG1 where significant QTL peaks
455 were found when analysing relative (but not absolute) trait values (Table 1 and
456 Supplementary Fig. 5).

457

458 *Simulations*

459 In the empirical data from the marine populations, IBD (the slope of the regression line
460 between linearised F_{ST} and geographic distance point distance in km) was 2.13^{-7} (95% of
461 the $1e^{6}$ bootstrap replicates were between 2.05^{-7} and 2.2^{-7}) for nine-spined sticklebacks,
462 and 1.53^{-8} (95% of the bootstrap replicates were between 1.10^{-8} and 1.96^{-8}) for three-
463 spined sticklebacks. Thus, the slope of the IBD regression line in this dataset was 13.9
464 times higher for nine-spined sticklebacks compared to three-spined sticklebacks (Fig. 4a).
465 The geographic distance (which is on an arbitrary scale) in the simulated data was scaled
466 based on the observed levels of IBD in the empirical data. This resulted in the distance
467 between the two simulated marine populations furthest away from each other (i.e. the
468 marine populations from which the focal freshwater populations were founded)
469 corresponding to 264 km and 352 km for three- and nine-spined sticklebacks,
470 respectively (Fig. 4a). Thus, with comparable levels of IBD as in the empirical data, our
471 simulations mimic the levels of parallel evolution that can be expected in three- and nine-
472 spined sticklebacks at relatively short geographic scales (<400 km). The difference in
473 IBD between the two species is also close to that in the empirical data (264/352 km =
474 0.75). In the empirical data, genetic differentiation between freshwater habitats for

475 populations <400 km from each other (mean F_{ST} = 0.19 and 0.49 for three- and nine-
476 spined sticklebacks, respectively) was also on par with the simulations (mean [across
477 simulation replicates] F_{ST} = 0.21 and 0.58, respectively). Notably, F_{ST} was >0.8
478 (linearised F_{ST}>4) between Pyöreälampi (no pelvic reduction) and Rytilampi (pelvic
479 reduction controlled by *Pitx1*/*Pel*), although these ponds are situated only 15 km from
480 each other (Fig. 4a; max and median F_{ST} = 0.96 and 0.62, respectively, for all pairwise
481 freshwater-freshwater comparisons). This is higher than the F_{ST} between any pair of
482 three-spined stickleback populations in the data (max = 0.78 and median = 0.26 for all
483 pairwise freshwater-freshwater comparisons).
484
485 *Do population demographic parameters influence local adaptation?*
486 In the simulations (Supplementary File 4), the relationship between marine-freshwater
487 F_{ST} and freshwater-freshwater F_{ST} depended on both the species and genetic architecture
488 (Fig. 4b). For instance, when the genetic architecture included one additive large effect
489 locus (architecture A; homozygotes for this locus were 100% locally adapted to
490 freshwater), this locus was often involved in parallel evolution when population structure
491 was parameterised based on three-spined sticklebacks (65% of replicates, n = 100), but
492 not when population structure was parameterised based on nine-spined sticklebacks (3%
493 of the replicates). In 20% of the replicates for both species, the freshwater allele for this
494 locus was fixed in only one of the focal freshwater populations (i.e. local adaptation, but
495 not parallel evolution). When the trait under selection was controlled by several medium
496 effect loci (architecture B), parallel evolution was more common in both stickleback

497 scenarios, particularly for the locus with the largest allelic value (20% and 80% for nine-
498 and three-spined sticklebacks, respectively; Fig. 4b, c; homozygotes for this locus were
499 60% locally adapted to freshwater). Local adaptation also occurred in 57% of the
500 replicates in the nine-spine stickleback-like scenario, and in 18% in the 3-sp (Fig. 4b, c).
501 For the locus with the second highest allelic value in architecture B (homozygotes for this
502 locus were 40% locally adapted to freshwater), the cases of both parallel evolution and
503 local adaptation collectively dropped to 38% and 43% three and nine-spined like
504 stickleback scenarios, respectively. With one non-additive large effect locus with
505 recessive alleles locally adapted to freshwater (architecture C), this locus was less likely
506 to be involved in parallel evolution in the three-spined stickleback-like scenario (48%),
507 compared to architecture A. However, results for the nine-spined stickleback-like
508 scenario was similar to architecture A (6% parallel evolution; Fig. 4b, c). Thus, at
509 relatively short (<400 km) geographical distances, parallel evolution is an expected
510 outcome in three- but not nine-spined stickleback-like scenarios, particularly when a
511 single additive large effect locus is responsible for freshwater adaptation. Results
512 addressing the questions of how local adaptation depend on ancestral allele frequency are
513 presented in Supplementary File 7 and Supplementary Figure 7, and QTL-use in parallel
514 evolution as a function of allelic effect size is shown in Supplementary Figure 8.
515

516 **Discussion**

517 The results demonstrate that pelvic reduction in nine-spined sticklebacks is not nearly as
518 common as in three-spined sticklebacks, and when it does exist, the genetic basis is more

519 variable. This is consistent with the re-analyses of neutral population genomic data
520 showing that both IBD in the sea and the genetic structure among marine populations are
521 >10 times higher in nine-spined sticklebacks, causing heterogeneity in the distribution of
522 standing genetic variation that is not present to the same extent in the marine three-spined
523 stickleback populations. Thus, the level and distribution of ancestral variation available
524 for local adaptation (and also for parallel phenotypic evolution) is likely a function of
525 population demographic parameters, with local adaptation being less likely in poorly
526 connected species, as suggested by Merilä (2013, 2014) and corroborated by our
527 simulations. However, other non-mutually exclusive factors, such as the genetic
528 architecture (i.e. dominance effects, heritability, mutation rates and numbers of causal
529 loci involved), as well as non-parallelism in phenotypic selection optima are also likely to
530 play roles. In the following, we discuss the possible causes of the discrepancy in pelvic
531 structure development between nine- and three-spined sticklebacks, as well as their
532 implications to our understanding of adaptive evolution in the wild.
533
534 *Can the genetic architecture of pelvic reduction be explained by population demographic*
535 *parameters?*
536 Together with earlier QTL studies (Shapiro *et al.* 2006, 2009; Shikano *et al.* 2013), we
537 show that multiple genomic regions (11 QTL, ten of which are novel to this study) are
538 associated with pelvic reduction in nine-spined sticklebacks across their distribution
539 range. Only one small effect QTL region (LG1; Table 1 and Supplementary Fig. 5) was
540 shared between any two crosses (HEL × BYN and HEL × PYÖ) in our study, but even

541 here it is not certain whether the underlying causal mutations are the same. Although the
542 high frequency of *Pel* deletions (disrupting pelvic armour development) means that most
543 of the deletions associated with pelvic reduction in three-spined sticklebacks are
544 independently derived (Xie *et al.* 2019), *Pitx1*/*Pel* is nevertheless predominantly
545 responsible for pelvic reduction in three-spined sticklebacks (Chan *et al.* 2010; Xie *et al.*
546 2019). In contrast, the major effect and recessive *Pitx1*/*Pel* allele in nine-spined
547 sticklebacks is known to be responsible for pelvic reduction in only one Canadian
548 population and Rytilampi. In both Bynästjärnen and Pyöreälampi, pelvic reduction is less
549 heritable, polygenic and additive. In addition, based on morphological data, it seems
550 likely that major effect loci not associated with *Pel*-deletions (Fig. 3) control pelvic
551 reduction in two Russian ponds, with one additional Alaskan population being controlled
552 by a large effect additive locus mapping to LG4 (Shapiro *et al.* 2009). We do not yet
553 know the mutation rate of the *Pel*-deletions in nine-spined sticklebacks so naturally, the
554 lack of *Pel*-deletions in nine-spined sticklebacks could be simply due to lower mutation
555 rates compared to three-spined sticklebacks. Regardless, pelvic reduction in the nine-
556 spined stickleback is much less common than in three-spined sticklebacks in general
557 (Klepaker *et al.* 2013), and when it occurs, the QTL that control pelvic reduction are
558 much more variable with respect to heritability, effect size distribution and dominance
559 relationships. Assuming that selection for pelvic reduction in nine-spined stickleback
560 freshwater populations is universal (see below), such a pattern is consistent with poor
561 connectivity, causing heterogeneity in the patterns of standing genetic variation. This
562 would restrict local adaptation, and hence, also parallel phenotypic evolution. However,

563 the focus of this study is on the role of population demographic parameters on local
564 adaptation from standing genetic variation, regardless of whether the alleles that confer
565 freshwater adaptations are identical by descent (e.g. EDA) or independently derived (e.g.
566 *Pel*-deletions in three-spined sticklebacks), and only secondarily on parallelism (i.e. when
567 the same trait is locally adapted in multiple independently colonised populations,
568 regardless of genetic architecture).

569 It has been shown that multiple independent alleles for the same adaptation can
570 temporarily co-exist in structured populations, causing heterogeneity in standing genetic
571 variation across the distribution range of a species. This can be true when considering the
572 dynamics of novel mutations (Ralph and Coop 2010) or neutral alleles segregating in the
573 population before they become adaptive (Ralph and Coop 2015a). However, these studies
574 assume uniform selection across single continuous populations and thus do not consider
575 the scenario where locally adapted freshwater alleles – which can be different in different
576 parts of the distribution range – are continually introduced to the sea. Furthermore,
577 simulations showing that the probability of local parallel evolution from standing genetic
578 variation depends on the migration rate from the sea (Galloway and Cresko 2019) or from
579 other populations of the same habitat (Ralph and Coop 2015a) implicitly assume a
580 homogenous pool of ancestral standing genetic variation. Our simulations show that local
581 adaptation is also a function of the allele frequency in the founding marine population;
582 with stronger IBD in the sea, standing genetic variation in the ancestral marine population
583 was a stronger bottleneck for parameters that resulted in nine-spined stickleback-like
584 population structure (9-sp) than for parameters that resulted in three-spined stickleback-

585 like population structure (3-sp; Supplementary Figures 7 and 8). Particularly for smaller
586 effect loci, there was a stronger dependence between the allele frequency in the ancestral
587 marine population and local adaptation in 9- compared to 3-sp, indicating that smaller
588 effect QTL have a stronger influence on local adaptation in 9- compared to 3-sp
589 (Supplementary Figures 7 and 8). This resulted in more polygenic and/or less complete
590 local adaptation in 9- compared to 3-sp (Supplementary Figures 7 and 8), indicating that
591 population demography is sufficient to preclude local adaptation.

592 While our simulation results match closely with estimates of population structure
593 and IBD in the empirical data, these simulations explore only a small proportion of the
594 possible parameter space with respect to selection intensity, effective population sizes
595 and migration rates (detailed in Supplementary File 4). We also assume that all parallel
596 evolution is due to standing genetic variation of alleles that are identical by descent (since
597 mutation rate is low), whereas it is known that pelvic reduction in three-spined
598 sticklebacks is, to a large extent, due to independently derived *Pel*-deletions attributable
599 to high mutation rates (Xie *et al.* 2019). Nevertheless, most three-spined stickleback
600 studies indicate that freshwater adaptation is indeed chiefly due to alleles identical by
601 descent that have segregated in the population for millions of years (e.g. Jones *et al.*
602 2012; Nelson and Cresko 2018), making *Pitx1* an exception. In the case of nine-spined
603 sticklebacks, *Pitx1* is with certainty associated with pelvic reduction in only two of all
604 studied populations, suggesting a minor role for recurrent mutation in determining pelvic
605 reduction in this species. While our simulations demonstrate that heterogeneous patterns
606 of standing genetic variation is sufficient for limiting local adaptation in species

607 characterized by low connectivity, further simulations with a wider parameter space
608 could be helpful in advancing our understanding of the limits of adaptation in small and
609 structured populations/species.

610 Interestingly, in the HEL × BYN cross complete spine reduction most likely
611 occurred when the allele responsible for spine reduction for the LG6 QTL was combined
612 with at least one allele causing spine reduction from the LG15 and LG16 QTL, indicating
613 epistatic interactions (Supplementary File 6). If a threshold number of alleles are needed
614 for complete pelvic reduction, this could also explain how standing genetic variation in
615 the sea is maintained, as the necessary multi-locus genotypes that cause sub-optimal
616 phenotypes in the sea are rarely formed, due to overall lower frequencies of spine-
617 reducing alleles in the sea. This is analogous to "epistatic shielding" that can contribute to
618 the persistence of disease alleles in populations (Phillips and Johnsson 1998; Phillips
619 2008). Consistent with this possibility, LG6 of the F_0 female of the HEL × BYN cross
620 (from the sea) was polymorphic for the pelvic spine QTL effect – evidently, a single
621 pelvic-reducing allele alone in this female was not enough to cause any pelvic reduction
622 at all (this female had a complete pelvis).

623

624 *Geographic heterogeneity in selection optima*

625 The high heterogeneity in the genetic architecture of pelvic reduction in the nine-spined
626 stickleback can alternatively be attributed to within habitat environmental variation
627 resulting in different selection optima in different pond populations (cf. Stuart *et al.* 2017;
628 Thompson *et al.* 2019). For example, the small differences between pelvic reduced

629 phenotypes in our study (e.g. in BYN and MAS/BOL spines are completely absent, while
630 in RYT they are only strongly reduced; Fig. 1, Supplementary Table 1 and
631 Supplementary Fig. 3) could in fact indicate different selection optima in the different
632 populations (Stuart *et al*. 2017; Thompson *et al*. 2019). It is possible to use available
633 phenotypic data to estimate the phenotypic optima of pelvic morphology (hypersphere) in
634 each of the populations using Fisher's geometric model (Stuart *et al.* 2017; Thompson *et*
635 *al*. 2019), where a strong overlap would suggest a higher probability of genetic
636 parallelism (Thompson *et al.* 2019). However, this assumes that the populations have
637 access to exactly the same ancestral variation and are free to evolve and reach their
638 optima, which is at odds with the results presented here. Without further detailed
639 environmental data or direct estimates of strength of selection on pelvic phenotypes,
640 disentangling the effects of gene flow and within habitat environmental variation
641 (assuming this leads to non-parallel angles of selection) is not possible (Stuart *et al.*
642 2017).

643 In a recent simulation study, Thompson *et al.* (2019) showed that genetic
644 parallelism from standing genetic variation rapidly declines as selection changes from
645 fully parallel (optima angle of 0°) to divergent (optima angle of 180°), especially when
646 the trait is polygenic. However, although selection was fully parallel in our simulations,
647 we did not observe strong genetic parallelism for smaller effect loci (with allelic effects <
648 6) in both the three- and nine-spined stickleback-like scenarios (Fig. 5b, c). This suggests
649 that the effects of the underlying genetic architecture on parallelism (in conjunction with
650 some IBD and population structuring) can be independent of the angle of optimal

651 phenotypes between two habitats. Thus, overlapping selection optima is necessary but
652 nevertheless not sufficient for parallel evolution to take place and it is important not to
653 disregard the population demographic setting as a factor that could severely restrict
654 heritability for adaptation and/or constrain adaptation to less optimal solutions.
655 Evolutionary studies of species with population demographic parameters comparable to
656 those typical for vulnerable or endangered species/populations, such as the nine-spined
657 stickleback, would be valuable to gain a better understanding of how such species may
658 respond to environmental changes and urbanisation (Thompson *et al.* 2018).
659
660 *Pelvic reduction outside marine-freshwater study systems*
661 While the evidence for genetic parallelism on large geographical scales in the marine-
662 freshwater stickleback model system is extensive (Colosimo *et al.* 2005; Jones *et al.*
663 2012; Terekhanova *et al.* 2014, 2019; Nelson and Cresko 2018; Fang *et al*. 2019), the
664 level of parallelism in lake-stream and pelagic-benthic ecotype pairs of three-spined
665 sticklebacks is much more diverse (Peichel *et al*. 2001, 2017; Conte *et al*. 2012, 2015;
666 Stuart *et al*. 2017; Rennison *et al*. 2019). For instance, Conte *et al*. (2015) found that
667 among benthic-limnetic three-spined stickleback pairs from Paxton and Priest lakes
668 (Vancouver Island, BC, Canada), 76% of 42 phenotypic traits diverged in the same
669 direction, whereas only 49% of the underlying QTL evolved in parallel in both lakes. For
670 highly parallel traits in two other pairs of benthic-limnetic sticklebacks, only 32% of the
671 underlying QTL were shared (Conte *et al*. 2012). Similarly, Stuart *et al*. (2017) found that
672 among 11 evolutionary independent replicate pairs of lake-stream three-spined

673 stickleback populations (Vancouver Island, BC, Canada), both within habitat variation
674 and constraints to gene flow contributed to the observed variation in levels of phenotypic
675 parallelism. Different lakes and streams likely do not have similar access to the same
676 global pool of ancestral variation as pairs of marine-freshwater three-spined sticklebacks,
677 where gene flow in the sea is high. This is consistent with the notion that more
678 heterogeneous access to ancestral variation can indeed limit genetic parallelism. This is
679 also true for one example of marine-freshwater three-spined stickleback divergence
680 among isolated insular freshwater populations in the Haida Gwaii archipelago off the
681 northern Pacific coast of Canada (Deagle *et al.* 2013). Here, similar to the nine-spined
682 sticklebacks in this study, several freshwater populations did not display any reduction in
683 pelvic armour. However, those populations that were fully plated were also genetically
684 more similar to adjacent marine individuals, suggesting that recent marine-freshwater
685 admixture and/or selection favouring plated freshwater individuals could explain this
686 pattern. Thus, with respect to access to ancestral variation available for freshwater
687 adaptation, nine-spined sticklebacks are likely closer to the three-spined stickleback lake-
688 stream and benthic-limnetic study systems than to the three-spined stickleback marine-
689 freshwater study system. The only notable exception is the lake-stream three-spined
690 sticklebacks mentioned above, where genetic structuring also is high.
691 In a small geographic region in Belgium, Raeymaekers et al. (2017) compared the
692 strength and nature of neutral and adaptive divergence along a salinity gradient in nine-
693 and three-spined sticklebacks. Since the lake and pond populations in this study were
694 connected by creaks (allowing contemporary gene flow), in contrast to our study – where

695 a much larger geographic area is considered – population structuring among the nine-
696 spined stickleback population was not stronger than in the three-spined sticklebacks.
697 Despite gene flow, both phenotypic and genetic parallelism was still lower in nine-
698 compared to three-spined sticklebacks. We suspect that lack of parallelism in nine-spined
699 sticklebacks in Raeymaekers et al. (2017) simply lies in lack of ancestral variation – it
700 does not matter how much gene flow there is if the necessary ancestral variation for local
701 adaptation is lacking. Based on our findings here it seems thus likely that ancestral
702 variation for local adaptation in nine- but not three-spined sticklebacks was already
703 lacking in the marine ancestral populations from which the Raeymaekers et al. (2017)
704 study area was colonised from.

705

706 *Candidate genes*

707 While the QTL peak for the *Pitx1*/*Pel* region in the HEL × RYT cross was narrow, this
708 was not the case for the other QTL we detected. Hence, due to the large QTL regions
709 detected by four-way single-mapping analyses, it was not meaningful to perform gene
710 ontology enrichment (GO) analyses – the QTL regions would have contained possibly
711 thousands of genes. Instead, we searched the literature for known candidate genes related
712 to pelvic reduction and found three (excluding *Pitx1*/*Pel*) that were clearly contained
713 within the identified QTL regions (Fig. 2 and Supplementary Table 4). Due to the
714 aforementioned large QTL regions, these can only be considered as highly putative
715 candidate genes for pelvic reduction and will not be discussed further. However, further
716 studies of pelvic reduction might find these candidates worthy of attention.

717

718 *Conclusions*

719 Our results show that the repeated parallel reduction in pelvic structures in freshwater
720 populations of nine-spined sticklebacks is due to a diverse set of genetic changes: only
721 one small effect QTL for pelvic reduction was shared between the three experimental
722 crosses in this study. In one cross, pelvic reduction was mapped to the previously
723 identified *Pitx1*/*Pel* regions, but in the other two crosses, the genetic basis of pelvic
724 reduction was polygenic, and mapped to many different chromosomes. In addition to
725 these, yet another large effect QTL different from the *Pitx1*/*Pel* locus likely segregates in
726 one nine-spined stickleback population, which is yet to be identified. The results also
727 shed light on the possible drivers of the observed genetic heterogeneity underlying pelvic
728 reduction; as shown by simulations, heterogeneous genetic architectures are more likely
729 to emerge when access to ancestral variation is limited by strong isolation by distance and
730 population structuring. This reinforces the role of the nine-spined sticklebacks as a useful
731 model system, alongside the three-spined stickleback, to study adaptive evolution in the
732 wild. Furthermore, since the population demographic characteristics of nine-spined
733 sticklebacks are similar to small and endangered species/populations, it is also likely to
734 be a well-suited model to study the genetics of adaptation in populations of conservation
735 concern.

736

737 **Data availably**

738 The data and code for analyses that support the findings of this study are openly

available in Dryad at DOI https://doi.org/10.5061/dryad.76hdr7str. Raw sequence reads have been submitted to NCBIs short read archives with accession numbers: PRJNA673430 and PRJNA672863.

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

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

799 Cresko, W. A., Amores, A., Wilson, C., Murphy, J., Currey, M., Phillips, P., … J.H.
800 Postlethwait. 2004. Parallel genetic basis for repeated evolution of armor loss in
801 Alaskan threespine stickleback populations. Proc. Natl. Acad. Sci. USA 101: 6050–
802 6055.

803 Deagle, B. E., Jones, F., Chan, Y. F., Absher, D. M., Kingsley, D. M., and T. E.
804 Reimchen. 2012. Population genomics of parallel phenotypic evolution in
805 stickleback across stream–lake ecological transitions. Proc. R. Soc. B 279: 1277–
806 1286 doi:10.1098/rspb.2011.1552

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

811 DeFaveri, J., Shikano, T., Ghani, N. I. A., and J. Merilä. 2012. Contrasting population
812 structures in two sympatric fishes in the Baltic Sea basin. Mar. Biol. 159: 1659–
813 1672. https://doi.org/10.1007/s00227-012-1951-4

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

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

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

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

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

827 Giles, N. 1983. The possible role of environmental calcium levels during the evolution of
828 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

861 Klepaker, T., Østbye, K., and M. A. Bell. 2013. Regressive evolution of the pelvic
862 complex in stickleback fishes: a study of convergent evolution. Evol. Ecol. Res. 15:
863 413–435.

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

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

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

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

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

876 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … T. Sam. 2009.
877 The Sequence Alignment / Map format and SAMtools. Bioinformatics 25: 2078–
878 2079. https://doi.org/10.1093/bioinformatics/btp352

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

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

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

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

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

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

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

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

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

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

Orr, H. A. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52:935-949. DOI: 10.2307/2411226

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

Peichel, C.L, Marques, D.A. 2017. The genetic and molecular architecture ofphenotypic diversity in sticklebacks. Phil.Trans. R. Soc. B372: 20150486. http://dx.doi.org/10.1098/rstb.2015.0486

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

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

921 Pritchard, J. R., and D. Schluter. 2001. Declining interspecific competition during
922 character displacement: Summoning the ghost of competition past. Evol. Ecol. Res.
923 3: 209–220.

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

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

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

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

935 Raeymaekers, J. A. M., Chaturvedi, A., Hablützel, P. I., Verdonck, I., Hellemans, B.,
936 Maes, G. E., … Volckaert, F. A. M. (2017). Adaptive and non-adaptive divergence
937 in a common landscape. Nature Communications, 8(1), 267. doi: 10.1038/s41467-
938 017-00256-6

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

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

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

946 Rennison, D. J., Stuart, Y. E., Bolnick, D. I., and Peichel, C. L. 2019. Ecological factors
947 and morphological traits are associated with repeated genomic differentiation
948 between lake and stream stickleback. Phil. Trans. R. Soc. B374:20180241.
949 https://doi.org/10.1098/rstb.2018.0241

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

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

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

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

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

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

967 Shapiro, Michael D, Summers, B. R., Balabhadra, S., Aldenhoven, J. T., Miller, A. L.,
968 Cunningham, C. B., … and D.M. Kingsley. 2009. The genetic architecture of
969 skeletal convergence and sex determination in ninespine sticklebacks. Curr. Biol. 19:
970 1140–1145. https://doi.org/10.1016/j.cub.2009.05.029.The

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

974 Shikano, T., Shimada, Y., Herczeg, G., and J. Merilä. 2010a. History vs. habitat type:
975 Explaining the genetic structure of European nine-spined stickleback (*Pungitius*
976 *pungitius*) populations. Mol. Ecol. 19: 1147–1161. https://doi.org/10.1111/j.1365-
977 294X.2010.04553.x

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

981 Sillanpää, M. J. 2011. On statistical methods for estimating heritability in wild
982 populations. Mol. Ecol. 20: 1324–1332. https://doi.org/10.1111/j.1365-
983 294X.2011.05021.x

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

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

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

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

996 Terekhanova, N. V, Barmintseva, A. E., Kondrashov, A. S., Bazykin, G. A., and N. S.
997 Mugue. 2019. Architecture of parallel adaptation in ten lacustrine threespine
998 stickleback populations from the White Sea area. Genome Biol. Evol. 11: 2605–
999 2618. https://doi.org/10.1093/gbe/evz175

1000 Terekhanova, N. V, Logacheva, M. D., Penin, A. A., Neretina, T. V, Barmintseva, A. E.,
1001 Bazykin, G. A., … N.S. Mugue. 2014. Fast evolution from precast bricks :
1002 Genomics of young freshwater populations of threespine stickleback *Gasterosteus*
1003 *aculeatus*. PLoS Genet. 10: e1004696. https://doi.org/10.1371/journal.pgen.1004696

1004 Thompson, K. A., Osmond, M. M., and D. Schluter. 2019. Parallel genetic evolution and
1005 speciation from standing variation. Evol. Lett. 3:129–141.
1006 https://doi.org/10.1002/evl3.106

1007 Thompson, K. A., Rieseberg, L. H., and D. Schluter. 2018. Speciation and the city.
1008 Trends Ecol. Evol 33: 815–826.

1009 Tibshirani, R. 1996. Regression shrinkage and selection via the Lasso. J. R. Stat. Soc.
1010 Series. B 58: 267–288.

1011 Varadharajan, S., Rastas, P., Löytynoja, A., Matschiner, M., Calboli, F. C. F., Guo, B., …
1012 and J. Merilä. 2019. A high-quality assembly of the nine-spined stickleback
1013 (Pungitius pungitius) genome. Genome Biol. Evol. 11: 3291–3308. evz240,
1014 https://doi.org/10.1093/gbe/evz240

1015 Van Ooijen, J. W. 2009. MapQTL v. 6.0: Software for the mapping of quantitative trait
1016 loci in experimental populations of diploid species. Kyazma BV, Wageningen, The
1017 Netherlands.

1018 Wang, Y., Zhao, Y., Wang, Y., Li, Z., Guo, B., and J. Merilä. 2020. Population
1019 transcriptomics reveals weak parallel genetic basis in repeated marine and
1020 freshwater divergence in nine-spined sticklebacks. Mol. Ecol.
1021 https://doi.org/10.1111/mec.15435

1022 Wolf, J. B., Brodie, E. D., and M. J. Wade. 2000. Epistasis and the Evolutionary Process.
1023 Oxford University Press, Oxford.

1024 Xie, K. T., Wang, G., Thompson, A. C., Wucherpfennig, J. I., Reimchen, T. E., Maccoll,
1025 A. D. C., … and P. Lake. 2019. DNA fragility in the parallel evolution of pelvic
1026 reduction in stickleback fish. Science 84: 81–84.

1027 Xu, S. 1996. Mapping quantitative trait loci using four-way crosses. Genet. Res. 68: 175–
1028 181.

1029 Xu, S. 2013a. Genetic mapping and genomic selection using recombination breakpoint
1030 data. Genetics 195: 1103–1115. https://doi.org/10.1534/genetics.113.155309

1031 Xu, S. 2013b. Principles of Statistical Genomics. Springer, New York.
1032 https://doi.org/DOI 10.1007/978-0-387-70807-2

1033 Yang, J., Guo, B., Shikano, T., Liu, X., and J. Merilä. 2016. Quantitative trait locus
1034 analysis of body shape divergence in nine-spined sticklebacks based on high-density
1035 SNP-panel. Sci. Rep. 6: 26632. https://doi.org/10.1038/srep26632

1036
1037

1038 **Table 1 | Summary of QTL-mapping analyses.** Each row corresponds to a significant
1039 (arbitrarily numbered) QTL region, with the proportion variance explained (PVE), jointly
1040 estimated for all PCs (one for each significant LD-cluster) extracted from such regions.
1041 Coding indicates whether the QTL was significant for the alleles inherited from the F_1
1042 female (♀), the F_1 male (♂) or the dominance effect (dom). Effect size (β) is based on the
1043 first PC extracted from all SNPs from each significant QTL region. "Std." indicates
1044 whether the trait was standardised (Yes) or not (No). P and P_{COR} represent nominal and
1045 corrected P-values from single-mapping four-way analyses, respectively. Results for
1046 standardised trait values are only shown for QTL that were not also significant for
1047 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

1048

1049 **Table 2 | Proportion phenotypic variance explained (PVE) in pelvic traits.**
1050 Percentages of total phenotypic variance explained by different linkage groups (LG), by
1051 all SNPs (Tot), by loci inherited from females (♀) and males (♂), as well as the
1052 dominance effect (Dom). Results are shown for each cross and trait separately, and for
1053 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	-

1054

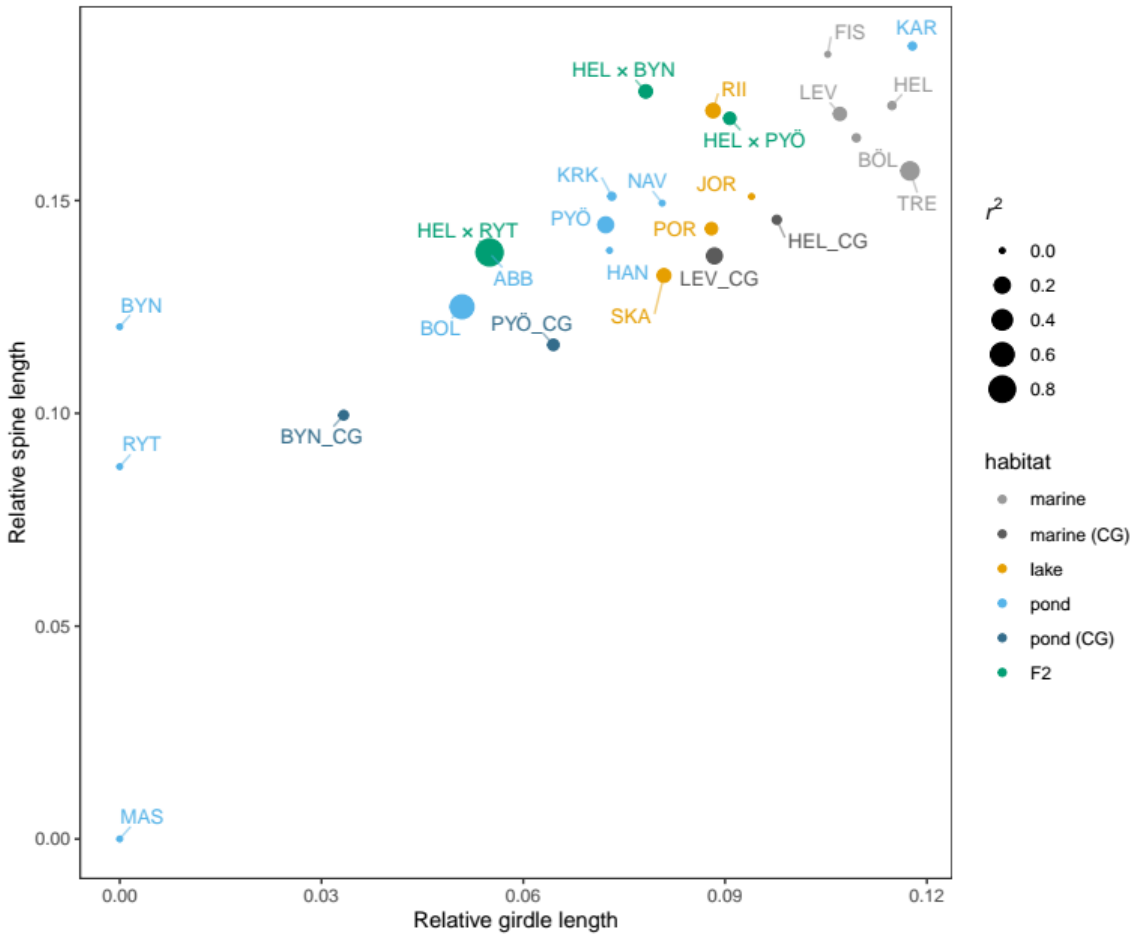

Figure 1 | Summary of previously published nine-spined stickleback pelvic phenotypes from the wild and common garden experiments. Scatterplot depicts relative (to standard body length) pelvic spine and girdle lengths colored according to population and habitat type and with size indicating Pearson's product moment correlation (r^2) between spine and girdle lengths for a given population. Populations with complete spine and/or girdle reductions (BYN, RYT and MAS) are here given r^2=0. Further details can be found in Supplementary Fig. 3 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.

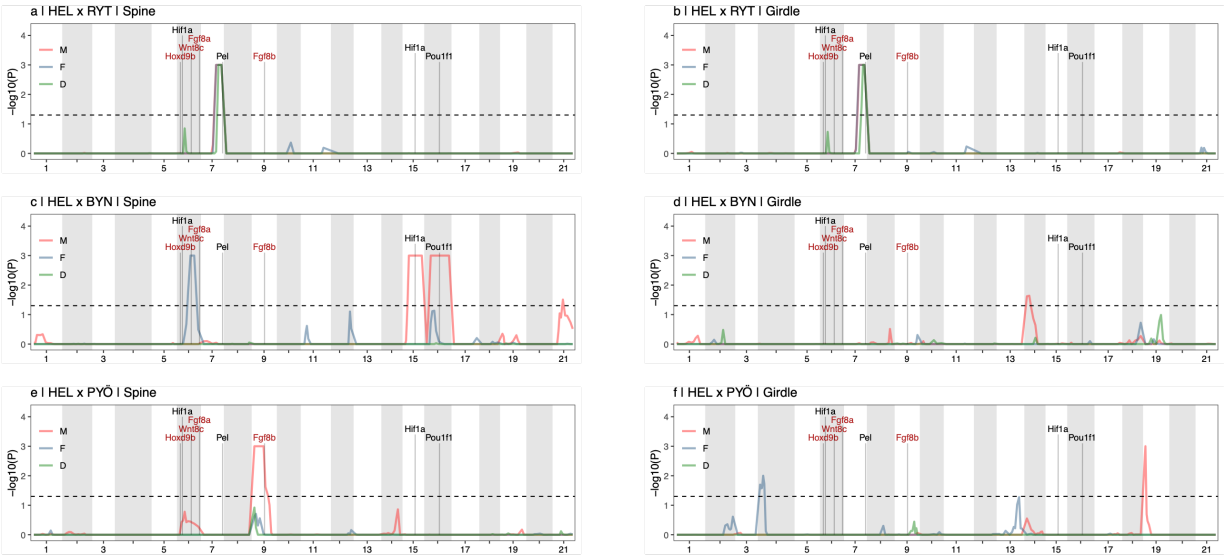

Figure 2 | Quantitative trait locus mapping of pelvic reduction in three independent stickleback crosses. Single-mapping four-way analyses of four morphological traits associated with pelvic reduction in (a-b) HEL × RYT cross, (c-d) HEL × BYN cross, and (e-f) HEL × PYÖ cross. QTL for pelvic spine length 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 3 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.

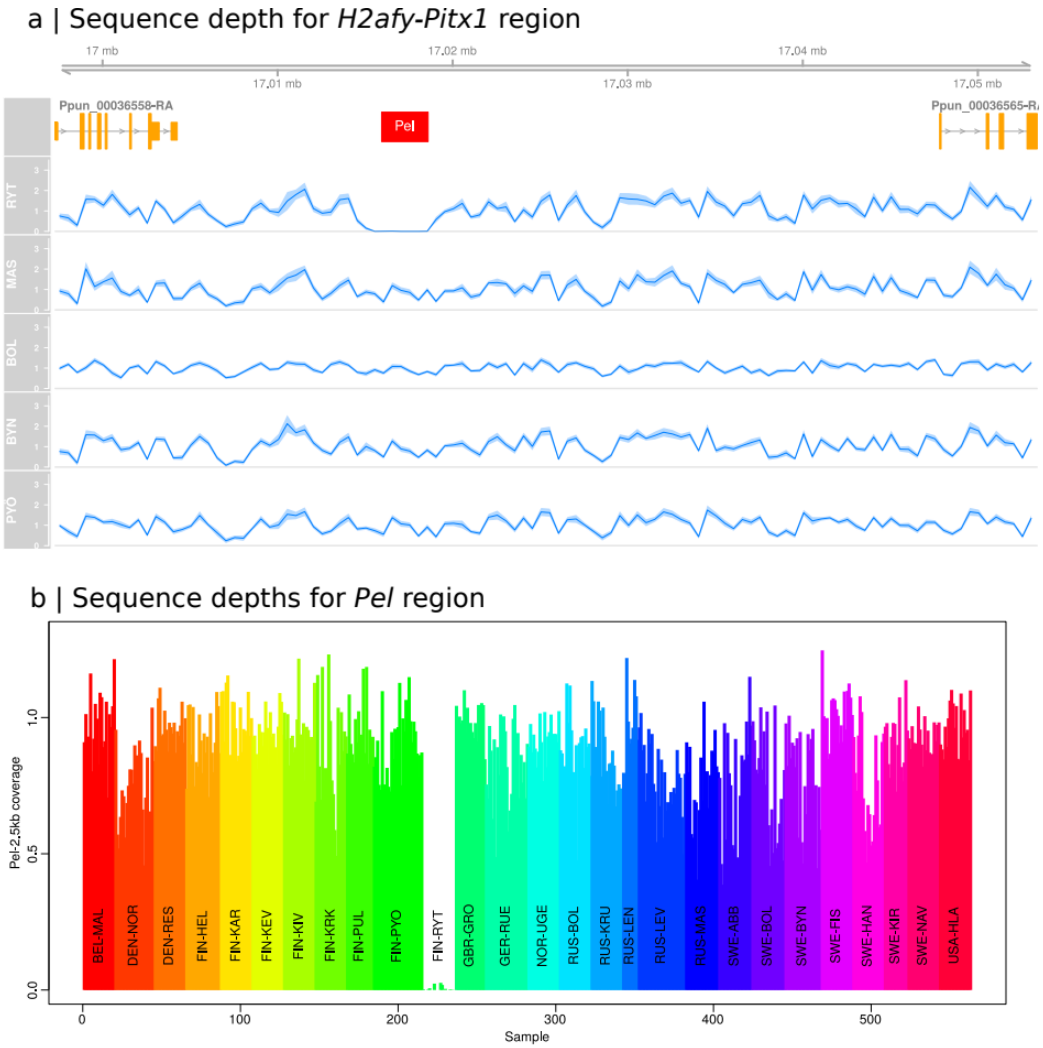

Figure 3 | Sequence depths for *H2afy-Pitx1* region. (a) Relative sequencing depth across the *H2afy-Pitx1* intergenic region for 10 individuals from five different populations. Blue line and shading indicate population average and 95% confidence intervals, respectively. The 3.5kb deletion in RYT, seen 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.

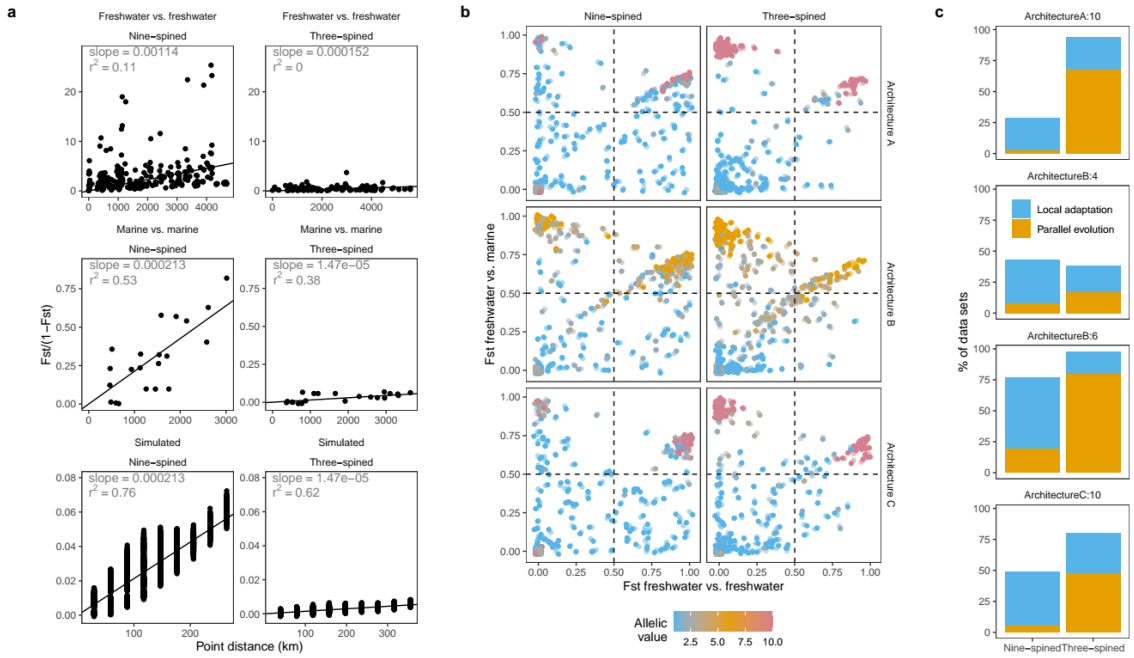

Figure 4 | Simulation results. (a) depicts linearised F_{ST} against geographic distance (IBD) for empirical and simulated data, with slope and squared Pearson's product moment correlation coefficient indicated. Geographic distance for simulated data (IBD in the sea) is scaled to match the slope for the IBD-plot in the sea in the empirical data. (b) depicts freshwater-freshwater F_{ST} against marine-freshwater F_{ST} from all QTL from all simulated data (n = 100), with allelic effect sizes indicated as shown in legend (phenotypic optima in marine and freshwater habitats are 0 and 20, respectively). 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.