Quick divergence but slow convergence during ecotype formation in lake and stream stickleback pairs of variable age

K. LUCEK*†, A. SIVASUNDAR*†, B. K. KRISTJÁNSSON‡, S. SKÚLASON‡ & O. SEEHAUSEN*†

*Aquatic Ecology and Evolution, Institute of Ecology & Evolution, University of Bern, Bern, Switzerland
†Department of Fish Ecology and Evolution, EAWAG Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and Biogeochemistry, Kastanienbaum, Switzerland
‡Department of Aquaculture and Fish Biology, Hólar University College, Sauðárkrókur, Iceland

Abstract

When genetic constraints restrict phenotypic evolution, diversification can be predicted to evolve along so-called lines of least resistance. To address the importance of such constraints and their resolution, studies of parallel phenotypic divergence that differ in their age are valuable. Here, we investigate the parapatric evolution of six lake and stream threespine stickleback systems from Iceland and Switzerland, ranging in age from a few decades to several millennia. Using phenotypic data, we test for parallelism in ecotypic divergence between parapatric lake and stream populations and compare the observed patterns to an ancestral-like marine population. We find strong and consistent phenotypic divergence, both among lake and stream populations and between our freshwater populations and the marine population. Interestingly, ecotypic divergence in low-dimensional phenotype space (i.e. single traits) is rapid and seems to be often completed within 100 years. Yet, the dimensionality of ecotypic divergence was highest in our oldest systems and only there parallel evolution of unrelated ecotypes was strong enough to overwrite phylogenetic contingency. Moreover, the dimensionality of divergence in different systems varies between trait complexes, suggesting different constraints and evolutionary pathways to their resolution among freshwater systems.

Introduction

If natural selection is the principal force governing evolutionary change, divergence among populations can be considered as the tracking of alternative adaptive peaks on the underlying fitness landscape (Wright, 1932; Lande & Arnold, 1983; Steppan *et al.*, 2002; Arnold *et al.*, 2008). The degree of divergence is then expected to depend on the time that has been available for selection to act, the strength of selection, the topology of the fitness landscape and the amount of adaptive standing genetic variation. All of these factors may affect both the rate and the direction of evolution. Addition-

ally, the strength of selection and/or the fitness landscape itself may fluctuate through time due to environmental variation (Jones *et al.*, 2004; Arnold *et al.*, 2008). Genetic drift and selection can reduce standing genetic variation, which may lead to different evolutionary outcomes across replicated cases of population divergence, even when selection is acting in a parallel manner (Barrett & Schluter, 2008). Consequently, strong parallel evolution is only expected if the selective regime, the relative level of standing genetic variation and the segregating alleles themselves are similar (Langerhans & DeWitt, 2004; Kaeuffer *et al.*, 2012) and if selection has enough time to overcome potential historical contingencies (Young *et al.*, 2009).

Evolution towards adaptive peaks can be influenced by genetic 'lines of least resistance' or g_{max}, which can be quantified as the leading eigenvector of the genetic variance–covariance matrix $\mathbf{G}$ (Lande & Arnold, 1983;

Correspondence: Kay Lucek, Aquatic Ecology and Evolution, Institute of Ecology & Evolution, University of Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland. Tel.: +41 316313131; fax: +41 587652168; e-mail: Kay.Lucek@eawag.ch

Schluter, 1996; Jones *et al.*, 2004; Arnold *et al.*, 2008). Biologically, this axis accounts for the largest proportion of genetic variance and is shaped by selection and drift, which then influence genetic constraints within a population (Lande & Arnold, 1983; Steppan *et al.*, 2002; Marroig & Cheverud, 2005). Different **G** matrices can be compared by calculating the angle θ between different g_{max} (Lande & Arnold, 1983; Schluter, 1996; Steppan *et al.*, 2002). Whereas genetic constraints may initially bias evolution towards g_{max} (Schluter, 1996), selection may alter the direction of g_{max} towards an existing or a new optimum on the adaptive landscape (Lande & Arnold, 1983; Jones *et al.*, 2004; Arnold *et al.*, 2008), for example during the colonization of new environments (Bacigalupe, 2009; Eroukhmanoff & Svensson, 2011). Similarly, genetic drift, bottlenecks or mutations may alter the **G** matrix and hence g_{max} (Chapuis *et al.*, 2008).

In the absence of quantitative genetic data, the **G** matrix might be approximated by the **P** matrix, which is based on phenotypic data from wild populations (Cheverud, 1988), especially when phenotypic traits are highly heritable (Lande, 1979), as has been suggested for several taxa (Cheverud, 1988; Kolbe *et al.*, 2011; Leinonen *et al.*, 2011). **P** is defined as the combination of the genetic and environmental covariance matrices, that is **G** + **E** (Lande, 1979; Arnold & Phillips, 1999), where both effects could also interact (**G** × **E**; Falconer, 1989). Consequently, **P** matrices include phenotypically plastic effects, which are differentially expressed in distinct environments (Pigliucci *et al.*, 1999). The leading eigenvector of a **P** matrix (p_{max}) therefore serves as an overall measure of phenotypic variation observed in the wild, combining both genetic and environmental effects. The changes of **P** and p_{max} towards novel adaptive peaks thus might occur rapidly through phenotypic plasticity (Lande, 2009; Draghi & Whitlock, 2012) or adaptation from standing genetic variation (Lande & Shannon, 1996; Barrett & Schluter, 2008). In either case, p_{max} of different replicated systems that vary in age should align, i.e. show a small or zero angle θ between them (Fig. S1). Alternatively, θ between mainly genetically determined p_{max} may evolve over time through selection and drift (Lande & Arnold, 1983; Jones *et al.*, 2004; Arnold *et al.*, 2008). θ is thus expected to subsequently increase over time between an ancestral p_{max} and the p_{max} of a derived population that evolves towards a new adaptive peak (Lande & Arnold, 1983; Jones *et al.*, 2004; Arnold *et al.*, 2008).

In threespine stickleback (*Gasterosteus aculeatus* species complex), the ancestral marine population repeatedly colonized freshwater throughout its distribution mainly after the last glacial maximum and subsequently adapted to different habitats such as streams and lakes. The result was a complex of phenotypically and ecologically divergent populations and, in some cases, even

sympatric or parapatric species (Bell & Foster, 1994; McKinnon & Rundle, 2002). Alongside the marine–freshwater transition, adaptive changes in both the **G** and the **P** matrix have been recorded (Berner *et al.*, 2010b; Leinonen *et al.*, 2011), where g_{max} and p_{max} are correlated (Leinonen *et al.*, 2011). The parallel evolution of distinct parapatric lake–stream pairs within freshwater has made this species complex an excellent system to investigate the process of ecological speciation. However, on an evolutionary timescale, most studies use relatively old systems and are often limited to one restricted geographical area (e.g. Reimchen *et al.*, 1985; Hendry & Taylor, 2004; Berner *et al.*, 2008; Kaeuffer *et al.*, 2012; Ravinet *et al.*, 2013; but see Berner *et al.*, 2010a; Ravinet *et al.*, 2013; Lucek *et al.*, 2013; Hendry *et al.*, 2013). In contrast, some lake–stream systems became only recently available to stickleback, for example due to contemporary translocations (Berner *et al.*, 2010a; Lucek *et al.*, 2010, 2012a; Moser *et al.*, 2012) or the creation of artificial lakes (Kristjánsson *et al.*, 2002a; Hendry *et al.*, 2013). Hence, stickleback provide a rare opportunity to study the evolution of parapatric divergence along the lake–stream habitat axis and the corresponding changes in the **P** matrix and p_{max} over a wide timescale, ranging from decades to millennia.

Here, we study replicated parapatric lake–stream stickleback from Switzerland and Iceland that are between 50 and 10 000 years old in relation to their putative ancestral marine population. Using this temporal gradient, we test whether phenotypic divergence emerges rapidly after the colonization of novel environments and whether lines of least resistance (p_{max}) diverge over time as suggested by several authors (Lande & Arnold, 1983; Jones *et al.*, 2004; Arnold *et al.*, 2008). Additionally, the large geographic scale coupled with the very different colonization histories of Iceland and Switzerland (Ólafsdóttir *et al.*, 2007a; Lucek *et al.*, 2010) allows us to test for parallel evolution. Specifically, we can test whether parapatric phenotypic divergence resulted in similar p_{max} and whether the degree and the direction of habitat-dependent phenotypic divergence are similar among our studied systems. We predict that habitat-dependent phenotypic changes in similar environments should result in similar p_{max} but that the degree of phenotypic divergence may differ due to different historical contingencies, the time for selection to act or differences in the selective regimes among our studied systems.

Materials and methods

Sample collection

We studied three Swiss lake–stream systems in the invasive range of stickleback that differ in their ages of stickleback colonization [Bern (Wohlen): ~50 years,

Constance: 140 years, Geneva: 140 years] and represent either independent introductions from different freshwater lineages (Constance, Geneva) or a case of recent admixture of these lineages (Bern; see Lucek *et al.*, 2010 for details). In addition, we studied three Icelandic lake–stream systems that differ in their geological age (Mývatn: 2500 years, Thingvallavatn: 8000–10 000 years; Saemundsson, 1992; Einarsson *et al.*, 2004) or are man-made (Hraunsfjördur: 50 years; Kristjánsson *et al.*, 2002b) and have been separately colonized by stickleback from ancestral marine populations. We also sampled two Icelandic marine populations (Table S1), presumably resembling the phenotypic marine ancestral state to most European freshwater stickleback (Mäkinen *et al.*, 2006; Jones *et al.*, 2012a). In Icelandic lakes, stickleback have been described to diverge in relation to benthic substrate (Kristjánsson *et al.*, 2002b). We thus sampled the largest potential habitats in each lake (see Fig. 1 and Table S1 for sampling locations).

Icelandic samples were obtained between August and September 2010 using minnow traps and by hand netting. Samples from Switzerland were similarly collected in 2007 and 2008 (Lucek *et al.*, 2010). In all cases, stream stickleback were obtained from inflowing streams (Table S1). All fish were killed with an overdose of clove oil and stored in 70% ethanol. A fin clip was additionally taken for genetic analysis and preserved in absolute ethanol. Sample size per site ranged from 17 to 62 (mean: 35 ± 10 SD) with a total of 918 individuals from 26 sites (Fig. 1, Table S1). Altitudinal difference and pairwise waterway distance between each stream site and the inflow of the stream into the lake were measured using GOOGLEEARTH (Google, USA).

Genetic analysis

We extracted DNA for individuals from the Marine 1 site and all freshwater sites, except for Mývatn, where only one of the three stream sites (MS1) was available for genetic analysis ($N_{Total} = 727$, Table S1). DNA was extracted using a 10% Chelex solution, following the manufacturer's protocol (Biorad, CA, USA). In some cases, additional individuals were included for which no phenotypic data were collected (Table S1). We amplified ten microsatellite markers in one multiplex set following the protocols of Raeymaekers *et al.* (2007). Three of these markers (Stn26, Stn96 and Stn130) have been shown to be associated with known QTLs for spine lengths (Peichel *et al.*, 2001). Detailed information on marker identity, the multiplexing set-up and the PCR protocol is provided as supplementary methods. We visualized alleles on an ABI 3130XL and scored them with GENEMAPPER 4.0 (Applied Biosystems, Zug, Switzerland). We generated a genetic tree-like relationship among sampling sites based on their Cavalli-Sforza distances of allelic frequencies using a

neighbour-joining algorithm implemented in PHYLIP 3.69 (Felsenstein 2012). Significance was estimated using 1000 bootstrapped resampling replicates. To test whether our markers conformed to neutral expectations, we conducted an F_{ST}-based outlier test using the software LOSITAN 1.0 (Antao *et al.*, 2008) separately for each lake–stream system. Using GENODIVE 2.0 (Meirmans & Van Tienderen, 2004), we calculated pairwise F_{ST} between parapatric lake and stream populations for all systems, pooling all sampling sites within a lake. We estimated significances using 1000 bootstrapped replicates as implemented in GENODIVE. Finally, we tested for a correlation of the obtained pairwise parapatric F_{ST} values with either the altitudinal difference or the geographical distance between a parapatric stream site and the lake using linear models. Models were compared using the Akaike information criterion corrected for small sample sizes (AICc).

Morphological data collection and analysis

We measured sixteen linear morphological traits (see Fig. S2 for details), many of which are known to be associated with ecological diversification in stickleback (see Kristjánsson *et al.*, 2002a; Mori & Takamura, 2004; Berner *et al.*, 2008; Leinonen *et al.*, 2011 and references therein), on the left side of each fish to the nearest 0.01 mm using a digital calliper. These traits were related to either antipredator defence (FSL, length of the first dorsal spine; DSL, length of the second dorsal spine; PSL, length of the pelvic spine; and PGL, length of the pelvic girdle), feeding (HL, head length; UJL, upper jaw length; SnL, snout length; SnW, snout width; and ED, eye diameter), or body shape and swimming performance (SL, standard length; PGW, width of the pelvic girdle; BD1, body depth measured after the first dorsal spine; BD2, body depth measured after the second dorsal spine; CPL, caudal peduncle length; BLA, basal length of the anal fin; BLD, basal length of the dorsal fin; and TLP, total length of the pelvic fin). We measured two additional feeding-related traits: the length of the lower gill arch (AL) and the length of the second gill raker (GRL2), as counted from the joint of the dorsal arch bone on the first lower gill arch (Berner *et al.*, 2008). Both measurements on the gill arch were taken using a micrometre mounted on a dissection microscope. Because all traits were significantly correlated with SL (results not shown), we size corrected the data using the residuals from a regression of each trait against SL. This regression was either performed pooling all individuals for the overall comparison of populations or separately (i.e. for each lake–stream system and for the marine population) for pairwise comparisons. By pooling all systems and populations, allometric information in some populations may be retained if the allometric trajectories differ among them. This allows, however, to estimate system- and
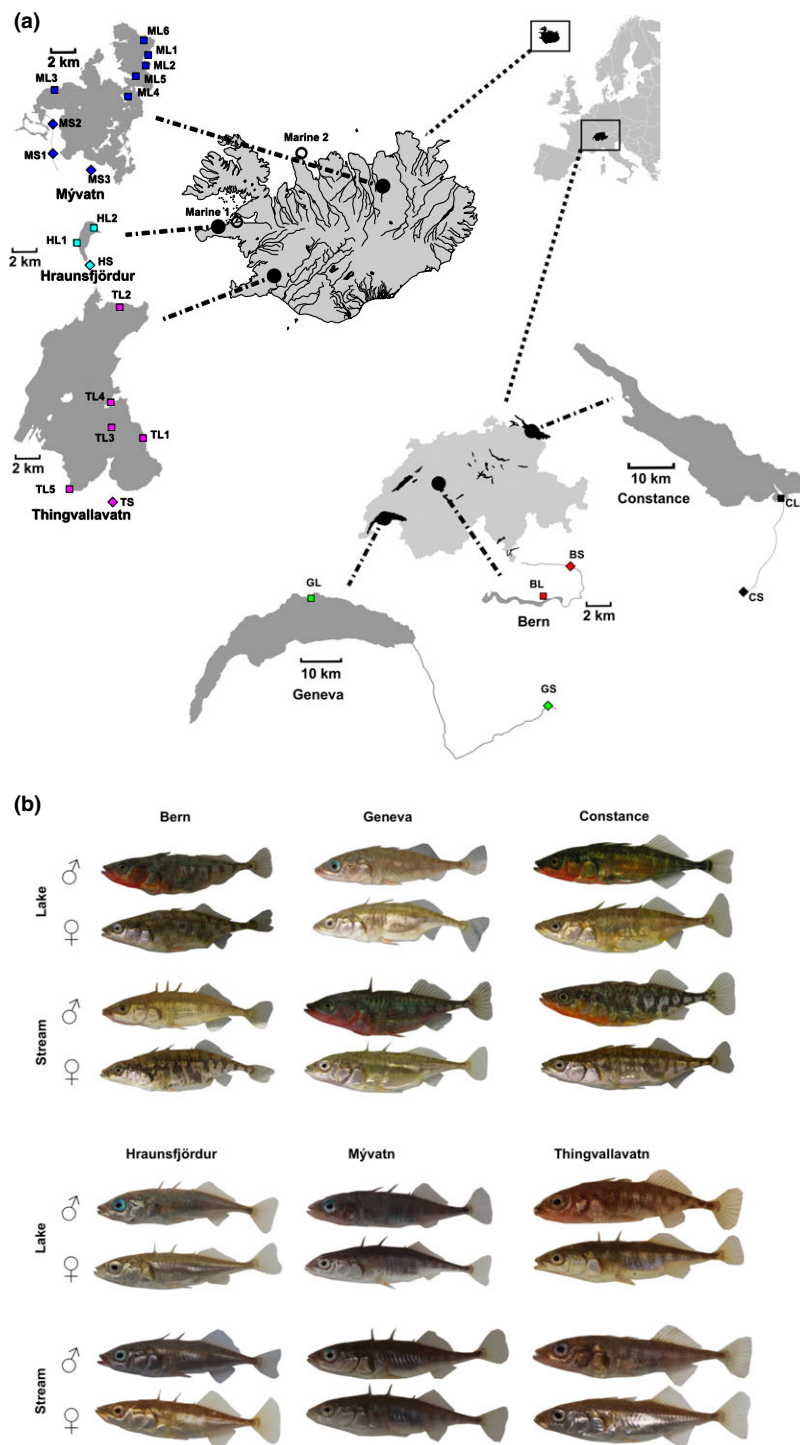
Fig. 1 Overview of the studied systems: (a) Sampled lakes and corresponding sampling sites (squares: lake populations; diamonds: stream populations; and circles: marine populations) for both Iceland (top) and Switzerland (bottom). (b) Representative examples of each sex for the different stickleback ecotypes of each lake–stream system.

population-specific components of phenotypic variation, which can be explained by different historical contingencies or differences in the selective regimes.

To estimate the relative contributions of *country* (Iceland or Switzerland), lake–stream *system* (Bern, Constance, Geneva, Hraunsfjördur, Mývatn, Thingvalla-

vatn), *habitat* (lake or stream) and the interaction of *system* x *habitat* on diversification within freshwater, we calculated the percentage of nonerror variance based on the respective sums of squares using a sequential ANOVA model (Langerhans & DeWitt, 2004; Eroukhmanoff *et al.*, 2009; Lucek *et al.*, 2013). Here, *country* should reflect variation due to different historical contingencies, which include differences in past selection regimes, or differences between the current selective regimes between Switzerland and Iceland. Similarly, *system* accounts for the variation among isolated lake–stream systems. The *habitat* term reflects the component of parapatric phenotypic divergence that is replicated among systems, that is parallel. Finally, the *system* x *habitat* interaction accounts for interactions between system differences (colonization history and environmental differences) and habitat-dependent phenotypic divergence within systems (Langerhans & DeWitt, 2004; Eroukhmanoff *et al.*, 2009). To further compare the overall phenotypic divergence among all sampled sites, we constructed a tree-like relationship using pairwise Mahalanobis distances based on the overall size-corrected phenotypic measurements.

Because local adaptation can lead to phenotypic differentiation between populations of the same ecotype (i.e. within one habitat type; Hendry & Taylor, 2004; Kaeuffer *et al.*, 2012; Ravinet *et al.*, 2013), all individuals from the same habitat were pooled if more than one site was sampled in a given lake or in a given stream system to estimate the overall degree of habitat-dependent phenotypic divergence. We estimated the parapatric phenotypic divergence using P_{ST}, an analog to Q_{ST} (Spitze, 1993), based on phenotypic measurements from wild individuals, which serves as a unit-less proportional measure of pairwise phenotypic divergence and is analogous to our measure of pairwise genetic divergence (F_{ST}). Following Leinonen *et al.* (2006), we estimated P_{ST} as $P_{ST} = \delta^2_{GB}/(\delta^2_{GB} + 2*(h^2*\delta^2_{GW}))$, where δ^2_{GB} and δ^2_{GW} are the between-population and within-population variance components for a specific trait and h^2 is heritability. For stickleback, only few heritability estimates are available, which differ among the studied populations (e.g. Baumgartner, 1995; Leinonen *et al.*, 2011). We thus assumed a full trait heritability ($h^2 = 1$), which provides a conservative estimate for P_{ST} (Leinonen *et al.*, 2006). For each P_{ST}, the 95% confidence interval was established using a resampling approach of 1000 replicates. P_{ST}s were either based on the residuals of the leading principal component (PC) axis, combining all traits or a combination of traits, related to functionally different groups (defence, feeding, body shape and swimming performance) as well as separately for each trait. For cases where the 95% confidence interval exceeded zero, the directionality of a P_{ST} was further assessed by comparing the mean trait values between the different ecotypes.

We calculated the leading eigenvector (p_{max}) of the phenotypic variance–covariance matrix **P** based on a PC analysis for each habitat and freshwater system using traits that were separately size corrected for each system and the marine population. For the marine populations, we pooled both sites to obtain a better estimate of the putative ancestral state of stickleback. In addition, we calculated the overall p_{max} for each freshwater system, where we pooled fish from lake and stream sites. By pooling individuals from distinct populations and habitats, the calculated **P** matrix may differ from the **P** matrices that were separately calculated for each habitat as traits may covary between populations from different habitats even if they do not covary within either of the populations. p_{max} of such overall **P** matrices may however be compared among each other to test whether parapatric phenotypic divergence results in similar and hence predictable patterns. We compared p_{max} of two **P** matrices by calculating the angle θ between them following Schluter (1996), where θ is the inversed cosine of the dot product of two leading eigenvectors that is divided by the summed length of both eigenvectors. We estimated θ between parapatric lake and stream ecotypes for all ecotype pairs, between the freshwater ecotypes and our marine populations, as well as between our lake–stream systems. In the latter case, we pooled all populations (lake and stream) from each lake–stream system. The significance of θ between p_{max} of all comparisons was estimated using a bootstrap procedure with 1000 replicates following Berner (2009). The obtained values for θ were then correlated with the time since stickleback colonized each system using linear models. All statistical analyses were performed in R 2.15.1 (R Development Core Team, 2012).

Results

Genetic divergence

The genetic tree indicates differentiation among our studied freshwater systems, where Iceland and Switzerland form distinct clades (Fig. 2a). In the Icelandic clade, the Marine 1 population falls next to the Hraunsfjördur branch, which is consistent with the very recent origin of this system (Kristjánsson *et al.*, 2002b; Ólafsdóttir *et al.*, 2007b), whereas both Mývatn and Thingvallavatn form distinct branches. Genetic substructure among the different sampling sites was furthermore indicated in all Icelandic systems. In the Swiss clade, the Bern system falls next to the Constance system, which are both distinct from the Geneva system, reflecting their introduction and admixture history (Lucek *et al.*, 2010).

The outlier tests performed separately for each lake–stream system suggested that none of the markers deviated from neutral expectations (results not shown). Therefore, all markers were retained for the population

(a) **Genetic differentiation**

(b) **Phenotypic differentiation**
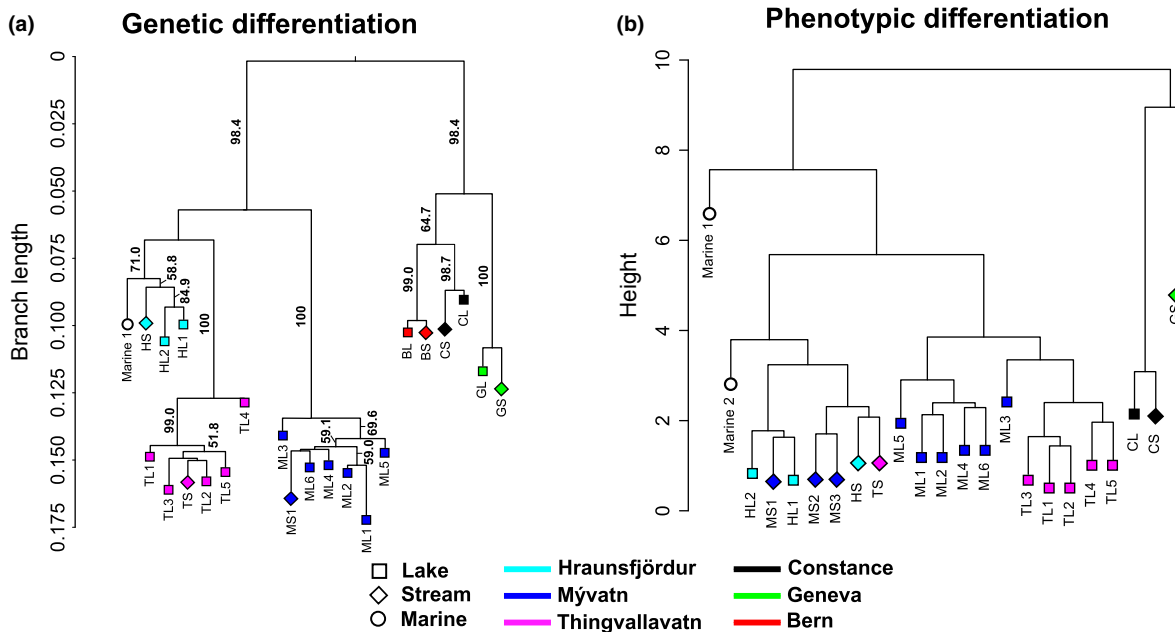
Fig. 2 Genetic and phenotypic relationship among sampling sites. Shape of tip labels indicates habitat (square: lake; diamond: stream; circle: marine), and colours represent different lake–stream systems. (a) Genetic differentiation among populations based on a neighbour-joining tree using Cavalli-Sforza distances among sampling sites included in this study (see Fig. 1), calculated from allele frequencies at 10 microsatellite loci. The tree is mid-point rooted. Numbers beside nodes indicate percent bootstrap support based on 1000 resampling replicates. Bootstrap values below 50% are not shown. (b) Dendrogram of phenotypic Mahalanobis distances among all sampling sites.

genetic analyses. Habitat-dependent parapatric genetic differentiation was highest in the Lake Geneva system in Switzerland (F_{ST}=0.053, $P < 0.001$), which also showed the greatest differences in altitude ($\Delta_{Altitude}$: 108 m) and distance to the lake (61 km). All parapatric ecotypes except Bern ($F_{ST} = 0.000$, n.s.) showed genetic differentiation (Constance: $F_{ST} = 0.018$, $P = 0.017$; Hraunsfjördur: $F_{ST} = 0.009$, $P = 0.006$; Mývatn: $F_{ST} = 0.028$, $P < 0.001$; and Thingvallavatn: $F_{ST} = 0.018$, $P = 0.009$). Pairwise F_{ST} between parapatric lake and stream populations was significantly correlated with both altitudinal differences between sites ($R^2 = 0.823$, $F_{1,4} = 18.6$, $P = 0.013$) and distance to the lake ($R^2 = 0.784$, $F_{1,4} = 14.5$, $P = 0.019$). These explanatory factors were significantly correlated with each other ($R^2 = 0.922$, $F_{1,4} = 47.2$, $P = 0.002$) and fitted the linear model equally well (ΔAICc = 1.22).

Historical contingency and divergence in freshwater

The trait-based ANOVA models all explained a significant amount of phenotypic variation (all $P < 0.001$, results not shown; Table 1). The highest proportion of nonerror variation was explained by historical contingency or differences in the current selective regimes between Iceland and Switzerland (*country*: 37.4% ± 23.0%; *system*: 35.1% ± 18.4; Table 1), where Swiss and Icelandic stickleback differed most strongly in defence-related

traits and to a lesser extent in feeding-related traits. Variation explained by *system* was highest for body-shape-related traits. System-specific components of parapatric lake–stream divergence occurred especially for feeding-related traits and to a lesser extent for body-shape-related traits as indicated by the *system × habitat* interaction. *Habitat* alone explained only a small fraction of the variance (4.4% ± 5.4%), where the traits TLP and BLA had the largest amount of variance explained.

The occurrence of individual trait-based parapatric phenotypic divergence (P_{ST}) and the overall dimensionality of parapatric divergence, measured as the number of traits with significant parapatric P_{ST}, differed among systems and countries. Similarly, the directionality of the trait divergence between lake and stream differed among Swiss and Icelandic ecotype pairs, and to a lesser extent also between ecotype pairs in different lake–stream systems within either country (Fig. 3). In the two oldest Icelandic lakes, P_{ST} exceeded zero for 14 (Thingvallavatn) and 16 (Mývatn) of 18 traits. However, even in the 50-year-old Hraunsfjördur system, this was true for six traits. In Switzerland, significant trait-specific P_{ST}s were observed only in the slightly older Constance and Geneva systems, especially for defence traits (four and seven traits, respectively), whereas P_{ST} did not exceed zero for any of the traits in the Bern system which stickleback colonized 50 years

Table 1 Nonerror variance components based on an ANOVA model with *Country* (Iceland, Switzerland), *System* (Bern, Geneva, Constance, Hraunsfjördur, Mývatn, Thingvallavatn), *Habitat* (Lake, Stream) and the interaction *System × Habitat* with their respective *F* value, degrees of freedom and significances (*P*) for each trait (see text for details). The R^2 values furthermore indicate the overall statistical support for each model (all $P < 0.001$).

	Trait	R^2	Country (%)	$F_{1,799}$	P	System (%)	$F_{4,799}$	P	Habitat (%)	$F_{1,799}$	P	System* habitat (%)	$F_{5,799}$	P
Feeding	HL	0.174	32.7	55.18	<0.001	47.2	19.90	<0.001	0.0	0.00	0.949	20.1	6.78	<0.001
	ED	0.403	60.9	329.08	<0.001	15.0	20.27	<0.001	9.2	49.67	<0.001	14.9	16.05	<0.001
	SnL	0.201	5.2	4.34	0.038	30.3	6.31	<0.001	4.7	3.90	0.049	59.8	9.95	<0.001
	UJL	0.160	33.9	60.82	<0.001	19.9	8.93	<0.001	7.1	12.68	<0.001	39.1	14.03	<0.001
	SnW	0.419	21.6	95.04	<0.001	64.3	70.79	<0.001	2.5	11.02	<0.001	11.6	10.19	<0.001
	GRL2	0.534	45.1	192.57	<0.001	33.9	36.15	<0.001	0.7	2.81	0.094	20.4	17.41	<0.001
	AL	0.527	51.2	141.28	<0.001	37.2	25.65	<0.001	3.9	10.79	0.001	7.6	4.21	<0.001
Defence	FSL	0.608	55.9	511.29	<0.001	24.7	56.41	<0.001	2.0	17.97	<0.001	17.5	32.01	<0.001
	DSL	0.507	58.1	517.04	<0.001	21.6	48.14	<0.001	3.5	31.35	<0.001	16.8	29.90	<0.001
	PSL	0.784	62.5	773.23	<0.001	20.8	64.49	<0.001	0.4	4.63	0.032	16.3	40.27	<0.001
	PGL	0.224	63.2	1835.81	<0.001	22.4	162.75	<0.001	3.3	95.75	<0.001	11.1	64.58	<0.001
Body	BD1	0.356	8.4	16.88	<0.001	41.8	21.04	<0.001	2.0	4.02	0.045	47.9	19.30	<0.001
	BD2	0.343	21.5	32.74	<0.001	36.0	13.70	<0.001	1.5	2.22	0.136	41.0	12.46	<0.001
	CPL	0.094	9.3	53.45	<0.001	88.5	127.49	<0.001	0.0	0.02	0.901	2.2	2.59	0.025
	PGW	0.183	75.4	618.94	<0.001	20.2	41.55	<0.001	0.0	0.32	0.574	4.3	7.11	<0.001
	TLP	0.355	0.5	1.13	0.287	23.5	13.59	<0.001	11.0	25.46	<0.001	65.0	30.02	<0.001
	BLA	0.348	27.5	121.42	<0.001	36.6	40.35	<0.001	22.1	97.33	<0.001	13.8	12.15	<0.001
	BLD	0.257	40.1	167.44	<0.001	47.8	49.97	<0.001	5.9	24.57	<0.001	6.2	5.17	<0.001

ago. The PC-based P_{ST}, combining either all traits or only defence-related traits, exceeded zero in all but the two youngest systems (Fig. 3). In contrast, P_{ST} exceeded zero only in Mývatn for feeding-related morphology and in the Icelandic systems for body-shape- and swimming performance-related traits. The magnitude of P_{ST} among parapatric ecotypes was not statistically associated with the altitudinal difference, with the waterway distance between sites, or with the age of a system for any trait combinations (all $P > 0.1$, results not shown).

The angle θ between p_{max} from parapatric lake and stream populations based on all phenotypic traits differed from zero in all cases except Thingvallavatn (Fig. 4, Table 2), whereas it was greatest in the two other Icelandic systems (Table 2), whose p_{max} was significantly differentiated from all other freshwater systems in the pairwise comparisons (Table 3). In the Swiss systems, the parapatric θs were significantly different from zero (Table 2), whereas p_{max} did not differ among the systems (Table 3). When traits were analysed by functional categories, the angle θ between parapatric lake–stream p_{max} differed across traits and systems (Fig. 4, Table 2). θ between parapatric ecotypes differed especially for feeding-related traits, albeit to a small degree (average θ: 9.3° ± 2.3° SD), whereas parapatric p_{max} differed less commonly for defence-, body-depth- and swimming performance-related traits (Table 2). p_{max} was furthermore comparable among lake–stream systems for feeding- and defence-related traits as suggested by the nonsignificant angle θ between them (Table 3). None of the angles between parapatric ecotypes were

statistically correlated with the time since stickleback colonization, the altitudinal difference or the geographical distance between the lake and the stream populations (all $P > 0.1$, results not shown).

Parallel adaptation trumps historical contingency late but not early in ecotype formation

Mahalanobis distances showed overall consistent morphological differentiation between Swiss and Icelandic freshwater stickleback populations (Fig. 2b). Despite the aforementioned evidence for consistent parapatric divergence, all populations of young lake–stream pairs, that is Hraunsfjördur in Iceland and all the Swiss systems, clustered by historical lineage rather than by ecotype. In contrast, the populations from the two oldest lake systems, that is Mývatn and Thingvallavatn, clustered strongly by ecotype despite being genetically more strongly differentiated than the lineages with young lake–stream pairs (Fig. 2a).

Phenotypic divergence during the marine–freshwater transition

The angle θ between p_{max} of the different freshwater lake populations differed significantly in all replicates when pooling all traits (average θ: 59.7° ± 6.5° SD; Fig. 5, Table 2), which was not true for stream populations (average θ: 29.7° ± 21.0° SD). Using only feeding-related traits, the freshwater p_{max} differs commonly from the marine one with relatively low angles θ (lake
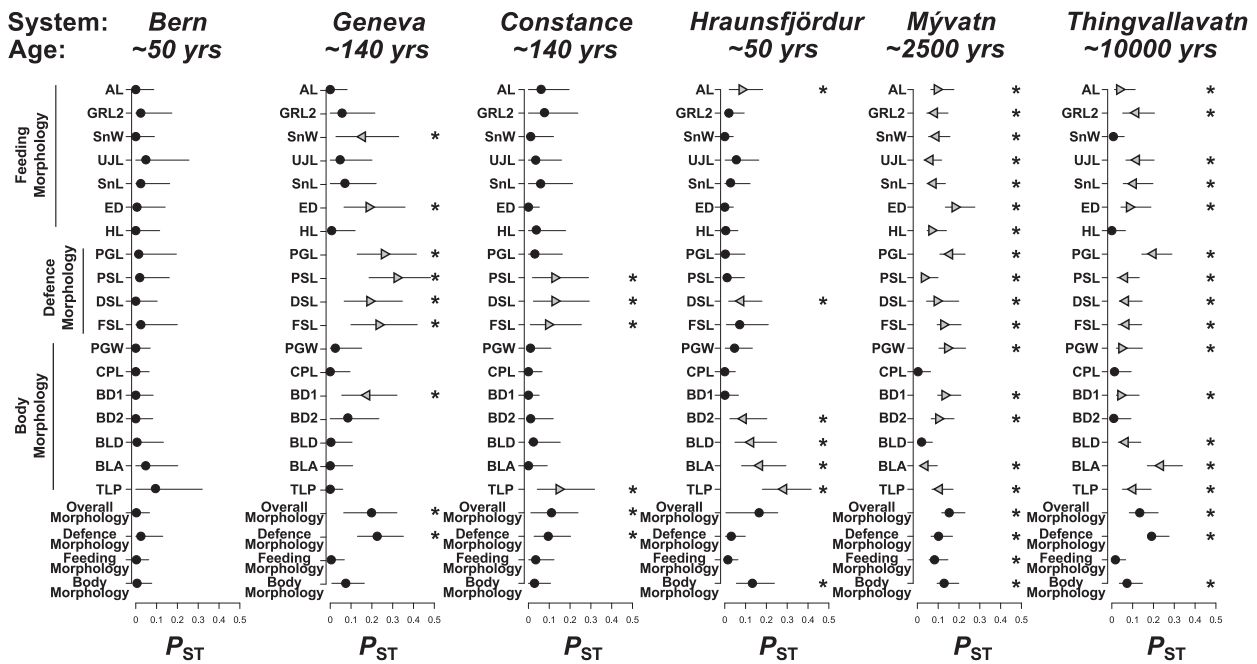
Fig. 3 Pairwise phenotypic divergence between lake and stream ecotypes ($P_{ST} \pm$ 95% CI) for each system, calculated for each trait separately, for all traits combined and for functionally distinct trait groups. P_{ST} for trait groups is based on scores of the first principal component axis for either all traits combined, defence-related traits, feeding-related traits, or body-shape-related traits (see text for details). Asterisks indicate cases where the 95% confidence interval for P_{ST} exceeds 0. For the latter traits, triangles indicate the directionality of the pairwise divergence, where a trait is larger (pointing right) or smaller (pointing left) in lake fish in comparison with stream fish. For a description of each trait and its abbreviation, see the main text and Fig. S2.

vs. marine: average θ = 9.2° ± 3.3° SD; stream vs. marine: average θ = 7.8° ± 2.0° SD). In contrast, the freshwater p_{max} based on defence- or body-shape- and swimming performance-related traits differed less commonly from the marine p_{max} (Fig. 5, Table 2). In all but one case (marine vs. lake populations using all traits combined: $F_{1,4}$ = 15.9, P = 0.016), the observed angle θ between a freshwater-derived p_{max} and the marine p_{max} was not statistically correlated with the relative age of each freshwater system (all P > 0.1, results not shown).

Discussion

The extent of parallel evolution of phenotypically similar ecotypes depends on the genetic constraints, the selective environment and the time for evolution to act (Schluter & Nagel, 1995; Langerhans & DeWitt, 2004; Nosil et al., 2009; Kaeuffer et al., 2012; Nosil, 2012). Nonparallel phenotypic features may thus occur between independently evolved yet ecologically similar ecotypes. The extent of convergent evolution and hence the degree to which two independent populations become more similar may further depend on the dimensionality with which ecotypic divergence is being measured. Rapid evolution may be inferred to be parallel when only few traits are being measured (e.g. Schluter et al., 2004), whereas parallel divergence in the multi-variate phenotype, leading to overall phenotypic convergence of parallel evolved ecotypes, may need much longer time (Young et al., 2009; Kolbe et al., 2011).

Comparing the phenotypic variance–covariance (**P**) matrices of different stickleback freshwater ecotypes and their marine ancestors, we find that phenotypic divergence can result in parallel lines of least resistance (p_{max}) both for the marine–freshwater transition (Fig. 5, Table 2) and the subsequent ecotype formation within freshwater (Fig. 4, Table 3). The extent and parallelism of parapatric ecotype formation within freshwaters seem to be driven by historical contingency, potential differences in the divergent selective regimes between lake and streams and the time available for evolution, where much of the phenotypic variation is explained by differences between Icelandic and Swiss sticklebacks (Fig. 2, Table 1). Parallel ecotypic divergence may trump historical contingency only in the oldest lakes, where the divergent selective regimes may moreover be strongest, with an increased dimensionality of ecotypic differentiation and a clustering of ecotypes despite being genetically very distinct (Figs 2 and 3).

The evolution of freshwater stickleback

The evolutionary transition between the marine and freshwater environment has been repeatedly studied in

Table 2 Angle between the leading eigenvector (θ) of either the marine population and the freshwater stream or the lake populations, as well as the angle between the parapatric lake–stream populations. The analysis was either performed combining all traits or separately for defence-, feeding- or body-shape-related traits. *P*-values are based on 1000 bootstrap replicates with the one-tailed 97.5% confidence limit (CL) indicated. Significant *P*-values (< 0.05) are given in bold; *P*-values (< 0.1 < *P* < 0.05) are in italic.

	Marine vs. Stream			Marine vs. Lake			Parapatric Lake–Stream		
	θ (°)	*P*	97.5% CL	θ (°)	*P*	97.5% CL	θ (°)	*P*	97.5% CL
All traits									
Bern	**84.89**	**<0.001**	**4.66**	**52.10**	**<0.001**	**33.99**	**58.67**	**<0.001**	**29.67**
Geneva	26.11	0.275	59.49	**56.77**	**<0.001**	**30.42**	**52.36**	**<0.001**	**34.42**
Constance	**70.65**	**<0.001**	**18.35**	**58.68**	**<0.001**	**29.44**	**58.14**	**<0.001**	**30.79**
Hraunsfjördur	**55.61**	**<0.001**	**31.15**	**55.97**	**<0.001**	**31.78**	**80.52**	**<0.001**	**8.55**
Mývatn	38.89	0.134	48.87	**65.11**	**<0.001**	**18.51**	**88.23**	**<0.001**	**0.87**
Thingvallavatn	**73.21**	**<0.001**	**15.69**	**69.72**	**<0.001**	**16.19**	29.88	0.243	50.52
Defence traits									
Bern	8.45	0.376	25.42	**39.34**	**0.016**	**36.00**	*31.69*	*0.058*	*46.52*
Geneva	5.67	0.548	29.68	14.48	0.147	27.16	*19.95*	*0.088*	*28.96*
Constance	14.04	0.151	29.49	27.41	0.147	50.16	17.10	0.298	59.16
Hraunsfjördur	**31.90**	**0.008**	**25.14**	2.02	0.921	38.75	**30.41**	**0.001**	**21.32**
Mývatn	7.16	0.517	35.50	*17.63*	*0.067*	*20.77*	11.43	0.203	27.26
Thingvallavatn	7.92	0.455	39.29	*18.84*	*0.054*	*20.98*	19.36	0.105	42.79
Feeding traits									
Bern	7.06	0.182	14.05	4.83	0.106	6.68	5.63	0.335	13.75
Geneva	4.81	0.289	8.48	**12.28**	**0.003**	**8.14**	**12.11**	**0.009**	**10.02**
Constance	**9.88**	**0.011**	**7.28**	**9.60**	**0.003**	**6.36**	**9.52**	**0.033**	**10.07**
Hraunsfjördur	*7.67*	*0.081*	*10.42*	**13.53**	**0.009**	**11.51**	*11.36*	*0.071*	*14.47*
Mývatn	**7.11**	**0.021**	**6.66**	**8.06**	**0.006**	**5.99**	**8.13**	**0.012**	**7.45**
Thingvallavatn	**10.01**	**0.034**	**10.73**	**6.79**	**0.007**	**5.90**	*8.83*	*0.055*	*10.23*
Body traits/Swimming performance									
Bern	**55.25**	**<0.001**	**32.94**	**56.50**	**<0.001**	**31.93**	13.80	0.437	38.91
Geneva	20.27	0.359	54.41	**87.37**	**<0.001**	**1.24**	**74.39**	**<0.001**	**13.89**
Constance	34.49	0.183	52.13	**77.28**	**<0.001**	**11.80**	**79.02**	**<0.001**	**9.53**
Hraunsfjördur	*39.42*	*0.094*	*47.27*	36.68	0.113	49.47	23.51	0.299	58.68
Mývatn	22.23	0.583	65.40	**38.47**	**0.042**	**43.57**	32.12	0.323	55.53
Thingvallavatn	**62.98**	**<0.001**	**25.71**	**39.41**	**0.035**	**42.75**	27.46	0.252	57.98

stickleback (e.g. Kristjánsson, 2005; Leinonen *et al.*, 2006, 2011; Wund *et al.*, 2008; Berner *et al.*, 2010b; Jones *et al.*, 2012a; Voje *et al.*, 2013). Colonizing freshwater habitats requires adaptation to new selective regimes, which may differ between distinct freshwater habitats (Gross, 1978; Gross & Anderson, 1984; Reimchen, 1994; Berner *et al.*, 2009, 2010b; Lucek *et al.*, 2014). The degree of phenotypic divergence from an ancestral-like marine population may consequently differ between distinct habitats and among traits due to differences in the selective regime and the colonization history. In concordance, we find that the degree of phenotypic and genetic differentiation differs among systems (Fig. 2). In the genetic tree, the two oldest lakes Mývatn and Thingvallavatn form distinct genetic clusters with the longest branch lengths, whereas the Swiss populations, albeit being genetically distinct from each other, form a separate branch. Conversely, ecotype-specific clusters occur for the old lakes in the phenotypic tree, whereas all populations of young ecotype pairs cluster in concordance with their genetic lineage. Lineage-dependent phenotypic constraints may have

consequently been retained in Switzerland as the Constance and Geneva systems were colonized about 140 years ago by genetically distinct freshwater lineages, and the Bern system lies in a hybrid zone between different lineages (Lucek *et al.*, 2010). In contrast, Icelandic freshwater populations likely derive from a common marine population, where in some cases, gene flow from the ancestral marine population may still be possible (Ólafsdóttir *et al.*, 2007c; Fig. 2a).

Historical contingency or potential differences in current selective regimes for different trait categories are further indicated by the changes in the **P** matrices among the different marine–freshwater comparisons. Here, patterns for the p_{max} of freshwater stickleback vary generally among all comparisons (Table 2, Fig. 5). Whereas p_{max} of freshwater populations differ commonly from the marine p_{max} when all phenotypic traits were combined, they are similar to the marine one in all comparisons for antipredator-related traits. Predator communities are thought to differ though, where marine and freshwater lake populations experience a predation regime dominated by gape-limited predators

Table 3 Pairwise angle (θ) between the leading eigenvector among freshwater systems (in degrees, lower triangular) with their corresponding *P*-value (upper triangular). Axes were calculated using either all phenotypic traits, or a subset of traits related to defence, feeding or body shape, respectively.

	Bern	Geneva	Constance	Hraunsfjördur	Mývatn	Thingvallavatn
All traits						
Bern	–	0.627	0.973	*0.078*	**0.007**	**0.025**
Geneva	6.31	–	0.635	**0.009**	**<0.001**	*0.081*
Constance	0.53	5.78	–	*0.052*	**0.001**	*0.055*
Hraunsfjördur	*21.85*	**28.16**	*22.38*	–	0.498	**<0.001**
Mývatn	**29.56**	**35.87**	**30.09**	7.72	–	**<0.001**
Thingvallavatn	**28.79**	*22.48*	*28.26*	**50.64**	**58.35**	–
Defence traits						
Bern	–	0.287	0.214	0.895	0.136	0.278
Geneva	11.75	–	0.737	0.232	0.308	0.939
Constance	14.59	2.85	–	0.194	0.596	0.774
Hraunsfjördur	1.28	10.47	13.32	–	*0.072*	0.257
Mývatn	20.26	8.52	5.67	*18.98*	–	0.304
Thingvallavatn	12.34	0.59	2.26	11.06	7.92	–
Feeding traits						
Bern	–	*0.059*	0.133	0.129	0.252	0.238
Geneva	*6.48*	–	0.395	0.809	0.259	0.279
Constance	3.89	2.59	–	0.597	0.576	0.765
Hraunsfjördur	5.73	0.75	1.84	–	0.403	0.447
Mývatn	2.49	3.99	1.40	3.24	–	0.738
Thingvallavatn	3.20	3.29	0.70	2.53	0.70	–
Body traits/Swimming performance						
Bern	–	**<0.001**	**<0.001**	0.406	0.284	0.260
Geneva	**60.59**	–	0.715	**<0.001**	**<0.001**	**<0.001**
Constance	**65.21**	4.62	–	**<0.001**	**<0.001**	**<0.001**
Hraunsfjördur	9.71	**50.88**	**55.50**	–	0.564	0.752
Mývatn	18.32	**42.27**	**46.90**	8.60	–	0.765
Thingvallavatn	13.66	**46.93**	**51.55**	3.95	4.65	–

Significant *P*-values (< 0.05) are given in bold; *P*-values ($< 0.1 < P < 0.05$) are in italic.

° P < 0.1   * P < 0.05   ** P < 0.01   *** P < 0.001
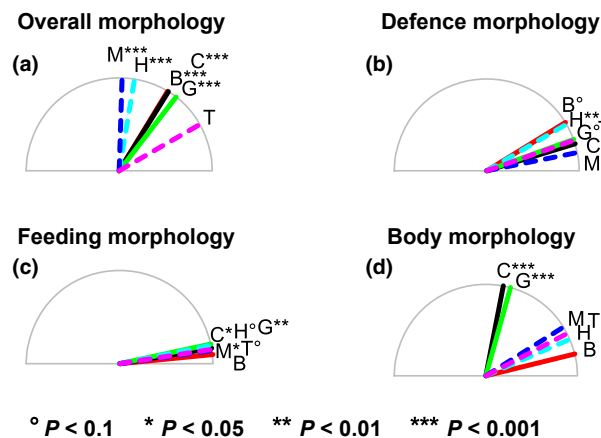
Fig. 4 Angles between the major axis of phenotypic variation (p_{max}) in parapatric lake vs. stream populations. Angles were calculated including either all phenotypic traits (a) or a subset of defence (b), feeding (c) or body shape/swimming performance related traits (d). Letters indicate the respective system: B, Bern (red); C, Constance (black); G, Geneva (green); H, Hraunsfjördur (blue); M, Mývatn (dark blue); and T, Thingvallavatn (pink). Dashed lines denote lake–stream systems from Iceland, solid lines systems from Switzerland.

such as birds and piscivorous fish (Gross, 1978; Reimchen, 1992), which shifts to increased insect predation in freshwater streams (Reimchen, 1994; Marchinko, 2009). Invertebrate predation may however be negligible in Iceland (Lucek *et al.*, 2012b), and empirical evidence for the role of invertebrate predators as a source of selection is mixed for Swiss populations (Zeller *et al.*, 2012a,b).

For trophic morphology on the other hand, freshwater p_{max} differ commonly from the marine one with small but significant angles θ, suggesting that the marine–freshwater transition may be generally associated with a change in the **P** matrix (Fig. 5). Habitat-dependent ecotypic differentiation in stickleback is indeed thought to be coupled with a change in diet and trophic morphology, where marine and some freshwater lake populations forage commonly on zooplankton in contrast to stream fish and some lake populations that feed on benthic prey (Gross & Anderson, 1984; Berner *et al.*, 2009; Kaeuffer *et al.*, 2012; Lucek *et al.*, 2012a). The small but significant angles θ, which we observe especially for lake populations, are consistent with prior findings in Canadian stickleback, where freshwater lake populations have been shown to have a p_{max} that has diverged from
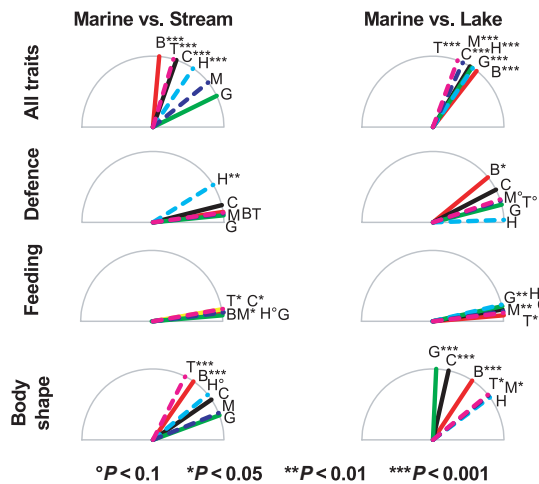
Fig. 5 Angles between the major axis of phenotypic variation (p_{max}) between the marine population and either the freshwater stream populations or the lake populations. Angles were calculated using (from top to bottom) the following: all phenotypic traits or a subset of defence-, feeding- or body-shape/swimming performance-related traits. For the overall divergence, indicated vectors are scaled according to the eigenvalue of the leading axis. Letters indicate the respective system: B, Bern (red); C, Constance (black); G, Geneva (green); H, Hraunsfjördur (blue); M, Mývatn (dark blue); and T, Thingvallavatn (pink). Dashed lines denote lake–stream systems from Iceland, straight lines systems from Switzerland.

the marine one, involving a shift in gill raker lengths (Berner *et al.*, 2010b). Lastly, differences in the selective regimes between our studied systems and countries may account for the observed changes in p_{max} for body–shape- and swimming-related traits that are linked to different foraging strategies in lakes and streams (Hendry & Taylor, 2004; Reid & Peichel, 2010; Hendry *et al.*, 2011). The wide range for θ for both the overall system and stream populations may further reflect different selection regimes for each stream due to environmental differences such as differences in flow regimes (Steppan *et al.*, 2002; Ravinet *et al.*, 2013).

Contingency, selection and parallelism during lake–stream divergence

Both the occurrence and the extent of parapatric divergence depend mainly on the underlying environmental and selective gradient and the time for evolution to act (Endler, 1977; Doebeli & Dieckmann, 2003; Nosil *et al.*, 2009). Parallel parapatric divergence is consequently only expected under comparable selective regimes (Kaeuffer *et al.*, 2012) where selection is acting on a similar gene pool (Barrett & Schluter, 2008). The repeated formation of parapatric lake–stream freshwater stickleback systems has been proposed to provide such a case (Reimchen *et al.*, 1985; Thompson *et al.*, 1997; Hendry & Taylor, 2004; Berner *et al.*, 2009; Lucek *et al.*,

2013). However, recent studies find nonparallelisms in the realized divergence that occur both on smaller geographical scales (Kaeuffer *et al.*, 2012; Hendry *et al.*, 2013; Lucek *et al.*, 2013; Ravinet *et al.*, 2013) and between continents (Berner *et al.*, 2010a). In the latter case, the authors suggested that genomic constraints could be responsible for the observed lower degree of divergence among Swiss populations and the evolutionary younger Atlantic stickleback lineage in general, where only the Constance system showed a level of divergence that is comparable to older Canadian systems (Berner *et al.*, 2010a; Ravinet *et al.*, 2013; but see Lucek *et al.*, 2013). However, the respective **P** matrices have not been compared.

Our results suggest that the evolution of parapatric lake–stream populations in stickleback can result in common and hence predictable p_{max} independent of the age of a system as it is indicated by the nonsignificant angles between the overall p_{max} of different lake–stream systems for defence- and feeding-related traits (Table 3). Thus, ecotype formation along parallel axes may start quickly. However, although parapatric lake–stream systems share similar p_{max}, only a relatively small fraction of the overall phenotypic variation can be attributed to parallel habitat-dependent differentiation (Table 1), where the directionality of parapatric divergence often differs between ecotype pairs in Switzerland and Iceland and sometimes also between pairs from different lake–stream systems within each country (Fig. 3). In contrast, a much larger fraction is explained by the system and habitat interaction and thus the combined effect of system-related historical contingency and/or system-related selection with ecotypic divergence (Langerhans & DeWitt, 2004; Eroukhmanoff *et al.*, 2009; Kaeuffer *et al.*, 2012). The increased dimensionality of parapatric differentiation in the two oldest lake systems, Mývatn and Thingvallavatn (Fig. 3), may either reflect stronger habitat-dependent divergent selection or that longer time is needed for a parallel evolutionary response to similar divergent selection during ecotype formation to trump historical contingency (Nosil *et al.*, 2009; Young *et al.*, 2009; Nosil, 2012). The increase in dimensionality further suggests that parallel phenotypic divergence is associated with increased phenotypic integration (Fig. 3), which is consistent with studies on older adaptive radiations that show increased convergence in multivariate trait dimensions in comparison with younger radiations (Young *et al.*, 2009; Kolbe *et al.*, 2011).

In contrast to the observed phenotypic divergence and convergence, the degree of neutral parapatric genetic differentiation is correlated with the parapatric environmental gradient rather than the evolutionary age of the system. Altitudinal gradients have similarly been found to explain the degree of parapatric genetic divergence in other freshwater systems (Caldera & Bolnick, 2008; Ravinet *et al.*, 2013) as well as during the

marine and freshwater transition (Deagle *et al.*, 2013, K. Lucek, B. K. Kristjánsson, S. Skúlason & O. Seehausen, unpublished) and may be linked to physical barriers restricting the potential for gene flow.

Rapid evolution vs. plasticity

Although phenotypic divergence was greatest in the oldest lakes, the observed differentiation in p_{max} was not associated with our studied temporal gradient. Hence, plasticity could have initially promoted the colonization of freshwater habitats (Smith & Skúlason, 1996) by rapidly shifting p_{max} (Lande, 2009; Draghi & Whitlock, 2012). Marine stickleback are known to be phenotypically plastic, allowing them to respond to different diets readily when colonizing new freshwater environments (Wund *et al.*, 2008). Plasticity can furthermore evolve in freshwater to initially promote a generalist lifestyle where divergent selection may then lead to canalization and a reduction in plasticity (Svanbäck & Schluter, 2012), matching theoretical predictions (Lande, 2009; Thibert-Plante & Hendry, 2011). In theory, however, phenotypic plasticity and thus p_{max} may evolve quite fast, that is over fewer generations than those separating the lake and stream populations in our youngest system (Lande, 2009; Draghi & Whitlock, 2012). p_{max} based on phenotypically plastic traits may thus align if populations experience a comparable selective regime as we observe for defence- and feeding-related traits.

Phenotypic shifts during the marine–freshwater transition as well as between distinct freshwater habitats in stickleback have similarly been suggested to occur through recurrent selection on standing genetic variation in the marine population (Deagle *et al.*, 2012; Jones *et al.*, 2012b). This is especially true for antipredator-related phenotypic shifts, where selection drives phenotypic divergence over only a few generations (Bell *et al.*, 2004; Barrett *et al.*, 2008; Schluter & Conte, 2009) and may similarly account for phenotypic divergence in other genetically determined traits such as gill rakers (Hermida *et al.*, 2002). Our observed parapatric divergence as well as the changes in the **P** matrix for these traits may therefore be a combined result of both plasticity and adaptation from standing genetic variation (Wund *et al.*, 2008; Eroukhmanoff & Svensson, 2011; Lucek *et al.*, 2014). Indeed, empirical evidence suggests that although p_{max} and g_{max} are correlated in stickleback, p_{max} can only approximate g_{max} and hence the underlying evolutionary constraints (Leinonen *et al.*, 2011).

Conclusions

Our results suggest that parapatric ecotype formation can result in parallel and hence predictable p_{max} for some trait combinations, that is trophic morphology, but that the directionality of change may differ for others due to historical contingency or environmental effects. Whereas changes in the **P** matrix during the marine–freshwater transition seem to evolve independently of our studied temporal axis, both the extent and the dimensionality of parapatric ecotype formation depend on the available time for evolution to occur. Thus, evolutionary changes towards novel adaptive peaks may occur readily during ecotype formation and may be aided by phenotypic plasticity, yet convergent phenotypic evolution needs time to overcome contingency.

Acknowledgments

References

Antao, T., Lopes, A., Lopes, R.J., Beja-Pereira, A. & Luikart, G. 2008. LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinf.* **9**: 323.

Arnold, S.J. & Phillips, P.C. 1999. Hierarchical comparison of genetic variance-covariance matrices. II. Coastal-inland divergence in the garter snake, *Thamnophis elegans*. *Evolution* **53**: 1516–1527.

Arnold, S.J., Bürger, R., Hohenlohe, P.A., Ajie, B.C. & Jones, A.G. 2008. Understanding the evolution and stability of the G-matrix. *Evolution* **62**: 2451–2461.

Bacigalupe, L.D. 2009. Biological invasions and phenotypic evolution: a quantitative genetic perspective. *Biol. Inv.* **11**: 2243–2250.

Barrett, R.D.H. & Schluter, D. 2008. Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**: 38–44.

Barrett, R.D.H., Rogers, S.M. & Schluter, D. 2008. Natural selection on a major armor gene in threespine stickleback. *Science* **322**: 255–257.

Baumgartner, J.V. 1995. Phenotypic, genetic, and environmental integration of morphology in a stream population of the threespine stickleback, *Gasterosteus aculeatus. Can. J. Fish. Aquat. Sci.* **52**: 1307–1317.

Bell, M.A. & Foster, S.A. 1994. *The Evolutionary Biology of the Threespine Stickleback*. Oxford University Press, Oxford, UK.

Bell, M.A., Aguirre, W. & Buck, N. 2004. Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution* **58**: 814–824.

Berner, D. 2009. Correction of a bootstrap approach to testing for evolution along lines of least resistance. *J. Evol. Biol.* **22**: 2563–2565.

Berner, D., Adams, D.C., Grandchamp, A.-C. & Hendry, A.P. 2008. Natural selection drives patterns of lake-stream divergence in stickleback foraging morphology. *J. Evol. Biol.* **21**: 1653–1665.

Berner, D., Grandchamp, A.-C. & Hendry, A.P. 2009. Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. *Evolution* **63**: 1740–1753.

Berner, D., Roesti, M., Hendry, A.P. & Salzburger, W. 2010a. Constraints on speciation suggested by comparing lake-stream stickleback divergence across two continents. *Mol. Ecol.* **19**: 4963–4978.

Berner, D., Stutz, W.E. & Bolnick, D.I. 2010b. Foraging trait (co)variances in stickleback evolve deterministically and do not predict trajectories of adaptive diversification. *Evolution* **64**: 2265–2277.

Caldera, E.J. & Bolnick, D.I. 2008. Effects of colonization history and landscape structure on genetic variation within and among threespine stickleback (*Gasterosteus aculeatus*) populations in a single watershed. *Evol. Ecol. Res.* **10**: 575–598.

Chapuis, E., Martin, G. & Goudet, J. 2008. Effects of selection and drift on G matrix evolution in a heterogeneous environment: a multivariate Qst-Fst Test with the freshwater snail *Galba truncatula*. *Genetics* **180**: 2151–2161.

Cheverud, J.M. 1988. A comparison of genetic and phenotypic correlations. *Evolution* **42**: 958–968.

Deagle, B.E., Jones, F.C., Chan, Y.F., Absher, D.M., Kingsley, D.M. & Reimchen, T.E. 2012. Population genomics of parallel phenotypic evolution in stickleback across stream-lake ecological transitions. *P. R. Soc. B* **279**: 1277–1286.

Deagle, B.E., Jones, F.C., Absher, D.M., Kingsley, D.M. & Reimchen, T.E. 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.

Doebeli, M. & Dieckmann, U. 2003. Speciation along environmental gradients. *Nature* **421**: 259–264.

Draghi, J.A. & Whitlock, M.C. 2012. Phenotypic plasticity facilitates mutational variance, genetic variance, and evolvability along the major axis of environmental variation. *Evolution* **66**: 2891–2902.

Einarsson, A., Stefansdottir, G., Johannesson, H., Olafsson, J., Gislason, G., Wakana, I. *et al.* 2004. The ecology of Lake Myvatn and the River Laxa: variation in space and time. *Aquat. Ecol.* **38**: 317–348.

Endler, J.A. 1977. *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, NY, USA.

Eroukhmanoff, F. & Svensson, E.I. 2011. Evolution and stability of the G-matrix during the colonization of a novel environment. *J. Evol. Biol.* **24**: 1363–1373.

Eroukhmanoff, F., Hargeby, A., Arnberg, N.N., Hellgren, O., Bensch, S. & Svensson, E.I. 2009. Parallelism and historical contingency during rapid ecotype divergence in an isopod. *J. Evol. Biol.* **22**: 1098–1110.

Falconer, D.S. 1989. *Introduction to Quantitative Genetics*, 3rd edn. John Wiley & Sons, New York, NY, USA.

Felsenstein, J. 2012. PHYLIP (Phylogeny Inference Package) version 3.69. Distributed by the author. Department of Genome Sciences. University of Washington, Seattle, WA.

Gross, H.P. 1978. Natural selection by predators on defensive apparatus of the three-spined stickleback, *Gasterosteus aculeatus* L. *Can. J. Zool.* **56**: 398–413.

Gross, H.P. & Anderson, J.M. 1984. Geographic variation in the gillrakers and the diet of European threespine stickleback, *Gasterosteus aculeatus*. *Copeia* **1**: 87–97.

Hendry, A.P. & Taylor, E.B. 2004. How much of the variation in adaptive divergence can be explained by gene flow? An

evaluation using lake-stream stickleback pairs. *Evolution* **58**: 2319–2331.

Hendry, A.P., Hudson, K., Walker, J.A., Räsänen, K. & Chapman, L.J. 2011. Genetic divergence in morphology-performance mapping between Misty Lake and inlet stickleback. *J. Evol. Biol.* **24**: 23–35.

Hendry, A.P., Hendry, A.S. & Hendry, C.A. 2013. Hendry vineyard stickleback: testing for contemporary lake-stream divergence. *Evol. Ecol. Res.* **15**: 1–32.

Hermida, M., Fernandez, C., Amaro, R. & San Miguel, E. 2002. Heritability and "evolvability" of meristic characters in a natural population of *Gasterosteus aculeatus*. *Can. J. Zool.* **80**: 532–541.

Jones, A.G.A., Arnold, S.J.S. & Bürger, R. 2004. Evolution and stability of the G-matrix on a landscape with a moving optimum. *Evolution* **58**: 1639–1654.

Jones, F.C., Chan, Y.F., Schmutz, J., Grimwood, J., Brady, S.D., Southwick, A.M. *et al.* 2012a. A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Curr. Biol.* **22**: 83–90.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, R. *et al.* 2012b. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**: 55–61.

Kaeuffer, R., Peichel, C.L., Bolnick, D.I. & Hendry, A.P. 2012. Parallel and nonparallel aspects of ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. *Evolution* **66**: 402–418.

Kolbe, J.J., Revell, L.J., Szekely, B., Brodie, E.D. & Losos, J.B. 2011. Convergent evolution of phenotypic integration and its alignment with morphological diversification in Caribbean *Anolis* ecomorphs. *Evolution* **65**: 3608–3624.

Kristjánsson, B.K. 2005. Rapid morphological changes in threespine stickleback, *Gasterosteus aculeatus*, in freshwater. *Environ. Biol. Fish* **74**: 357–363.

Kristjánsson, B.K., Skúlason, S. & Noakes, D. 2002a. Morphological segregation of Icelandic threespine stickleback (*Gasterosteus aculeatus* L). *Biol. J. Linn. Soc.* **76**: 247–257.

Kristjánsson, B.K., Skúlason, S. & Noakes, D. 2002b. Rapid divergence in a recently isolated population of threespine stickleback (*Gasterosteus aculeatus* L.). *Evol. Ecol. Res.* **4**: 659–672.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain – body size allometry. *Evolution* **33**: 402–416.

Lande, R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *J. Evol. Biol.* **22**: 1435–1446.

Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1210–1226.

Lande, R. & Shannon, S. 1996. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* **50**: 434–437.

Langerhans, R.B. & DeWitt, T.J. 2004. Shared and unique features of evolutionary diversification. *Am. Nat.* **164**: 335–349.

Leinonen, T., Cano, J.M., Mäkinen, H.S. & Merilä, J. 2006. Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *J. Evol. Biol.* **19**: 1803–1812.

Leinonen, T., Cano, J.M. & Merilä, J. 2011. Genetics of body shape and armour variation in threespine sticklebacks. *J. Evol. Biol.* **24**: 206–218.

Lucek, K., Roy, D., Bezault, E., Sivasundar, A. & Seehausen, O. 2010. Hybridization between distant lineages increases

adaptive variation during a biological invasion: stickleback in Switzerland. *Mol. Ecol.* **19**: 3995–4011.

Lucek, K., Sivasundar, A. & Seehausen, O. 2012a. Evidence of adaptive evolutionary divergence during biological invasion. *PLoS ONE* **7**: e49377.

Lucek, K., Haesler, M.P. & Sivasundar, A. 2012b. When phenotypes do not match genotypes – unexpected phenotypic diversity and potential environmental constraints in Icelandic stickleback. *J. Hered.* **103**: 579–584.

Lucek, K., Sivasundar, A., Roy, D. & Seehausen, O. 2013. Repeated and predictable patterns of ecotypic differentiation during a biological invasion: lake-stream divergence in parapatric Swiss stickleback. *J. Evol. Biol.* **26**: 2691–2709.

Lucek, K., Sivasundar, A. & Seehausen, O. 2014. Disentangling the role of phenotypic plasticity and genetic divergence in contemporary ecotype formation during a biological invasion. *Evolution*. doi: 10.1111/evo.12443.

Mäkinen, H.S., Cano, J.M. & Merilä, J. 2006. Genetic relationships among marine and freshwater populations of the European three-spined stickleback (*Gasterosteus aculeatus*) revealed by microsatellites. *Mol. Ecol.* **15**: 1519–1534.

Marchinko, K.B. 2009. Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. *Evolution* **63**: 127–138.

Marroig, G. & Cheverud, J.M. 2005. Size as a line of least evolutionary resistance: diet and adaptive morphological radiation in New World monkeys. *Evolution* **59**: 1128–1142.

McKinnon, J.S. & Rundle, H. 2002. Speciation in nature: the threespine stickleback model systems. *Trends Ecol. Evol.* **17**: 480–488.

Meirmans, P.G. & Van Tienderen, P. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* **4**: 792–794.

Mori, S. & Takamura, N. 2004. Changes in morphological characteristics of an introduced population of the threespine stickleback *Gasterosteus aculeatus* in Lake Towada, northern Japan. *Ichthyol. Res.* **51**: 295–300.

Moser, D., Roesti, M. & Berner, D. 2012. Repeated lake-stream divergence in stickleback life history within a Central European lake basin. *PLoS ONE* **7**: e50620.

Nosil, P. 2012. *Ecological Speciation*. Oxford University Press, Oxford, UK.

Nosil, P., Harmon, L.J. & Seehausen, O. 2009. Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol.* **24**: 145–156.

Ólafsdóttir, G.A., Snorrason, S.S. & Ritchie, M.G. 2007a. Postglacial intra-lacustrine divergence of Icelandic threespine stickleback morphs in three neovolcanic lakes. *J. Evol. Biol.* **20**: 1870–1881.

Ólafsdóttir, G.A., Snorrason, S.S. & Ritchie, M.G. 2007b. Parallel evolution? Microsatellite variation of recently isolated marine and freshwater three-spined stickleback. *J. Fish Biol.* **70**: 125–131.

Ólafsdóttir, G.A., Snorrason, S.S. & Ritchie, M.G. 2007c. Morphological and genetic divergence of intralacustrine stickleback morphs in Iceland: a case for selective differentiation? *J. Evol. Biol.* **20**: 603–616.

Peichel, C.L., Nereng, K.S., Ohgi, K.A., Cole, B.L., Colosimo, P.F., Buerkle, C.A. *et al.* 2001. The genetic architecture of divergence between threespine stickleback species. *Nature* **414**: 901–905.

Pigliucci, M., Cammell, K. & Schmitt, J. 1999. Evolution of phenotypic plasticity a comparative approach in the phyloge-

netic neighbourhood of *Arabidopsis thaliana*. *J. Evol. Biol.* **12**: 779–791.

Raeymaekers, J.A.M., Van Houdt, J.K.J., Larmuseau, M.H.D., Geldof, S. & Volckaert, F.A.M. 2007. Divergent selection as revealed by P(ST) and QTL-based F(ST) in three-spined stickleback (*Gasterosteus aculeatus*) populations along a coastal-inland gradient. *Mol. Ecol.* **16**: 891–905.

Ravinet, M., Prodöhl, P.A. & Harrod, C. 2013. Parallel and nonparallel ecological, morphological and genetic divergence in lake-stream stickleback from a single catchment. *J. Evol. Biol.* **26**: 186–204.

R Development Core Team 2012. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL www.r-project.org.

Reid, D.T. & Peichel, C.L. 2010. Perspectives on the genetic architecture of divergence in body shape in sticklebacks. *Integr. Comp. Biol.* **50**: 1057–1066.

Reimchen, T.E. 1992. Injuries on stickleback from attacks by a toothed predator (*Oncorhynchus*) and implications for the evolution of lateral plates. *Evolution* **46**: 1224–1230.

Reimchen, T.E. 1994. Predators and morphological evolution in threespine stickleback. In: *The Evolutionary Biology of the Threespine Stickleback* (M.A. Bell & S.A. Foster, eds), pp. 240–276. Oxford University Press, Oxford, UK.

Reimchen, T.E., Stinson, E.M. & Nelson, J.S. 1985. Multivariate differentiation of parapatric and allopatric populations of threespine stickleback in the Sangan river watershed, Queen-Charlotte-Islands. *Can. J. Zool.* **63**: 2944–2951.

Saemundsson, K. 1992. Geology of the Thingvallavatn area. *Oikos* **64**: 40–68.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**: 1766–1774.

Schluter, D. & Conte, G.L. 2009. Genetics and ecological speciation. *Proc. Natl. Acad. Sci. USA* **106**(Suppl. 1): 9955–9962.

Schluter, D. & Nagel, L.M. 1995. Parallel speciation by natural selection. *Am. Nat.* **146**: 292–301.

Schluter, D., Clifford, E.A., Nemethy, M. & McKinnon, J.S. 2004. Parallel evolution and inheritance of quantitative traits. *Am. Nat.* **163**: 809–822.

Smith, T.B. & Skúlason, S. 1996. Evolutionary significance of resource polymorphisms in fishes, amphibians, and birds. *Annu. Rev. Ecol. Evol. S* **27**: 111–133.

Spitze, K. 1993. Population structure in *Daphnia obtusa* – quantitative genetic and allozymic variation. *Genetics* **135**: 367–374.

Steppan, S., Phillips, P.C. & Houle, D. 2002. Comparative quantitative genetics: evolution of the G matrix. *Trends Ecol. Evol.* **17**: 320–327.

Svanbäck, R. & Schluter, D. 2012. Niche specialization influences adaptive phenotypic plasticity in the threespine stickleback. *Am. Nat.* **180**: 50–59.

Thibert-Plante, X. & Hendry, A.P. 2011. The consequences of phenotypic plasticity for ecological speciation. *J. Evol. Biol.* **24**: 326–342.

Thompson, C., Taylor, E.B. & McPhail, J.D. 1997. Parallel evolution of lake-stream pairs of threespine sticklebacks (*Gasterosteus*) inferred from mitochondrial DNA variation. *Evolution* **51**: 1955–1965.

Voje, K.L., Mazzarella, A.B., Hansen, T.F., Ostbye, K., Klepaker, T.O., Bass, A. *et al.* 2013. Adaptation and constraint in a stickleback radiation. *J. Evol. Biol.* **26**: 2396–2414.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *P. 6th Int. Cong. Genet.* **1**: 356–366.

Wund, M.A., Baker, J.A., Clancy, B., Golub, J.L. & Foster, S.A. 2008. A test of the "Flexible stem" model of evolution: ancestral plasticity, genetic accommodation, and morphological divergence in the threespine stickleback radiation. *Am. Nat.* **172**: 449–462.

Young, K.A., Snoeks, J. & Seehausen, O. 2009. Morphological diversity and the roles of contingency, chance and determinism in African cichlid radiations. *PLoS ONE* **4**: e4740.

Zeller, M., Lucek, K., Haesler, M.P. & Seehausen, O. 2012a. Little evidence for a selective advantage of armour-reduced threespined stickleback individuals in an invertebrate predation experiment. *Evol. Ecol.* **26**: 1293–1309.

Zeller, M., Lucek, K., Haesler, M.P., Seehausen, O. & Sivasundar, A. 2012b. Signals of predation-induced directional and disruptive selection in the threespine stickleback. *Evol. Ecol. Res.* **14**: 193–205.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1 Sampling sites with their respective geographic location, country of origin and the habitat with the number of individuals used for the phenotypic and the genetic data analysis as well as the coordinates where they were obtained.

Table S2 Technical summary on the microsatellites used in this study.

Figure S1 The evolution of the P matrix over time through gradual (a) or rapid changes (b).

Figure S2 Summary of all linear measurements used in this study that were either obtained on the left side (a), the gill arch (b), or from the ventral side of each individuals (c).