doi:10.1111/j.1420-9101.2009.01780.x

Evolutionary history shapes the association between developmental instability and population-level genetic variation in three-spined sticklebacks

S. VAN DONGEN,* L. LENS,† E. PAPE,‡,[1] F. A. M. VOLCKAERT‡ & J. A. M. RAEYMAEKERS‡

*Evolutionary Ecology Group, University of Antwerp, Antwerp, Belgium
†TEREC, Ghent University, Gent, Belgium
‡Laboratory of Animal Diversity and Systematics, Katholieke Universiteit Leuven, Leuven, Belgium

Keywords:

developmental instability;
evolution;
F_{ST};
fish;
fluctuating asymmetry;
genetic variation;
inbreeding;
P_{ST}.

Abstract

Developmental instability (DI) is the sensitivity of a developing trait to random noise and can be measured by degrees of directionally random asymmetry [fluctuating asymmetry (FA)]. FA has been shown to increase with loss of genetic variation and inbreeding as measures of genetic stress, but associations vary among studies. Directional selection and evolutionary change of traits have been hypothesized to increase the average levels of FA of these traits and to increase the association strength between FA and population-level genetic variation. We test these two hypotheses in three-spined stickleback (*Gasterosteus aculeatus* L.) populations that recently colonized the freshwater habitat. Some traits, like lateral bone plates, length of the pelvic spine, frontal gill rakers and eye size, evolved in response to selection regimes during colonization. Other traits, like distal gill rakers and number of pelvic fin rays, did not show such phenotypic shifts. Contrary to *a priori* predictions, average FA did not systematically increase in traits that were under presumed directional selection, and the increases observed in a few traits were likely to be attributable to other factors. However, traits under directional selection did show a weak but significantly stronger negative association between FA and selectively neutral genetic variation at the population level compared with the traits that did not show an evolutionary change during colonization. These results support our second prediction, providing evidence that selection history can shape associations between DI and population-level genetic variation at neutral markers, which potentially reflect genetic stress. We argue that this might explain at least some of the observed heterogeneities in the patterns of asymmetry.

Introduction

Developmental instability (DI) is the sensitivity of a developing trait to random perturbations of developmental processes (Klingenberg, 2003a). The phenotypic outcome of DI is that developmental trajectories deviate from their predetermined path (determined by genotype and environment). Consequently, observable trait values will deviate from their (unknown) expectations. Most commonly, the degree of asymmetry in bilaterally symmetric traits is used to estimate DI. In doing so, the development of both sides should reflect two independent replicates of the same developmental event. In the absence of any DI, a perfectly symmetric phenotype should emerge unless a genetic predisposition towards a particular direction (i.e. directional asymmetry) occurs. Higher variation in directionally *random* asymmetry is expected with increasing levels of DI (Van Dongen, 2006). This type of asymmetry, termed fluctuating asymmetry (FA), has received much attention in ecology

Correspondence: Dr Stefan Van Dongen, Evolutionary Ecology Group, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerp, Belgium. Tel.: +32 3 265 33 36; fax: +32 3 265 34 74;
e-mail: stefan.vandongen@ua.ac.be
[1] Current address: Marine Biology Section, Ghent University, Krijgslaan 281 S8, B-9000 Ghent, Belgium.

and evolution because it has been shown to increase with various (putative) measures of stress and to correlate negatively with fitness estimates like fecundity, attractiveness, competitive ability, parasite resistance and survival (see reviews in for example Lens *et al.*, 2002a; Polak, 2003; Van Dongen, 2006). These associations have prompted the suggestion that levels of asymmetry could serve as a general measure of individual and population (genetic) quality and health (e.g. Møller and Swaddle, 1997). However, patterns are highly heterogeneous and very little is known about the underlying causes of this heterogeneity.

Interest in FA and DI in evolutionary biology is not only limited to its potential as a measure of stress and fitness but also in relation to the evolutionary potential of traits. Because the stability of development may have the capacity to reduce morphological variation and conceal genetic variation (e.g. Breuker *et al.*, 2006), it has received recent attention in evolutionary biology. Indeed, the process of DI and stress effects upon it affect evolutionary processes and possibly speciation (Badyaev & Foresman, 2000). It is therefore important to understand if, and when, levels of DI increase because of stress. Attempts to explain the observed heterogeneity in DI–stress associations have been both conceptual and statistical (Lens *et al.*, 2002a; Van Dongen, 2006; Polak & Taylor, 2007). Some studies have indicated that the effects of different types of stress on the levels of DI may not be additive (e.g. Lens *et al.*, 2000). Furthermore, statistical analyses of FA often have relatively low power at both the population (e.g. Van Dongen, 1999) and individual level. Much of the observed variation in single trait asymmetry at the individual level reflects statistical noise only, and a high amount of between-individual variation in DI is indispensible before FA can reliably reflect DI (e.g. Gangestad & Thornhill, 1999; Houle, 1997; Van Dongen, 2006; Whitlock, 1996). A general explanation for the heterogeneity in patterns in FA might therefore be that many populations do not harbour sufficient variation in DI, a case in which the association between the observable asymmetry and underlying DI is extremely low (Polak & Taylor, 2007).

Traits that experienced a recent history of sustained directional selection may have been enriched in DI and its variation through the loss of canalizing modifiers (Polak & Taylor, 2007). Such conditions may increase the likelihood to show associations with measures of stress and/or fitness (Polak & Taylor, 2007). Thus, (recent) evolutionary changes may lead to higher DI and/or stronger associations between DI and indicators of stress. Support for this hypothesis has been mainly derived from the study of traits under sexual selection, where strong directional selection is acting on traits within one sex. Meta-analyses have shown that sexually selected traits often show high levels of asymmetry (e.g. Møller & Höglund, 1991) and possibly show stronger associations with measures of quality and fitness (Møller & Thornhill, 1998; but see Palmer, 1999 for a critical re-evaluation and a reply by Møller *et al.*, 1999). In gorilla, FA and environmental stress correlated positively in male canines only, a trait under sexual selection. This association appeared to be absent in female canines and both male and female second premolars (Manning & Chamberlain, 1994). Similar patterns occurred in grizzly bears (Badyaev, 1998). Furthermore, a link between DI and sexual selection was relatively strong in a population of *Drosophila bipectinata* undergoing recent evolutionary change in a sexually secondary trait (Polak & Taylor, 2007). Söderman *et al.* (2007) found higher asymmetries in traits with the highest level of sexual dimorphism in *Rana arvalis*. Yet, these traits did not show a stronger increase in FA with acidification than less dimorphic ones. Similarly, Bjorksten *et al.* (2000) did neither find increased FA under stress in a sexual selected trait in stalk-eyed flies, nor did sustained directional selection on *Drosophila melanogaster* wing shape result in higher wing asymmetry (Pelabon *et al.*, 2006). The above overview of the literature clearly calls for studies that compare levels of DI as well as associations between DI and measures of stress among traits with different selection histories. We predict higher mean levels of FA and stronger associations between FA and putative measures of stress in traits that experienced recent directional asymmetry compared with traits under stabilizing selection.

The three-spined stickleback (*Gasterosteus aculeatus* L.; Gasterosteidae; Teleostei) forms a highly suitable model system to compare levels of DI and the association between DI and measures of stress among traits that underwent different recent evolutionary changes. This species has a marine origin, but has repeatedly and independently invaded the freshwater environment, undergoing morphological changes in some traits often related to body armour (Bell and Foster, 1994; Leinonen *et al.*, 2006; Marchinko & Schluter, 2007; Cano *et al.*, 2008; and Raeymaekers *et al.*, 2007 for our area). In order to investigate the role of directional selection in determining levels of DI and in affecting associations between DI and other covariates potentially related to stress, we performed two analyses in the three-spined stickleback. First, we compared levels of DI in coastal (lowland) and freshwater (upland) populations. Secondly, we compared associations between FA and population-level neutral genetic variation between traits with different selection history. We studied three-spined stickleback from the Scheldt and Meuse basin and the coastal regions of Belgium and the Netherlands (Fig. 1). In this area, upland freshwater populations diverged postglacially (10 000–16 000 years BP) from a lowland ancestor, and observed genetic differentiation suggests a recent evolutionary event (Raeymaekers *et al.*, 2005). In the upland region, the populations living upstream typically experienced a loss of genetic variation due to low effective population sizes and migration barriers (Raeymaekers *et al.*, 2007, 2008). Reduced levels of

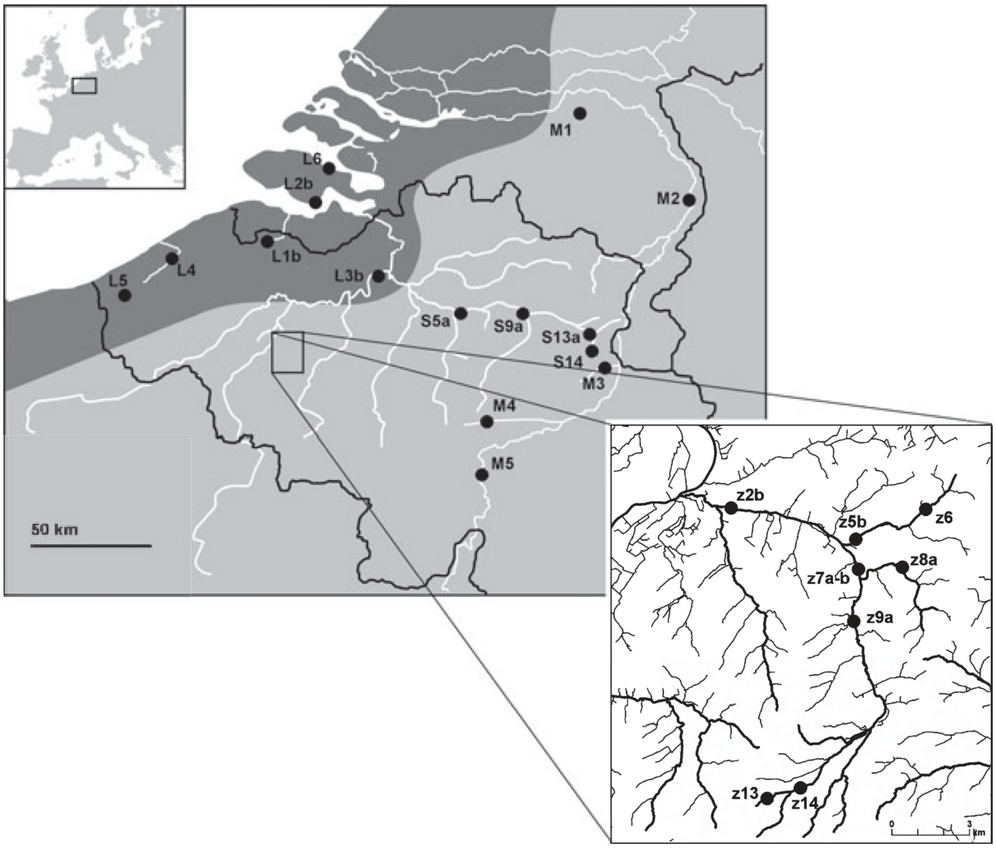

Fig. 1 Map of the study area, including sample locations and distribution of lowland (dark grey) and upland (light grey) populations.

genetic variation, as a measure of increased levels of inbreeding, may increase levels of DI and FA as shown in several studies. More generally and especially in fishes, many studies have shown a positive association between heterozygosity and measures of fitness. However, associations have been weak on average and heterogeneous among study systems, and the underlying mechanism is still under discussion (David, 1998; Thelen & Allendorf, 2001). Although there is still much debate about the biological relevance of measures of genetic variation at presumed selectively neutral loci (e.g. Amos & Balmford, 2001), a review has shown a correlation coefficient of 0.43 between population-level fitness and genetic variation (Reed & Frankham, 2003). Furthermore, FA and DI may act as an early warning system, where levels of DI would increase with stress before fitness is strongly affected (Lens *et al.*, 2002b). In addition, fitness effects of the loss of genetic variation may only become apparent in combination with other forms of stress (Lens *et al.*, 2000; Amos & Balmford, 2001), suggesting that fitness should not necessarily decrease with loss of genetic variation in any situation. As genetic variation at neutral markers may also reflect the evolutionary potential of quantitative traits, it is a biologically relevant characteristic of many natural populations. We therefore argue that associations between FA and neutral genetic variation may be relevant even when the amount of loss in genetic variation under the prevailing environmental conditions may not have resulted in a loss in fitness.

When comparing neutral and phenotypic differentiation between lowland and upland regions, some traits have experienced a recent evolutionary change. The upland ecotype is characterised by the typical loss of body armour, shorter pelvic spines and smaller eyes (Raeymaekers *et al.*, 2007). On the other hand, changes in other traits like gill rakers and number of rays in the pectoral fin were minor (Raeymaekers *et al.*, 2007). We hypothesize that for traits showing high differentiation between up- and lowland, FA and associations between FA and genetic variation will be higher in the upland populations.

In this study, we measured asymmetry in a range of traits in a total of 24 lowland and upland populations. First, we compare morphological variation between both sets of populations and determine the degree of phenotypic differentiation (for this particular set of populations) as a proxy for the strength of directional selection. Secondly, we explicitly compare levels of asymmetry between these two sets of populations in order to test if traits under directional selection have elevated levels of DI in the upland populations. Thirdly, we compare associations between FA and genetic variation in the upland populations between traits under varying degrees of directional selection and predict stronger associations for traits that diverged from the ancestral lowland state.

Materials and methods

Study population and sampling methods

Three-spined sticklebacks were sampled at 24 sites in Belgium and the Netherlands in the spring of 2004 (Fig. 1). Six lowland (coastal) samples (coded L) were caught in the polders bordering the Scheldt estuary (L1b, L2b, L3a and L6) and the North Sea (L4 and L5). Five samples originated from the Meuse basin (coded M) and 13 samples originated from the Scheldt basin (coded S or Z). About 50 adult (>30 mm) individuals per site were caught with a dip net or by electrofishing, and flash frozen on dry ice. All individuals were weighed (±0.01 g) and measured for standard length (SL) and total length (TL) (±0.1 cm). We took fin clips and stored them in 100% ethanol for DNA analysis.

Measurements of asymmetry

Pectoral fin rays (PFR) were counted bilaterally on thawed fish (no data available for the Scheldt populations coded with Z). Next, we fixed fish bodies in a 4% formalin solution. Two months later, we rinsed the sticklebacks with water for 72 h, bleached for 4 h (1% KOH bleach solution), buffered for 24 h (30% borax buffer, pH 7.0) and stained with alizarin red S to facilitate plate counts and plate morph registration (Taylor & Van Dyke, 1985). After staining, we determined the diameter of both eyes (coded EYE) with a digital calliper (±0.01 mm). We measured eye diameter along a vertical straight line on the bony structure of the eye. We removed both pelvic spines (PS) from the fish and stored separately in 100% glycerol. Later on, they were photographed with a Matrox Meteor camera connected with the software program Das (Cam2Disk 2.2) and digitally measured in Image-Pro Plus 5.0 (Media Cybernetics, Silver Spring, MD, USA). Spine length was determined as the straight line connecting the lowest point of the articulation to the spine tip (Mazzi *et al.*, 2002). The number and position of the lateral plates on the right and left side were determined. We considered anterior non-structural plates (P1P3), central structural plates (P4P8) and posterior non-structural plates (P9PN) as distinct units (Bergstrom & Reimchen, 2003). We dissected gills after removal of the gill cover. Large (LGR) and small (SGR) gill rakers were counted on the frontal (LGR-1; SGR-1) and distal (LGR-2; SGR-2) part of the first branchial arch (Peichel *et al.*, 2001). Final values for all traits were obtained after averaging over two independent measurements, whenever available. Because the length of the pelvic spine and the eye diameter depend on body size, they were divided by the standard length. No such correction was needed for the meristic traits.

We analysed degrees of measurement error and directional asymmetry (systematic asymmetry in a particular direction) using a mixed regression model and descriptive statistics as outlined in Van Dongen (2006). To investigate if different traits represent independent developmental events, we studied correlations in the signed asymmetries (left–right trait value) to detect developmental integration, and thus traits that share developmental errors. In the case of significant developmental integration, such traits share developmental events.

Genetic analyses

We extracted genomic DNA from fin clips using a silica-based purification method (Elphinstone *et al.* 2003) and assessed allelic variation at fifteen microsatellite loci following procedures in Raeymaekers *et al.* (2007), using the Qiagen® Multiplex PCR Kit (Qiagen, Venlo, the Netherlands). We visualized PCR products on an ABI3130 Avant Genetic analyzer (Applied Biosystems, Foster City, CA, USA). Allele sizes were determined by means of an internal GeneScan 500-LIZ size standard and genotypes were scored using GENEMAPPER v. 3.7 (Applied Biosystems). We checked genotypes for scoring errors that might be attributable to stutter-products, large allele dropout or to the presence of null-alleles, using the software MICRO-CHECKER v. 2.2.3 (Van Oosterhout *et al.*, 2004). Multilocus heterozygosity (MLH) and allelic richness (AR) were calculated based on seven putatively neutral microsatellites (Gac1097, Gac4170, Gac5196, Stn174, Stn23, Stn37, and Stn84) using GENETIX v. 4.04 (Belkhir *et al.*, 2002).

Phenotypic differentiation

We compared the extent of divergence for meristic and metric traits (standardized for size), quantified as P_{ST}, with neutral genetic divergence (F_{ST}). If selection regimes differ among locations, P_{ST} will be larger than expected based on of neutral loci (Leinonen *et al.*, 2008). We calculated P_{ST} and F_{ST} among each population pair following procedures in Raeymaekers *et al.* (2007).

Statistical analyses of patterns in asymmetry

We performed two sets of analyses to test the two hypotheses introduced above. First, we test whether levels of FA were higher in upland compared to lowland populations, especially in traits under directional selection. Secondly, we compare associations between FA and genetic variation at the population level in the upland populations between traits that showed phenotypic divergence under directional selection and those that did not change.

Comparing average asymmetry between lowland and upland populations for different traits

To test our first prediction, we compared average asymmetries across all populations in the lowland and upland regions between traits that did or did not experience recent evolutionary changes. We constructed a linear model with average population FA as dependent variable and a factor CHANGE and the factor upland versus lowland (UP/LOW) and their interaction as fixed. We added trait (nested within the factor CHANGE) and the trait-by-UP/LOW interaction as well as population as random effects. A significant interaction between CHANGE and UP/LOW supports our prediction of higher FA in the upland populations for traits under recent directional selection. The interaction between trait and UP/LOW was also explicitly tested to explore if differences between lowland and upland vary across traits.

To better understand these possible trait-specific effects, a further exploration was performed. The proportion of trait-specific asymmetries (number of asymmetric elements over the mean number elements across sides) in plate numbers (P1P3, P4P8 and P9PN), gill rakers (LGR-1, LGR-2, SGR-1, and SGR-2) and number of fin rays in the pectoral fin (PFR) at the individual level were analysed using generalized linear mixed models with logit link and binomial error structure. The factors upland vs. lowland (UP/LOW), trait and their two-way interaction were tested as fixed effects. We decided to analyse the three categories of plate numbers and the four categories of gill raker numbers in two separate analyses. This approach will allow testing for a two-way interaction between trait and UP/LOW and to compare increases in asymmetry between traits that may differ in functional importance (structural vs. non-structural lateral plates) and/or phenotypic divergence (frontal vs. distal gill rakers). For PFR a separate analysis was performed. We added population, nested within the category upland vs. lowland, as random effect to take into account the dependency of the data within populations. For eye diameter (EYE) and length of the pelvic spine (PS), the relative asymmetry (unsigned asymmetry divided by trait value) was analysed separately in a similar model (UP/LOW and trait as fixed effects, population as random effect), assuming approximate normality. To rule out that differences in asymmetry between the lowland and upland populations could be confounded by the loss of genetic variation in the upland populations or by differences in trait size between upland and lowland populations, population-level multi-locus heterozygosity (MLH) and uncorrected individual trait size were added to the above models as covariates.

Comparing associations between FA and genetic variation between traits in upland populations

Average degrees of asymmetry (either a proportion or relative asymmetry for the meristic and metric traits, respectively) for each trait were determined for all 18 upland populations. As only 27 individuals developed plates caudally of plate 8, we omitted P9PN from further analyses. In order to combine metric and meristic traits in the statistical analyses, we standardised average population FA by subtracting the grand mean and dividing by the overall standard deviation for each trait within a single analysis. To be able to compare the degree of association between asymmetry and levels of genetic variation between the groups of traits that did and did not experience a recent evolutionary change during the colonization of the freshwater environment, we used a mixed ANCOVA model. Average standardized asymmetry was used as dependent variable and average population-level genetic variation (MLH) or allelic richness (AR) as covariates. We also added a factor to indicate which traits experienced an evolutionary change or not (CHANGE) as well as its interaction with MLH (or AR) as fixed effect. To take into account that patterns may vary across traits, we added trait (nested within CHANGE) and its interaction with MLH or AR as random effects. Denominator degrees of freedom and standard errors of test of fixed effects were approximated using Kenward–Rogers method. In this way, the tests of fixed effects explicitly take between-trait variation in the associations into account.

Results

Suitability of asymmetry measurements as measures of developmental instability

Measurement error was very low for all traits except for eye diameter where asymmetry and measurement error were almost equally large (Table 1). The repeatability of nearly 50% indicates that about half of the variation in the observed asymmetry of EYE is due to real variation in FA. As we mainly focus on population-level analyses, where ME is averaged across all individuals, we retained this trait for further analyses. Furthermore, it is unlikely that the higher ME would generate associations with genetic variation.

In two traits (LGR-2 and EYE), tests for directional asymmetry were statistically significant, and the test remained significant for LGR-2 after Bonferroni correction. However, the degree of directional asymmetry was low relative to the total variation in asymmetry (i.e. low effect sizes in Table 1) such that a correction for directional asymmetry will hardly affect results. This lack of directional asymmetry (in spite of high sample sizes) in most traits and a very small level for LGR-2 (i.e. 0.15 SD units, Table 1), also shows that there is little evidence for

Table 1 Overview of descriptive statistics of asymmetry measurements in three-spined stickleback. Levels of asymmetry (FA) relative to measurement error (ME) were obtained from mixed regression model analyses and formed the basis to calculate repeatability [r^2 = FA/(FA + ME)]. A significance test for directional asymmetry (DA) and its relative importance [effect size = DA/SD(signed asymmetry)] were also provided. Statistically significant results ($P < 0.05$) are indicated in bold.

Trait	FA	ME	r^2 (%)	DA (left–right)	Effect size of DA
LGR-1	0.0042	0.00003	99	$t_{1185} = -1.47$, $P = 0.14$	−0.03
LGR-2	0.022	0.00002	99	**$t_{1138} = 5.09$, $P < 0.0001$**	0.15
SGR-1	0.0069	0.00004	99	$t_{1182} = 1.05$, $P = 0.29$	0.03
SGR-2	0.011	0	100	$t_{1155} = 1.68$, $P = 0.09$	0.05
P1P3	0.050	0	100	$t_{1123} = 1.42$, $P = 0.16$	0.04
P4P8	0.0037	0	100	$t_{1195} = 0.03$, $P = 0.98$	0.00
P9PN	0.25	0.006	99	$t_{251} = 0.91$, $P = 0.37$	0.05
PFR	0.001	0	100	$t_{742} = -0.65$, $P = 0.52$	−0.05
PS	0.001	0.00012	90	$t_{977} = -0.65$, $P = 0.52$	−0.02
EYE	0.00043	0.00049	47	**$t_{1193} = 2.78$, $P = 0.006$**	0.08

a genetic basis of asymmetry in a particular direction in any trait. Below, we report analyses for data uncorrected for directional asymmetry only, but results for corrected data were almost identical.

Between-trait correlations in signed and unsigned asymmetries (absolute value of signed asymmetry) were very low for most trait combinations (Table 2). The correlation in the signed asymmetry between SGR-1 and SGR-2 was statistically significant and negative, indicating that there is an interaction between those two traits within sides during development. For correlations in the unsigned asymmetries, P9PN appeared to correlate positively with P1P3, P4P8 and SGR-2, suggesting that similar external factors affected levels of DI in these traits. However, because there were no correlations in the signed asymmetries, we regarded the different traits as independent developmental events (Klingenberg, 2003b). Hence, we treated them as separate units in our analyses.

One argument against the separation of the number of plates in different traits could be that directional selection acted on the major gene underlying plate development (*Eda*, Colosimo *et al.*, 2005), and thus that the evolutionary change did not reflect an independent evolutionary event. In spite of the fact that the developmental pathways of the plates appear independent, we explore the effect of treating them as separate traits on the association between DI and genetic variation, by deleting one of them separately and testing the interaction between CHANGE and MLH.

Levels of genetic variation

Average MLH equalled 0.83 (SD = 0.03, range: 0.78–0.87) in the lowland populations and 0.65 (SD = 0.10, range: 0.46–0.80) in the upland populations, a difference that was highly significant ($t_{22} = 4.3$, $P = 0.0003$). Average AR equalled 12.4 (SD = 0.93, range: 11.0–13.5) in the lowland populations and 7.37 (SD = 1.74, range: 3.83–10.1) in the upland populations, a difference that was highly significant as well ($t_{22} = 6.67$, $P < 0.0001$). The upland populations show a large range of genetic variation, which is the result of lower effective population sizes and the presence of barriers in upstream populations (Raeymaekers *et al.*, 2008). As MLH and AR were highly correlated ($r = 0.94$, $P < 0.0001$), the observed patterns will likely be very similar.

Phenotypic differentiation

Mean trait values and between-population variations are provided in Table 3. As shown before in a smaller dataset (Raeymaekers *et al.*, 2007), upland populations appeared to be characterized by lower plate number (P1P3, P4P8

Table 2 Correlations in the signed (above diagonal) and unsigned (below diagonal) asymmetry between traits of three-spined stickleback. Significant correlations after correction for multiple testing are indicated in bold.

	LGR-1	LGR-2	SGR-1	SGR-2	P1P3	P4P8	P9PN	PFR	PS	EYE
LGR-1	–	−0.03	0.02	−0.02	0.01	−0.02	−0.02	−0.03	−0.01	0.01
LGR-2	0.06	–	0.03	0.01	0.06	0.03	−0.01	0.02	0.04	0.02
SGR-1	−0.03	−0.03	–	**−0.22**	−0.02	0.05	−0.05	−0.03	0.01	−0.01
SGR-2	0.06	0.03	0.02	–	0.00	−0.01	−0.12	−0.05	0.04	0.03
P1P3	0.04	−0.06	−0.03	−0.01	–	−0.01	0.12	−0.09	−0.02	0.00
P4P8	−0.04	0.03	0.03	0.01	0.04	–	−0.07	0.05	0.04	0.05
P9PN	−0.02	−0.00	0.00	**0.20**	**0.14**	**0.36**	–	−0.03	0.03	−0.02
PFR	−0.02	−0.05	−0.02	0.03	0.01	−0.01	0.03	–	0.00	0.06
PS	0.00	−0.03	−0.00	−0.02	0.02	0.03	0.05	−0.02	–	0.06
EYE	0.04	0.04	−0.03	0.00	0.01	0.05	0.03	−0.02	0.06	–

Table 3 Means and standard deviations among populations and range of the different traits in the lowland and upland populations of three-spined stickleback. For each trait the degree of phenotypic differentiation (P_{ST} and its 95% confidence interval) is also provided. Traits in which we considered the phenotypic differentiation to be higher compared to the neutral expectations (F_{ST} = 0.11, 95% CI = 0.10–0.12) are indicated in bold.

	Lowland		Upland		P_{ST}	
Traits	Mean (SD)	Range	Mean (SD)	range	Mean	95% CI
LGR-1	**14.0 (0.26)**	**13.6–14.3**	**13.3 (0.31)**	**12.8–13.9**	**0.21**	**0.18–0.24**
LGR-2	6.18 (0.39)	5.68–6.75	6.04 (0.24)	5.55–6.44	0.07	0.05–0.08
SGR-1	**11.9 (0.22)**	**11.5–12.1**	**11.3 (0.28)**	**10.9–12.0**	**0.16**	**0.14–0.19**
SGR-2	3.74 (0.21)	3.55–4.10	3.88 (0.35)	3.38–4.58	0.11	0.09–0.14
P1P3	**2.62 (0.21)**	**2.34–2.90**	**1.61 (0.29)**	**1.24–2.40**	**0.47**	**0.43–0.51**
P4P8	**4.93 (0.05)**	**4.84–4.98**	**4.34 (0.21)**	**4.12–4.78**	**0.47**	**0.44–0.51**
P9PN	**10.7 (1.73)**	**8.14–12.8**	**1.52 (5.23)**	**0–10.1**	**0.43**	**0.40–0.46**
PS	**1.48 (0.08)**	**1.40–1.59**	**1.30 (0.17)**	**1.08–1.71**	**0.46**	**0.42–0.51**
EYE	**1.07 (0.09)**	**0.97–1.17**	**1.00 (0.10)**	**0.86–1.19**	**0.35**	**0.30–0.40**
PFR	10.1 (0.07)	10.0–10.2	10.1 (0.04)	9.96–10.2	0.02	0.01–0.03

and P9PN), shorter relative pelvic spines (PS), smaller relative eye diameter (EYE) and slightly fewer large and small frontal gill rakers (LGR-1 and SGR-1) (Table 3; Fig. 2).

Differences in levels of asymmetry between lowland and upland populations

Average levels of asymmetry did not differ between lowland and upland populations (no significant UP/LOW effect: $F_{1,7}$ = 0.64, P = 0.45), and showed no differences between traits that did or did not show a recent evolutionary change (no significant CHANGE effect: $F_{1,7}$ = 0.54, P = 0.48). Furthermore, there was no interaction between UP/LOW and CHANGE ($F_{1,7}$ = 1.35, P = 0.28). However, differences between lowland and upland populations appeared to vary across traits (χ^2_1 = 35.1, P < 0.0001). This was explored in the next set of analyses.

Changes in the average degree of asymmetry between lowland and upland populations differed between traits (Fig. 3). For gill rakers, changes were on average small but not similar for all traits (statistically significant interaction Table 4). This interaction appears to be the result of the slight but significant increase of asymmetry in the small frontal gill rakers (SGR-1) in upland populations (Fig. 3). Asymmetry in lateral plates was higher in the upland populations, and this difference was comparable among traits (no significant interaction; Table 4). Asymmetry in the three other traits did not differ between low- and upland populations (Table 4).

Associations between asymmetry and levels of genetic variation

Associations between population-level asymmetry and multi-locus heterozygosity differed significantly between traits that experienced either a recent evolutionary change or not (CHANGE: $F_{1,149}$ = 0.02, P = 0.88; MLH: $F_{1,149}$ = 0.18, P = 0.67; CHANGE*MLH: $F_{1,149}$ = 7.20, P = 0.008). This interaction between CHANGE and MLH remained significant when either P1P3 ($F_{1,131}$ = 6.48, P = 0.01) or P4P8 ($F_{1,131}$ = 8.41, P = 0.004) were removed from the analyses. Thus, patterns were not an artefact of treating the different plates as independent developmental events. Furthermore, this interaction was also significant after removing EYE (having moderate degrees of measurement error, $F_{1,131}$ = 6.92, P = 0.01) or PS (potentially showing directional asymmetry in fish with reduced spine lengths as a result of the action of the *PitX1* gene (Shapiro *et al.*, 2004, $F_{1,131}$ = 5.34, P = 0.02). Thus, the stronger associations between FA and MLH in traits that experienced a recent evolutionary change was not an artefact of including one

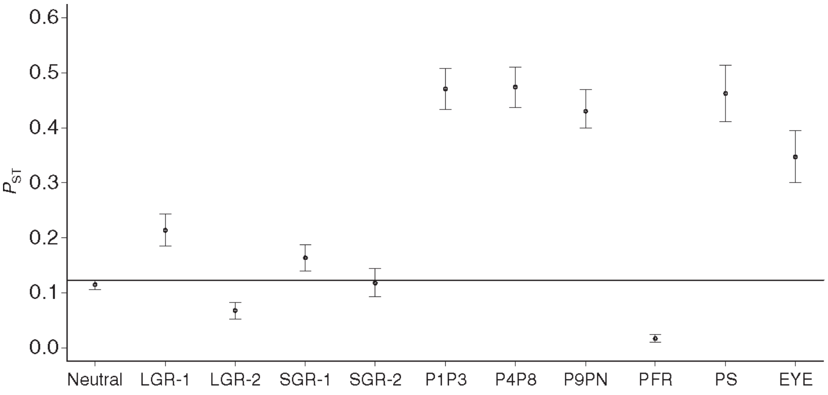

Fig. 2 Comparison of pair-wise neutral genetic differentiation of three-spined stickleback between lowland and upland populations and pair-wise differentiation on the basis of phenotypic traits (P_{ST}). Tails denote 95% confidence intervals and the horizontal line delineates the upper 95% confidence interval for F_{ST} as estimated from seven microsatellite loci.

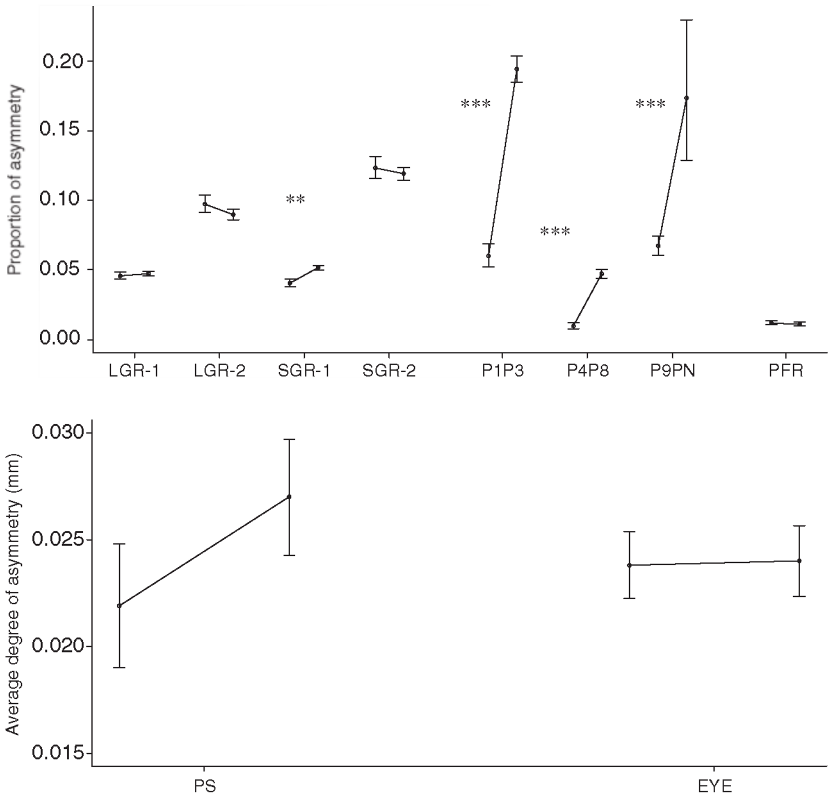

Fig. 3 Mean (±SE) levels of asymmetry of three-spined stickleback in lowland (left) and upland (right) populations for eight meristic traits (upper panel) and two continuous traits (bottom panel). For each trait, means in lowland and upland were connected by line pieces to illustrate the differences. Significant differences were indicated as $^{**}P < 0.01$ and $^{***}P < 0.001$.

Table 4 Overview of tests of mixed models comparing levels of asymmetry between upland and lowland populations (UP/LOW). Gill rakers (LGR-1, LGR-2, SGR-1, SGR-2) and lateral plates (P1P3, P4P8, P9PN) were combined in single analyses.

Factor in model	Gill rakers	Lateral plates	PFR	EYE/PS
UP/LOW	**$\chi^2_1 = 4.02^{*}$**	**$\chi^2_1 = 45.6^{***}$**	$\chi^2_1 = 0.41$	$F_{1,26} = 0.06$
Trait	**$\chi^2_3 = 104^{***}$**	**$\chi^2_2 = 193^{***}$**	–	**$F_{1,1795} = 16.3^{***}$**
UP/LOW*trait	**$\chi^2_3 = 12.8^{**}$**	$\chi^2_2 = 5.90$	–	$F_{1,2123} = 0.34$
MLH	$\chi^2_1 = 0.35$	$\chi^2_1 = 1.29$	$\chi^2_1 = 0.31$	$F_{1,26} = 1.77$
Trait value	**$\chi^2_1 = 37.8^{***}$**	**$\chi^2_1 = 121^{***}$**	**$\chi^2_1 = 52.5^{***}$**	**$F_{1,961} = 24.3^{***}$**

Significant effects are indicated in bold ($^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$).

particular trait in the analysis. Furthermore, the difference in degree of association between traits that did or did not experience a recent evolutionary change appeared to be consistent among traits as the variance component of the random effects interaction between trait and multilocus heterozygosity equalled zero and thus was not statistically significant. Between-trait variation in FA also equalled zero because FA levels were standardized.

Estimated slopes of the regression between average asymmetry and multi-locus heterozygosity equalled −0.28 (SE = 0.09, $t_{149} = -2.95$, $P = 0.004$) and 0.20 (SE = 0.15, $t_{149} = 1.33$, $P = 0.19$) for traits that experienced either a recent phenotypic change or not respectively. The negative association between FA and MLH remained significant after excluding P1P3 ($t_{131} = -2.59$, $P = 0.01$), P4P8 ($t_{131} = -5.99$, $P = 0.003$), EYE ($t_{131} = -4.61$, $P = 0.01$) and PS ($t_{131} = -2.18$, $P = 0.04$). Thus, as predicted, associations were stronger for traits that experienced a recent evolutionary change (Fig. 4). This interpretation was also supported by the negative association between the degree of phenotypic differentiation between the lowland and upland populations (P_{ST}) and the correlation between asymmetry and multilocus heterozygosity (Fig. 5). Of all single-trait correlations between multilocus heterozygosity and asymmetry, only PS was statistically significant (PFR: $r = 0.59$, $P = 0.10$; LGR-1: $r = -0.41$, $P = 0.09$; LGR-2: $r = 0.28$, $P = 0.26$; SGR-1: $r = -0.11$, $P = 0.66$; SGR-2: $r = 0.04$, $P = 0.86$; PS: $r = -0.53$, $P = 0.02$; EYE: $r = -0.26$, $P = 0.30$; P1P3: $r = -0.33$, $P = 0.18$; P4P8: $r = -0.04$, $P = 0.89$). The association between multilocus heterozygosity and asym-

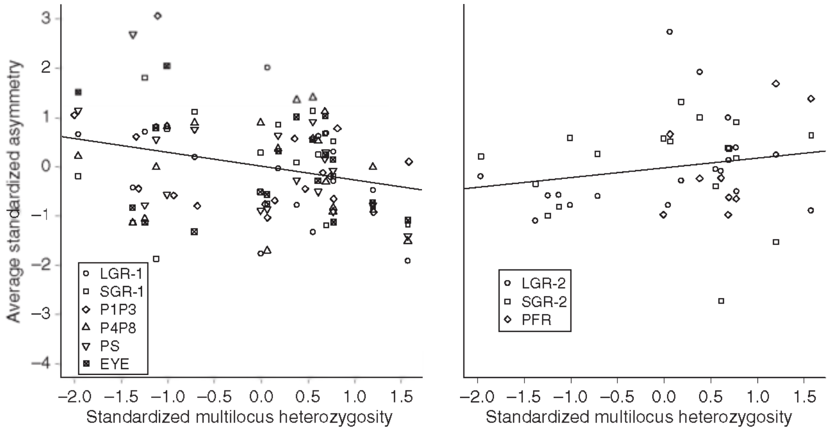

Fig. 4 Associations between population-level multi-locus heterozygosity and standardized asymmetry in 18 upland three-spined stickleback populations for traits that did (left) and did not (right) experience a recent evolutionary change during the postglacial colonization of the freshwater habitat. Estimated regression lines from the mixed ANCOVA model (see text for details) were also added.

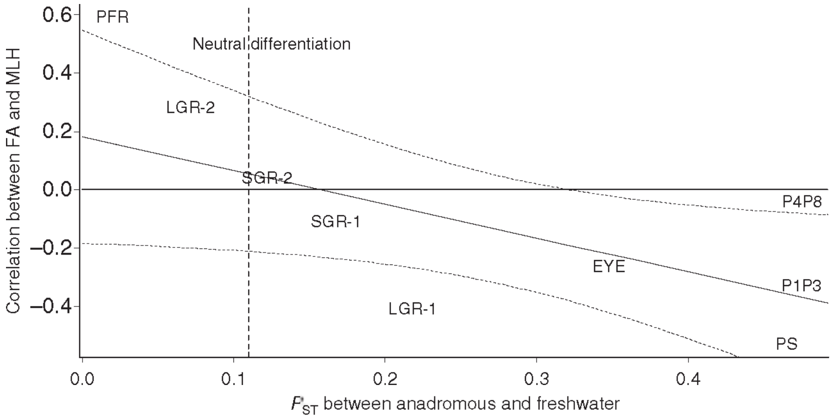

Fig. 5 Association between the degree of differentiation between lowland and upland populations (P_{ST}) of three-spined stickleback as a measure of the strength of directional selection on the one hand and the degree of correlation between asymmetry and microsatellite multilocus heterozygosity in the upland populations on the other hand. The linear regression line (correlation coefficient = 0.18–1.16 × P_{ST}; $F_{1,7} = 5.9$, $P = 0.04$) and 95% confidence bands are provided. The vertical dashed line indicates the observed amount of genetically neutral differentiation between lowland and upland populations (F_{ST}).

metry in PS was no longer significant after Bonferroni correction. The positive trend in PFR, albeit not statistically significant, is remarkable. However, PFR was only measured in a subset of populations, and removing this trait from the analysis did not alter the significant interaction between CHANGE and MLH ($F_{1,140} = 5.43$, $P = 0.02$).

As expected based on the high correlation between MLH and AR, patterns were very similar for AR. Indeed the strength of association between asymmetry and AR differed significantly between traits that experienced a recent evolutionary change or not (CHANGE: $F_{1,149} = 0.05$, $P = 0.82$; AR: $F_{1,8} = 0.01$, $P = 0.93$; CHANGE × MLH: $F_{1,8} = 8.23$, $P = 0.01$), and estimated slopes of the regression between average asymmetry and AR equalled −0.14 (SE = 0.05, $t_{149} = -2.64$, $P = 0.01$) and 0.15 (SE = 0.09, $t_{149} = 1.74$, $P = 0.08$) for traits that experienced a recent phenotypic change or not, respectively. As for MLH, of all single-trait correlations between AR and asymmetry, only PS was statistically significant (PFR: $r = 0.65$, $P = 0.06$; LGR-1: $r = -0.31$, $P = 0.20$; LGR-2: $r = 0.40$, $P = 0.09$; SGR-1: $r = 0.00$, $P = 0.99$; SGR-2: $r = 0.05$, $P = 0.84$; PS: $r = -0.64$, $P = 0.004$; EYE: $r = -0.15$, $P = 0.54$; P1P3: $r = -0.44$, $P = 0.07$; P4P8: $r = 0.05$, $P = 0.83$). After Bonferroni correction, the single significant correlation in PS remained statistically significant ($P = 0.004 \times 9 = 0.036$). As for the analysis of MLH, excluding PFR did not affect the statistical significance of the interaction between CHANGE and AR ($F_{1,140} = 4.86$, $P = 0.04$).

Discussion

General conclusions

We predicted two patterns in developmental instability for traits with a recent history of directional selection. First, levels of FA should increase, and second, FA in traits that experienced recent evolutionary change should be more sensitive to stress. We found evidence that some but not all traits under presumed directional selection during colonization of the freshwater habitat suffered from increased levels of FA. We also detected a weak but statistically significant negative association between FA and genetic variation in microsatellite markers, but only in traits that experienced this recent evolutionary change. Generally, the results presented here fit the framework proposed by Polak & Taylor (2007) that associations between FA and measures of stress and fitness should most likely be found in populations and traits showing recent evolutionary changes. The reason why for example Bjorksten *et al.* (2000) did not find an association between stress and FA in a sexually selected trait in stalk eyed flies, and Söderman *et al.* (2007) did not detect any differences in FA-stress association among traits that differ in degree of sexual dimorphism might be that the directional selection in these studies might be of more ancient origin. Evolutionary changes in the three-spined stickleback here, and in many other areas, are relatively recent and these changes can be extremely rapid. This recent and fast evolution may be a prerequisite for the increase in the sensitivity of a developing system to stress factors.

Increase of FA in freshwater populations in traits under directional selection

The expected increase in asymmetry under directional selection was not statistically significant overall and only observed for lateral plate numbers (P1P3, P4P8, P9PN) and to a smaller extent for the frontal short gill rakers (SGR-1). In contrast, frontal long gill rakers (LGR-1), length of the pelvic spine (PS) and eye diameter (EYE) did not show this expected increase although P_{ST} values were above neutral levels of genetic differentiation (F_{ST}). This heterogeneity does not seem to be related to the degree of phenotypic divergence (P_{ST}) as the pelvic spine had a relatively high P_{ST} estimate comparable with the bone plate numbers, and estimates were somewhat lower for both SGR-1 and LGR-1. Furthermore, the increased FA in these four traits did not appear to be a direct effect of loss in genetic variation alone as suggested as a possible mechanism by Pelabon *et al.* (2006), because differences in asymmetry between upland and lowland populations remained significant after correction for genetic variation at neutral markers. Trait functionality and fitness consequences of asymmetry also seem to fail to explain this heterogeneity as the higher symmetry in structural plates P4 to P8, which we observe here as well, have been interpreted as evidence for a higher functional importance for structural integrity of predatory defences (Bergstrom & Reimchen, 2000). Yet, as we find a similar increase in asymmetry for both the structural and the non-structural plates, a role of functionality of trait symmetry against predation in determining the changes in asymmetry between upland and lowland populations does not seem to be supported by our data. It could be argued that the lower asymmetry in the lowland populations is the result of a higher predation rate in more asymmetric individuals (Bergstrom & Reimchen, 2003). However, the fact that both structural and non-structural plates showed a comparable increase in asymmetry in the freshwater populations (no significant interaction between the factor UP/LOW and trait, Table 4) indeed does not support a direct role of a decrease in predation risk in the observed increase in FA. An alternative mechanism for the increased asymmetry in bone plates in upland populations may be the limitation of calcium in the freshwater environment. This hypothesis was further supported by the positive association between size-corrected FA and the number of plates for P4P8 ($r = 0.10$, $n = 896$, $P = 0.002$), but not the other plates (P1P3: $r = -0.18$, $n = 825$, $P < 0.0001$; P9PN: $r = -0.89$, $n = 30$, $P < 0.0001$). Producing more plates appeared to result in an increased probability of asymmetric expression of the structural lateral plates. For the non-structural plates, it is possible that low plated fish with fewer armour plates simply have more scope for asymmetry because their sides are not fully covered with plates. Thus, the generally higher FA in plates in the upland populations is more likely attributable to factor(s) unrelated to the recent directional selection but it remains difficult to identify the exact mechanism. In conclusion, our data do not support a general increase in DI for traits under recent directional selection. This appears to confirm earlier experimental results in *Drosophila* (Pelabon *et al.*, 2006) and theoretical arguments against the more traditional view that directional selection would generally act as a decanalizing force (Hansen *et al.*, 2006).

Associations between FA and genetic variation in traits under directional selection

The observed correlations between FA and genetic variation, a putative marker of inbreeding depression, suggest that FA has some predictive power to detect population-level stress in three-spined stickleback, and that selection history may shape the DI–stress associations. This interpretation critically depends on the assumption that our measures of genetic variation reflect levels of inbreeding and the expected associated loss in fitness. As many other authors have failed to find a correlation between FA and genetic variation or inbreeding at the individual level while they did at the population level (e.g. Hartl *et al.*, 1995; Zachos *et al.*, 2007), sampling effect might play a role. For example, sampling of the genome may require at least 20 loci in samples of more than 100 individuals to attain sufficient statistical power (Slate & Pemberton, 2002). Such conditions have come within reach with the latest developments in environmental genomics (e.g. Ellegren, 2008) and offer interesting perspectives for future studies of patterns at the individual level. Because at this point we scored only seven neutral markers, we restricted our analyses to population-level measures of FA and genetic variation. Nevertheless, it allowed to explicitly compare patterns between traits under directional selection or not and revealed significant differences between these two categories. We do not have any direct fitness estimates to confirm that the loss of genetic variation would indeed reflect stress and would result in inbreeding depression. Fitness is difficult to assess in the wild and the effects of reduced genetic variation may only become apparent in co-occurrence with other forms of stress (Lens *et al.*, 2000; Amos & Balmford, 2001). The fact that we observed a loss of about 50% in genetic variation can be regarded as a considerable reduction that likely has caused inbreeding depression in some populations. Indeed, a one-generation inbreeding experiment showed that brother-sister mating resulted in reduced fertilization success and survival in three-spined stickleback (Frommen *et al.*, 2008). More generally, several studies have found associations between heterozygosity at putative selectively neutral markers and fitness (Reed & Frankham, 2003). Nevertheless, results are heterogeneous and it is not clear what the underlying mechanism is (David, 1998; Thelen & Allendorf, 2001). Even if fitness effects of the observed loss in genetic variation would be limited, associations with DI and FA could be predicted because FA has been proposed as an early warning mechanism which might already increase before severe fitness loss is observed (e.g. Lens *et al.*, 2002b). In the most unlikely situation, a reduction to zero genetic variation in our populations might still not result in a fitness reduction. In that situation, we only observed FA-MLH associations, which cannot be directly interpreted in relation to fitness. Nevertheless, increased levels of FA are also of interest in evolutionary processes because it can generate morphological variation, reveal the presence of otherwise hidden genetic variation and affect morphological changes during speciation (e.g. Badyaev & Foresman, 2000; Breuker *et al.*, 2006). Thus, increased levels of FA in genetically eroded populations may affect the evolutionary potential of these traits.

It is important for the interpretation of the results that not all traits show the same strength of association between asymmetry and population-level genetic variation at neutral markers. More specifically, traits having experienced a recent evolutionary change appeared more sensitive to the presumed stress effects. This pattern could, for example, explain why Mazzi *et al.* (2002) did not find increased levels of FA in lateral bone plates after experimental inbreeding, because they studied a marine population in which the number of plates probably did not change much for long periods. It is also important to note that the associations were relatively weak, which could in part be due to the fact that only seven microsatellite markers were scored, but may also be an inherent property of the studied patterns. Furthermore, other factors, like water quality could affect levels of asymmetry as well, hereby weakening associations between FA and genetic variation.

Associations between genetic variation and asymmetry for the traits under directional selection appeared comparable among traits as the estimated variance in slopes among traits equalled to zero. However, the power to detect between-trait variation was probably very low. Associations between genetic variation and asymmetry were relatively weak, and single-trait associations were only statistically significant for pelvic spines. Visual inspection of the association between the degree of correlation between MLH and FA and P_{ST} (Fig. 5) suggest that there may also be a role for trait functionality since the correlation is relatively weak for P4P8. Such comparisons, however, will require a much higher number of populations.

We relied on phenotypic divergence (P_{ST}) relative to differentiation at neutral genetic markers (F_{ST}) to infer degrees of divergent selection. Such an interpretation is only valid under restricted and probably unrealistic assumptions (e.g. Raeymaekers *et al.*, 2007). Raeymaekers *et al.* (2007) were able to provide evidence of divergent selection between lowland and upland populations for plate number in the same study system by comparing the level of divergence at *Eda*, the gene underlying the development of the lateral plates (Colosimo *et al.*, 2005), with neutral genetic divergence. Several other studies have found indications that the rapid and parallel evolutionary changes in the three-spined stickleback have a genetic basis (e.g. Hatfield, 1997; Peichel *et al.*, 2001; Shapiro *et al.*, 2004) and changes may be adaptive (e.g. Bergstrom & Reimchen, 2000, 2003; Marchinko & Schluter, 2007). Thus, it seems warranted to conclude that the observed phenotypic differences between lowland and upland freshwater populations in our study area have a genetic basis, although the role of additive and non-additive effects may differ across traits.

Our results are not confounded by directional asymmetry, which may have a genetic basis, since levels of directional asymmetry were very small and only significant in one trait (Table 1). It has been shown that the reduction of the length of the pelvic spine may be associated with the gene *PitX1*, which in turn causes directional asymmetry (Shapiro *et al.*, 2004). Our results might suggest that this gene is either not involved in the reduction of spine length in our populations, or that at least it hardly affects levels of directional asymmetry. Moreover, for the other traits, there is no evidence for a genetic tendency to be more asymmetrical in a particular direction.

We thus provide evidence of a role of recent evolutionary changes shaping associations between developmental instability and population-level genetic variation. Future studies will be required to confirm if this is a general phenomenon and to investigate if associations at the individual level are affected in a similar way. Finally, studies on the fitness consequences of the reduced genetic variation observed here will be required to elucidate the role of DI as a putative marker of genetic stress.

Acknowledgements

We thank S. Geldof for field support and technical assistance. JAMR is a post-doctoral fellow of FWO-Vlaanderen. Research was sponsored by project G.0142.03 of FWO-Vlaanderen. We thank two anonymous reviewers for their very constructive suggestions and comments on earlier versions of this manuscript.

References

Amos, W. & Balmford, A. 2001. When does conservation genetics matter? *Heredity* **87**: 257–265.

Badyaev, A.V. 1998. Environmental stress and developmental stability in dentition of the Yellowstone grizzly bear. *Behav. Ecol.* **9**: 339–344.

Badyaev, A.V. & Foresman, K.R. 2000. Extreme environmental change and evolution: stress-induced morphological variation is strongly concordant with patterns of evolutionary divergence in shrew mandibles. *Proc. R. Soc. Lond. Ser. B* **267**: 371–377.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. (2002) *GENETIX 4.04: Logiciel Sous Windows Pour la Génétique Des Populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier.

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

Bergstrom, C.A. & Reimchen, T.E. 2000. Functional implications of fluctuating asymmetry among endemic populations of *Gasterosteus aculeatus*. *Behaviour* **137**: 1097–1112.

Bergstrom, C.A. & Reimchen, T.E. 2003. Asymmetry in structural defenses: insights into selective predation in the wild. *Evolution* **57**: 2128–2138.

Bjorksten, T., David, P., Pomiankowski, A. & Fowler, K. 2000. Fluctuating asymmetry of sexual and nonsexual traits in stalk-eyed flies: a poor indicator of developmental stress and genetic quality. *J. Evol. Biol.* **13**: 89–97.

Breuker, C.J., Patterson, J.S. & Klingenberg, C.P. 2006. A single basis for developmental buffering of *Drosophila* wing shape. *PLoS ONE* **1**: e7. doi:10.1371/journal.pone.0000007.

Cano, J.M., Makinen, H.S., Leinonen, T., Freyhof, J. & Merila, J. 2008. Extreme neutral genetic and morphological divergence supports classification of Adriatic three-spined stickleback (*Gasterosteus aculeatus*) populations as distinct conservation units. *Biol. Conserv.* **141**: 1055–1066.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G. Jr, Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D. & Kingsley, D.M. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of *Ectodysplasin* alleles. *Science* **307**: 1928–1933.

David, P. 1998. Heterozygosity-fitness correlations: new perspectives on old problems. *Heredity* **80**: 531–537.

Ellegren, H. 2008. Sequencing goes 454 and takes large-scale genomics into the wild. *Mol. Ecol.* **17**: 629–631.

Elphinstone, M.S., Hinten, G.N., Anderson, M.J. & Nock, C.J. 2003. An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. *Molecular Ecology Notes* **3**: 317–320.

Frommen, J.G., Luz, G., Mazzi, D. & Bakker, T.C.M. 2008. Inbreeding depression affects fertilization success and survival but not breeding coloration in threespined sticklebacks. *Behaviour* **145**: 425–442.

Gangestad, S.W. & Thornhill, R. 1999. Individual differences in developmental precision and fluctuating asymmetry: a model and its implications. *J. Evol. Biol.* **12**: 402–416.

Hansen, T.F., Carter, A.J.R & Pélabon, C. 2006. On adaptive accuracy and precision in natural populations. *American Naturalist* **168**: 168–181.

Hartl, G.B., Suchentrunk, F., Willing, R. & Petznek, R. 1995. Allozyme heterozygosity and fluctuating asymmetry in the brown hare (*Lepus europaeus*): a test of the developmental homeostasis hypothesis. *Phil. Trans. R. Soc. Lond. B* **350**: 313–323.

Hatfield, T. 1997. Genetic divergence in adaptive characters between sympatric species of stickleback. *American Naturalist* **149**: 1009–1029.

Houle, D. 1997. Comment on 'A meta-analysis of the heritability of developmental stability' by Møller and Thornhill. *J. Evol. Biol.* **10**: 17–20.

Klingenberg, C.P. 2003a. A developmental perspective on developmental instability: theory, models and mechanisms. In: *Developmental Instability: Causes and Consequences* (M. Polak, ed), pp. 14–34. Oxford University Press, New York.

Klingenberg, C.P. 2003b. Developmental instability as a research tool: using patterns of fluctuating asymmetry to infer the developmental origins of morphological integration. In: *Developmental Instability: Causes and Consequences* (M. Polak, ed), pp. 427–442. Oxford University Press, New York.

Leinonen, T., Cano, J.M., Mäkinen, H. & 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., O'Hara, R.B., Cano, J.M. & Merilä, J. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *J. Evol. Biol.* **21**: 1–17.

Lens, L., Van Dongen, S., Galbusera, P., Schenck, T., Matthysen, E. & Van de Casteele, T. 2000. Developmental instability and inbreeding in natural bird populations exposed to different levels of habitat disturbance. *J. Evol. Biol.* **13**: 889–896.

Lens, L., Van Dongen, S., Kark, S. & Matthysen, E. 2002a. Fluctuating asymmetry as an indicator of fitness: can we bridge the gap between studies? *Biol. Rev.* **77**: 27–38.

Lens, L., Van Dongen, S. & Matthysen, E. 2002b. Fluctuating asymmetry as early warning system in the critically endangered Taita thrush. *Conserv. Biol.* **13**: 479–487.

Manning, J.T. & Chamberlain, A.T. 1994. Fluctuating asymmetry in Gorilla canines: a sensitive indicator of environmental stress. *Proc. R. Soc. Lond. Ser. B* **255**: 189–193.

Marchinko, K.B. & Schluter, D. 2007. Parallel evolution by correlated response: Lateral plate reduction in three-spined stickleback. *Evolution* **61**: 1084–1090.

Mazzi, D., Largiadèr, C.R. & Bakker, T.C.M. 2002. Inbreeding and developmental stability in three-spined sticklebacks (*Gasterosteus aculeatus* L.). *Heredity* **89**: 293–299.

Møller, A.P. & Höglund, J. 1991. Patterns of fluctuating asymmetry in avian feather ornaments – implications for models of sexual selection. *Proc. R. Soc. Lond. Ser. B* **245**: 1–5.

Møller, A.P. & Swaddle, J.P. 1997. *Asymmetry, Developmental Stability, and Evolution*. Oxford series in Ecology and Evolution. Oxford University Press, Oxford.

Møller, A.P. & Thornhill, R. 1998. Bilateral asymmetry and sexual selection: a meta-analysis. *Am. Nat.* **151**: 174–192.

Møller, A.P., Thornhill, R. & Gangestad, S. 1999. The biological significance of fluctuating asymmetry and sexual selection: a reply to Palmer. *Am. Nat.* **154**: 234–241.

Palmer, A.R. 1999. Detecting publication bias in meta-analyses: a case study of fluctuating asymmetry and sexual selection. *Am. Nat.* **154**: 220–233.

Pelabon, C., Hansen, T.F., Carter, A.J.R. & Houle, D. 2006. Response of fluctuating and directional asymmetry to selection on wing shape in *Drosophila melanogaster*. *J. Evol. Biol.* **19**: 764–776.

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

Polak, M. (Editor), 2003. *Developmental Instability: Causes and Consequences*. Oxford University Press, New York

Polak, M. & Taylor, P.W. 2007. A primary role of developmental instability in sexual selection. *Proc. R. Soc. Lond. Ser. B* **274**: 3133–3140.

Raeymaekers, J.A.M., Maes, G.E., Audenaert, E. & Volckaert, F.A.M. 2005. Detecting Holocene divergence in the anadromous-freshwater three-spined stickleback (*Gasterosteus aculeatus*) system. *Mol. Ecol.* **14**: 1001–1014.

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.

Raeymaekers, J.A.M., Maes, G.E., Geldof, S., Hontis, I., Nackaerts, K. & Volckaert, F.A.M. 2008. Modeling genetic connectivity in sticklebacks as a guideline for river restoration. *Evol. Appl.* **1**: 475–488.

Reed, D.H. & Frankham, R. 2003. Correlation between fitness and genetic diversity. *Conserv. Biol.* **17**: 230–237.

Shapiro, M.D., Marks, M.E., Peichel, C.L. *et al.* 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*, **428**: 717–723.

Slate, J. & Pemberton, J.M. 2002. Comparing molecular measures for detecting inbreeding depression. *J. Evol. Biol.* **15**: 20–31.

Söderman, F., Van Dongen, S., Pakkasmaa, S. & Merilä, J. 2007. Asymmetry in the moor frog *Rana arvalis*. *Oecologia* **151**: 593–604.

Taylor, W.R. & Van Dyke, G.C. 1985. Revised procedures for staining and clearing small fishes and other vertebrates for bone and cartilage study. *Cybium* **9**: 107–119.

Thelen, G.C. & Allendorf, F.W. 2001. Heterozygosity-fitness correlations in rainbow trout: effects of allozyme loci or associative overdominance? *Evolution* **55**: 1180–1187.

Van Dongen, S. 1999. Accuracy and power in fluctuating asymmetry studies: effects of sample size and number of within subject repeats. *J. Evol. Biol.* **12**: 547–550.

Van Dongen, S. 2006. Fluctuating asymmetry and developmental instability in evolutionary biology: past, present and future. *J. Evol. Biol.* **19**: 1727–1743.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. & Shipley, P. 2004. Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**: 535–538.

Whitlock, M. 1996. The heritability of fluctuating asymmetry and the genetic control of developmental stability. *Proc. R. Soc. Lond. Ser. B* **263**: 849–854.

Zachos, F.E., Hartl, G.B. & Suchentrunk, F. 2007. Fluctuating asymmetry and genetic variability in the roe deer (*Capreolus capreolus*): a test of the developmental stability hypothesis in mammals using neutral molecular markers. *Heredity* **98**: 392–400.

Received 18 September 2008; revised 28 April 2009; accepted 5 May 2009