# Detecting Holocene divergence in the anadromous–freshwater three-spined stickleback (*Gasterosteus aculeatus*) system

J. A. M. RAEYMAEKERS, G. E. MAES, E. AUDENAERT and F. A. M. VOLCKAERT
*Katholieke Universiteit Leuven, Laboratory of Aquatic Ecology, Ch. de Bériotstraat, 32, B-3000 Leuven, Belgium*

## Abstract

**The anadromous–freshwater three-spined stickleback (*Gasterosteus aculeatus*) system allows for inferring the role of adaptation in speciation with a high level of accuracy because the freshwater ecotype has evolved multiple times from a uniform anadromous ancestor. A cause for concern is that independent evolution among drainages is not guaranteed in areas with a poorly resolved glacial history. This is the case for the west European great rivers, whose downstream valleys flanked the southern limit of the late Pleistocene ice sheet. We tested for independent and postglacial colonization of these valleys hypothesizing that the relationships among anadromous and freshwater sticklebacks correspond to a raceme structure. We compared the reduction in plate number accompanying this colonization to the genetic differentiation using 13 allozyme and five microsatellite loci in 350 individuals. Overall microsatellite differentiation ($F_{ST} = 0.147$) was twice as large as allozyme differentiation ($F_{ST} = 0.066$). Although habitat-specific gene flow may mask the ancestral relationships among both ecotypes, levels of microsatellite differentiation supported the hypothesis of raceme-like divergence, reflecting independent colonizations rather than the presence of two distinct evolutionary clades. Under an infinite alleles model and in the absence of gene flow, the observed freshwater divergence might be reached after 440 (microsatellites) to 4500 (allozymes) generations. Hence, the anadromous–freshwater stickleback system most likely diverged postglacially. We conclude that the reduction in plate number in two freshwater basins probably occurred independently, and that its considerable variation among populations is not in agreement with the time since divergence.**

*Keywords*: allozymes, divergence, isolation by distance, microsatellites, raceme, species pairs

*Received 26 September 2004; revision received 5 December 2004; accepted 5 December 2004*

## Introduction

The remarkable variation in phenotypes of the three-spined stickleback (*Gasterosteus aculeatus* L.; Gasterosteidae: Teleostei) has drawn considerable attention (McKinnon & Rundle 2002). Morphological variation among anadromous and freshwater stickleback populations resulted in one of the earliest attempts to define the evolutionary significance of phenotypic polymorphism (Heuts 1947). Currently, the anadromous–freshwater stickleback system is known as one of six models for rapid speciation within the stickleback complex (reviewed in McKinnon & Rundle 2002). Each system is the result of

genetic divergence of ancestral genotypes into morphologically distinct pairs of ecotypes.

The anadromous–freshwater system is a striking example of how evolutionary mechanisms act on sticklebacks. While other stickleback systems represent unique cases with few replicates, anadromous–freshwater stickleback pairs occur on a remarkably wide scale throughout the northern hemisphere (Bell & Foster 1994). Replicate evolution of similar ecotypic forms is a common phenomenon among northern temperate species complexes (Bernatchez *et al*. 1999). Resident freshwater sticklebacks are believed to have originated repeatedly from anadromous ancestors, often tracking glacial retreat (Bell & Foster 1994; McPhail 1994). The stickleback complex is thought to form a phylogenetic raceme (Williams 1992; Bell & Andrews 1997). Anadromous populations form an ancestral core that

Correspondence: Joost Raeymaekers, Fax: + 32 16 32 45 75; E-mail: joost.raeymaekers@bio.kuleuven.ac.be

persists in the marine environment, where distribution ranges may shift in response to each glacial cycle. In contrast, extinction of freshwater populations is a probable outcome of climate change, and freshwater isolates iteratively diverge from the same ancestor, rather than from other divergent isolates. In a contemporary, regional anadromous–freshwater system, the raceme model predicts closely related, ancestor-like anadromous populations, and independent evolution of freshwater populations among different drainages. Such colonizations should also lead to a reduction of genetic diversity at neutral loci (Bell & Foster 1994).

Episodes of stickleback colonization are quickly followed by divergence in many traits, including morphology, behaviour, physiology and life history (Heuts 1947; McPhail 1994; McKinnon & Rundle 2002; Bell *et al*. 2004). The comparison of phenotypic traits among populations makes it possible to detect the effects of natural selection (Foster *et al*. 1992), provided that common ancestry can be ruled out (Reusch *et al*. 2001). The raceme-like diversification within the anadromous–freshwater stickleback system provides an unusual opportunity to minimize the effects of common ancestry (Bell & Andrews 1997). The repeated origin of freshwater populations allows for a high accuracy of adaptive inference, as derived character states can be considered to have evolved independently.

On the other hand, the raceme model or the parallel evolution of phenotypic traits among freshwater populations cannot just be assumed. The minimal degree of isolation required to do so is unknown (McPhail 1994), and varies with geographical relief, tectonic stability and habitat age (Bell & Foster 1994). McKinnon *et al*. (2004) genetically

confirmed the independent evolution of freshwater populations, maximizing such probability using samples from geographically distant and recently glaciated regions. It remains a challenge to find evidence for the raceme model on a regional scale and in areas where recent glaciation is no certainty. In this study, we revisit Heuts's (1947) study area, situated in the downstream section of the west European great rivers (Meuse, Scheldt and Rhine) flowing into the southern bight of the North Sea (Fig. 1). Palaeoclimate suggests that most freshwater habitats were recolonized by sticklebacks after the last glacial (Bell & Foster 1994). However, at that time the downstream valleys of the great rivers and the coast flanked the southern limit of the ice sheet, while the coast line traced a largely drained North Sea basin (Van der Molen & de Swart 2001). Therefore, the presence of stickleback habitat in these valleys before postglacial relaxation remains unknown. Hence, independent evolution and the possible adaptive significance of phenotypic traits cannot be derived accurately from the historical context, but requires genetic validation.

Here, we compare neutral genetic differentiation with the level of phenotypic divergence (number of lateral plates) accompanying the colonization of freshwater habitat. First, we checked whether freshwater populations show reduced plate number in comparison to anadromous populations, which is congruent with the results of Heuts (1947) and other examples of anadromous–freshwater divergence. Second, we tested whether stickleback populations are characterized by the genetic predictions of the raceme model, and whether the divergence is of Holocene ancestry. Because of an uncertain age of divergence, we used both allozymes and microsatellites to evaluate the hypothesis of
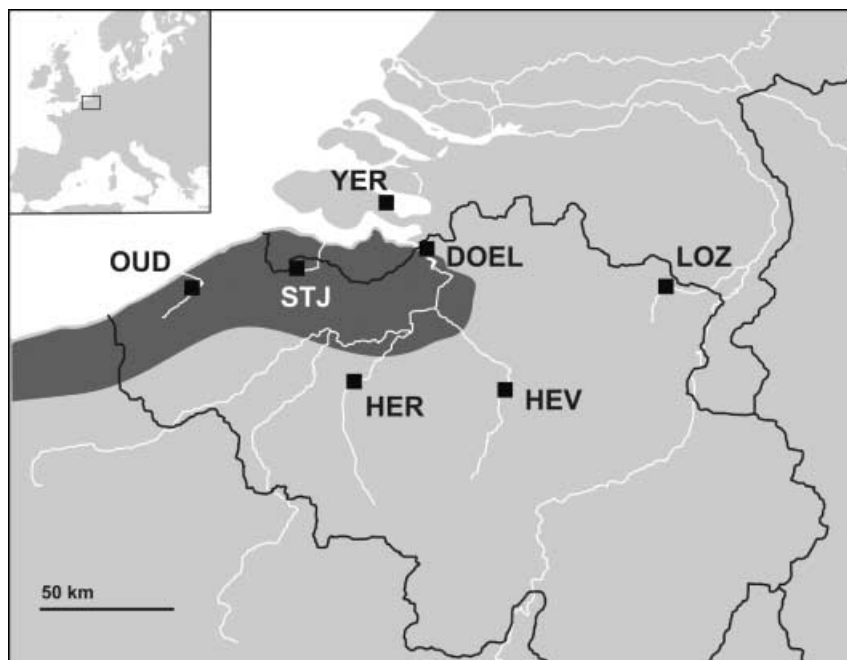


**Fig. 1** Sampling locations of four anadromous (OUD, YER, STJ and DOEL) and three freshwater (HER, HEV and LOZ) three-spined stickleback (*Gasterosteus aculeatus*) populations in the downstream section of the west European great rivers (Meuse, Scheldt and Rhine). The shaded area indicates the occurrence of anadromous populations south of the Westerschelde estuary. Abbreviations as in Table 1.

a postglacial raceme structure. We expect allozymes and microsatellites to uncover processes with a different temporal resolution. The performance of both types of markers was compared on the same individuals, providing an opportunity to check the benefits of microsatellites to assess stickleback genetic structure if divergence is rapid (see Reusch *et al.* 2001). Finally, we verify whether the divergence in plate number among colonization events can be considered independent and adaptive. Several abiotic and ecological factors have been associated with variation for this trait among freshwater basins (Bergstrom 2002), but time since divergence also seems to play a role (Klepaker 1995). Hence, any explanation of phenotypic polymorphism within the anadromous–freshwater stickleback system may have to take into account both selective forces and the time span during which these forces were acting.

## Materials and methods

### Sample collection and identification of plate morphs

Seven populations of three-spined stickleback were sampled in Belgium and the Netherlands between April 1997 and June 1998 (Fig. 1). Anadromous samples ($n = 4$) were collected in the Scheldt estuary (Yerseke and Doel) and in coastal polders (Oudenburg and Sint-Jan-in-Eremo). Freshwater samples ($n = 3$) came from affluents of the rivers Scheldt (Herzele and Heverlee) and Meuse/Maas (Lozen). The Meuse and the Scheldt represent separate basins; Herzele and Heverlee are hydrographically dependent, but the interconnecting waterway crosses the distribution zone of anadromous populations (Heuts 1947; J.A.M. Raeymaekers, unpublished; Fig. 1). About 50 individuals per population were caught either with a hand net or electric fishing gear, and flash frozen in liquid nitrogen. Prior to tissue sampling, each fish was thawed on ice and measured. Muscle tissue was sampled and stored at –80 °C for allozyme electrophoresis. Fin clips were taken and stored in 100% ethanol for DNA analysis. Fish bodies were fixed in a 7% formalin solution, bleached, and stained with alizarin red S to facilitate plate counts and plate morph registration according to Taylor & Van Dyke (1985). Only specimens larger than 20 mm standard length were examined because ontogenetic stages of small specimens might be confounded with genetically determined plate morphs (Klepaker 1996). Anterior plates were counted on the left side. Terms describing the lateral plate morphs follow Ziuganov (1983). Specimens categorized as low had 10 or fewer anterior plates, partials had 11–20 anterior plates, and completely plated specimens had 21–30 anterior plates.

### Allozyme electrophoresis

Muscle tissue was homogenized in an equal volume of distilled water at 4 °C. The samples were centrifuged at 10 000 *g* at 4 °C for 10 min. The supernatant was used to detect genotypes by horizontal starch gel electrophoresis (SGE) and cellulose acetate gel electrophoresis (CAGE). Staining procedures followed Hebert & Beaton (1989) and the nomenclature used for enzymes followed Shaklee *et al.* (1990). Buffers used in the electrophoretic analyses were Poulik (P), tris-citrate EDTA (TCE), tris-glycine (TG) and tris-maleate (TM). The following nine enzyme systems coding for 13 loci were scored: aspartate aminotransferase (AAT*, EC 2.6.1.1, CAGE, TM), creatine kinase (CK*, EC 2.7.3.2, SGE, TCE), fumarate hydratase (FH*, EC 4.2.1.2, CAGE, TM), glucose-6-phosphate isomerase (GPI-1*, GPI-2*, EC 5.3.1.9, SGE, P), isocitrate dehydrogenase (IDH*, EC 1.1.1.42, SGE, TCE), L-lactate dehydrogenase (LDH*, EC 1.1.1.27, SGE, TCE), malate dehydrogenase (MDH-1*, MDH-2*, EC 1.1.1.37, SGE, TCE), mannose-6-phosphate isomerase (MPI-1*, MPI-2*, EC 5.3.1.8, CAGE, TG) and phospho-glucomutase (PGM-1*, PGM-2*, EC 5.4.2.2, CAGE, TM).

### DNA extraction and microsatellite amplification

Allelic variation was assessed at five microsatellite loci (Gacµ4, Gacµ7, Gacµ9, Gacµ10 and Gacµ14; Taylor 1998). DNA samples were extracted from fin clips using 100 µL of a 5% Chelex (Bio-Rad) extraction solution. The samples were heated to 96 °C for 20 min, cooled on ice, vortexed and centrifuged (10 000 *g*) for 5 min. For all primer sets, 1 µL of DNA was amplified in 10 µL polymerase chain reactions (PCR). The reaction consisted of an initial denaturation step of 3 min at 95 °C, followed by 35 cycles of 45 s at 95 °C, 30 s at annealing (56 °C or 60 °C according to Taylor 1998), and 30 s extension (72 °C). A final elongation step of 10 min at 72 °C was performed. Single reactions contained 1× PCR buffer (supplied with polymerase), $MgCl_2$ at a concentration of 1.5 mM or 2 mM (Gacµ9), 200 µM of dNTP, 0.5 µM (Gacµ4, Gacµ9 and Gacµ10) or 1 µM (Gacµ7 and Gacµ14) of forward primer (IRD700 or IRD800-labelled, Westburg), the same amounts of nonlabelled reverse primer, 0.25 U of Goldstar *Taq* polymerase (Eurogentec) and 10–100 ηg of gDNA. Double distilled water was added up to 10 µL. PCR products were electrophoresed and detected on an automated sequencer (LICOR 4200) using the software E-SEQ 2.00 (LICOR). Products were scored using the software GENE IMAGIR 4.03 (Scanalytics).

### Data analysis

A morphometric index (Bell 1982) was used to quantify variation in plate morphology. The index is calculated per sample using the equation:

$$\text{Morph index} = \frac{L + 2LK + 3P + 4C}{N} \qquad \text{(eqn 1)}$$

where $N$ is the total number of specimens; $L$ is the number of low-plated specimens without a caudal keel; $LK$ is the number of low-plated, keeled specimens; $P$ is the number of partially plated specimens; and $C$ is the number of completely plated specimens. This index measures the plate armour of a sample expressed as its 'averaged' plate morph, with values ranging from 1 (exclusively low-plated, unkeeled specimens) to 4 (completely plated specimens).

Genetic diversity was evaluated based on genotype and allele frequencies, the level of polymorphism, the observed and unbiased expected heterozygosities ($H_O$ and $H_E$), the number of alleles and the mean number of alleles per locus using GENETIX 4.04 (Belkhir *et al.* 2002). Departures from Hardy–Weinberg equilibrium (HWE) were calculated and tested with GENEPOP 3.4 (Raymond & Rousset 1995). Tests for linkage disequilibrium (LD) were performed according to the method of Black & Krafsur (1985) implemented in GENETIX.

Population differentiation was quantified in GENETIX using the standardized allelic variance $F_{ST}$, estimated as $\theta$ for the allozyme and microsatellite loci (Weir & Cockerham 1984). Pairwise genetic distances ($D_{CE}$) were calculated according to Cavalli-Sforza & Edwards (1967), considering drift as the only force acting on genetic variability with fluctuating effective population size. $F_{ST}$s and $D_{CE}$s were tested for significance against $10^4$ random permutations of the data in GENETIX. For microsatellite loci, genetic structure was also assessed by the specific distance ($d\mu)^2$ (Goldstein *et al.* 1995) obtained with SPAGEDI 1.1 (Hardy & Vekemans 2002). The resulting distance matrices were visualized by two-dimensional classical metric multidimensional scaling (CMDS) plots with the function CMDSCALE in S-PLUS (MathSoft). A phylogenetic tree was constructed using the maximum likelihood algorithm CONTML in PHYLIP 3.5 (Felsenstein 1995). CONTML estimates phylogenies under a model in which all divergence is a result of genetic drift in the absence of new mutations. Allele frequencies were bootstrapped 1000 times for assessing the significance of each tree cluster using the PHYLIP subprograms SEQBOOT, CONTML and CONSENSE.

Three scenarios for the origin of the observed genetic structure were considered: isolation by distance, the raceme model, and a 'two clades' model under which anadromous and freshwater populations represent separate clades, allowing for only a single origin for the freshwater ecotype. In order to test for isolation by distance, the correlation between genetic distances and geographical distances along waterways among populations was tested using a simple Mantel test module (Mantel 1967) programmed in S-PLUS. To check the predictions of the raceme model, we compared the genetic distance matrices with a hypothetical matrix, similar to Lemaire *et al.* (2000). Under raceme-like differentiation, we expect the relative homogeneity of the anadromous group, and freshwater populations to be more related to the anadromous populations than to other freshwater populations. This translates in hypothetical dissimilarities with values 0, 1, and 2 assigned to anadromous, anadromous–freshwater and freshwater population pairs, respectively. These values correspond to the number of freshwater populations in each pair. Positive correlations between genetic distances and hypothetical dissimilarities were tested with a simple Mantel test. To control for redundancy between the dissimilarity matrix and geographical distance, we calculated and tested partial correlations between genetic distances and raceme dissimilarities after correction for geographical distance. This was done with a partial Mantel test module permuting under a null model (Legendre 2000) in S-PLUS. Because the number of Mantel test permutations is small for small sample sizes ($n = 7$) (Legendre 2000), complete enumeration of all possible $7! = 5040$ permutations was carried out for all simple and partial tests. As the null model partial Mantel test has an inflated type I error under non-normality (Legendre 2000), genetic distance matrices were inspected for kurtosis and outliers.

In order to test the validity of the raceme model vs. the two clades model, phylogenetic trees were constructed to simulate those models and tested against the most likely tree topology with CONTML in PHYLIP 3.5 (Kishino-Hasegawa test; Felsenstein 1995). We also assessed population structure with Bayesian Markov chain Monte Carlo (MCMC) assignment methods using multilocus genotypes, in order to compare the predictions of each model with the most likely population structure and with the assignment performance under a predefined raceme structure. The most likely structure was calculated assuming no admixture (given the existence of ecotypes) and correlated allele frequencies with STRUCTURE 2.1 (Pritchard *et al.* 2000). Each run, considering population structure according to a specific number of groups ($1 = K = 7$), consisted of three chains of $10^5$ MCMC replicates, initiated by $10^4$ burn-in steps. The congruence of the most likely result with the raceme model and the two clades model in terms of number of groups and assignment success was evaluated. The predefined raceme structure (i.e. separate freshwater populations and a single anadromous cluster) was assessed in GENECLASS 2.0 (Piry *et al.* 2004). Freshwater individuals were assigned with and without their population of origin included in the raceme. In the first case, individuals should be assigned primarily towards their own population. However, if the raceme model is valid, the anadromous cluster should represent a good second best destination. In the second case, assignments should be directed primarily to the anadromous cluster. Under the two clades model, assignments should reveal strong affinities among freshwater populations. Finally, this approach was also used to assign freshwater individuals to predefined anadromous plate morph groups.

## Results

### Plate counts and plate morphs

Plate number (Table 1) clearly differed among samples (Kruskal–Wallis ANOVA, $P < 0.0001$) and anadromous populations tended to be more armoured than freshwater populations (Mann–Whitney $U$-test, $P < 0.0001$). In two anadromous samples, Oudenburg and Doel, each of the three morphs was represented at moderate frequencies, resulting in mean plate numbers around 18 and a morph index of more than 3. Two freshwater samples, Heverlee and Herzele, were monomorphic for the low-plated morph, with mean plate numbers below 5 and a morph index of 2. Interestingly, one anadromous (Yerseke) and one freshwater sample (Lozen) showed a similar pattern with an intermediate morph index around 2.5 and mean plate numbers around 11. This was significantly lower than in other anadromous samples for the former case and significantly higher than in other freshwater samples for the latter case (Mann–Whitney $U$-tests, $P < 0.0001$). In Sint-Jan-in-Eremo, the single fish sufficiently large for scoring the number of plates belonged to the complete-plated morph.

### Genetic variability and Hardy–Weinberg equilibrium

*Allozymes.* Ten out of 13 loci were polymorphic (Appendix I). The total number of alleles per locus ranged from one to four. Within locations we found 1.08 (Herzele) to 1.54 (Oudenburg) alleles per locus (Table 2), and the average observed heterozygosity ($H_O$) was 0.045. Anadromous samples showed a higher genetic diversity than freshwater

**Table 1** Sampling location, standard length (means ± SD), plate number range and plate morph variation in four anadromous (A) and three freshwater (F) samples of *Gasterosteus aculeatus* from Belgium (B) and the Netherlands (NL). The morph distribution is presented in percentages and as a morph index

| Sampling station; system | Code | Standard length (mm) | Range | Complete | Partial | Low | Morph index |
|---|---|---|---|---|---|---|---|
| Oudenburg (B/A); coastal polders | OUD | 47.55 ± 4.79 | 6–31 | 38.78 | 36.73 | 24.49 | 3.14 |
| Yerseke (NL/A); Oosterschelde estuary | YER | 31.56 ± 3.78 | 6–28 | 10 | 34 | 56 | 2.54 |
| Sint-Jan-in-Eremo (B/A); Westerschelde polders | STJ | 21.16 ± 3.42 | — | 1 | — | — | (4) |
| Doel (B/A); Scheldt estuary | DOEL | 52.62 ± 6.90 | 6–25 | 34 | 56 | 10 | 3.24 |
| Herzele (B/F); Scheldt | HER | 46.66 ± 4.31 | 3–6 | 0 | 0 | 100 | 2 |
| Heverlee (B/F); Dijle/Scheldt | HEV | 26.90 ± 3.88 | 2–9 | 0 | 0 | 100 | 2 |
| Lozen (B/F); Aabeek/Meuse | LOZ | 39.16 ± 2.38 | 5–28 | 4 | 36 | 60 | 2.44 |

**Table 2** Genetic variability estimates at 13 allozyme loci (top) and five microsatellite loci (bottom) in seven samples of *Gasterosteus aculeatus*: average number of alleles per locus (MNA), observed and expected (nonbiased or nb) heterozygosities ($H_O$ and $H_E$; means ± SD) and percentage of polymorphic loci (0.95 and 0.99 criteria). Sample abbreviations as in Table 1

| | Sample | MNA | $H_O$ | $H_E$ (nb) | $P_{0.95}$ | $P_{0.99}$ |
|---|---|---|---|---|---|---|
| Allozymes | OUD | 1.539 | 0.051 ± 0.109 | 0.061 ± 0.133 | 0.154 | 0.385 |
| | YER | 1.231 | 0.052 ± 0.115 | 0.049 ± 0.108 | 0.231 | 0.231 |
| | STJ | 1.308 | 0.059 ± 0.145 | 0.059 ± 0.138 | 0.154 | 0.308 |
| | DOEL | 1.385 | 0.056 ± 0.112 | 0.062 ± 0.134 | 0.231 | 0.231 |
| | HER | 1.077 | 0.039 ± 0.139 | 0.039 ± 0.140 | 0.077 | 0.077 |
| | HEV | 1.308 | 0.023 ± 0.064 | 0.027 ± 0.068 | 0.077 | 0.154 |
| | LOZ | 1.154 | 0.037 ± 0.127 | 0.039 ± 0.133 | 0.077 | 0.077 |
| | All samples | 2.230 | 0.045 ± 0.107 | 0.051 ± 0.121 | 0.154 | 0.231 |
| Microsatellites | OUD | 13.2 | 0.717 ± 0.203 | 0.797 ± 0.182 | 1.000 | 1.000 |
| | YER | 13 | 0.695 ± 0.202 | 0.796 ± 0.184 | 1.000 | 1.000 |
| | STJ | 13.2 | 0.713 ± 0.244 | 0.811 ± 0.150 | 1.000 | 1.000 |
| | DOEL | 14.6 | 0.724 ± 0.185 | 0.808 ± 0.140 | 1.000 | 1.000 |
| | HER | 6.2 | 0.604 ± 0.201 | 0.619 ± 0.226 | 1.000 | 1.000 |
| | HEV | 8.8 | 0.536 ± 0.329 | 0.559 ± 0.340 | 0.800 | 0.800 |
| | LOZ | 6.4 | 0.460 ± 0.362 | 0.492 ± 0.344 | 0.800 | 0.800 |
| | All samples | 20.4 | 0.632 ± 0.241 | 0.793 ± 0.198 | 1.000 | 1.000 |

samples in terms of $H_O$ ($P = 0.034$), but not in terms of allelic richness ($P = 0.083$). Rare alleles were relatively important in the Scheldt population at Heverlee (Table 2). Only one locus pair (MDH-2* and PGM-2*) was in LD after correction for multiple testing over populations and locus pairs in Oudenburg ($P = 0.0001$).

*Microsatellites.* The total number of alleles per locus ranged from six (Gacµ14) to 42 (Gacµ10) and within locations from 6.2 (Herzele) to 14.6 (Doel) per locus (Table 2; Appendix II). All loci were polymorphic in all populations, except Gacµ14 in the freshwater samples Lozen and Heverlee. Anadromous samples showed a higher genetic diversity than freshwater samples in terms of $H_O$ (0.712 vs. 0.534; $P < 0.05$) and in terms of allelic richness ($P < 0.05$). All anadromous populations deviated from HWE, which was a result of just one locus, Gacµ9. Significant linkage between Gacµ9 and Gacµ10 occurred in all anadromous populations. A null allele with an estimated frequency of 21%–34% might be responsible for the disequilibria associated with Gacµ9 in these populations. The locus was retained in all analyses, as neither its exclusion, nor corrected allele frequencies strongly affected the differentiation with and among anadromous populations.

### Genetic structure

Overall genetic differentiation at allozyme loci was moderate ($F_{ST} = 0.066$; $P < 0.01$). Pairwise $F_{ST}$s over all loci were significant for nine out of 21 tests after Bonferroni correction (Table 3). All pairwise genetic distances ($D_{CE}$) between samples were significant, except Herzele vs. Lozen. Lozen appeared as the least differentiated freshwater population. Most allozyme divergence concentrated in the loci MDH-2*, PGM-2* and GPI-1*. For microsatellites, overall genetic

differentiation was high ($F_{ST} = 0.147$; $P < 0.0001$). Pairwise $F_{ST}$s and pairwise $D_{CE}$s between all samples were highly significant after Bonferroni correction ($P < 0.0001$; Table 3). The anadromous genetic differentiation (allozymes: $F_{ST} = 0.038$; microsatellites: $F_{ST} = 0.047$) was considerably lower than the freshwater genetic differentiation (allozymes: $F_{ST} = 0.118$; microsatellites: $F_{ST} = 0.255$), especially for microsatellites ($P = 0.022$). There was no evidence for differentiation among the three anadromous plate morph groups. Complete lists of allele frequencies can be obtained from the Author for correspondence upon request.

### Model testing

Unlike allozymes, microsatellites showed a clear relationship between geographical distance and genetic divergence (as linearized $F_{ST}$ Rousset 1997; Fig. 2). However, when accounting for anadromous, anadromous–freshwater and freshwater population pairs, linearized microsatellite $F_{ST}$s (Fig. 2b) seemed to increase according to the number of freshwater populations involved. In addition, the scatter of population–pair categories around the isolation-by-distance least squares fit is biased, as anadromous and freshwater pairs, respectively, tend to be lower and higher than the average isolation by distance trend. There was no such pattern for allozymes. Figure 3 shows a visual comparison of linearized genetic distances and raceme model dissimilarities with CMDS plots. The Mantel correlation between linearized allozyme $F_{ST}$s and raceme model dissimilarities was not significant ($r_M = 0.49$; $P = 0.076$); anadromous populations and Lozen clustered in between both Scheldt populations (Fig. 3b). For linearized microsatellites $F_{ST}$s, the resemblance of the CMDS plot with the raceme model CMDS plot was good (Fig. 3c). The

| | | OUD | YER | STJ | DOEL | HER | HEV | LOZ |
|---|---|---|---|---|---|---|---|---|
| Allozymes | OUD | — | 0.057* | 0.044 | 0.035 | 0.133* | 0.070* | 0.075* |
| | YER | 0.054* | — | 0.053 | 0.019 | 0.124* | 0.029 | 0.052 |
| | STJ | 0.057* | 0.043* | — | 0.031 | 0.039 | 0.101* | 0.012 |
| | DOEL | 0.068* | 0.033* | 0.053* | — | 0.103* | 0.047 | 0.046 |
| | HER | 0.045* | 0.029* | 0.023* | 0.044* | — | 0.228* | 0.013 |
| | HEV | 0.043* | 0.020* | 0.036* | 0.037* | 0.027* | — | 0.120* |
| | LOZ | 0.040* | 0.016* | 0.022* | 0.033* | 0.003 | 0.016* | — |
| Microsatellites | OUD | — | 0.018* | 0.027* | 0.064* | 0.116* | 0.109* | 0.183* |
| | YER | 0.031* | — | 0.037* | 0.075* | 0.124* | 0.123* | 0.173* |
| | STJ | 0.040* | 0.039* | — | 0.071* | 0.149* | 0.166* | 0.224* |
| | DOEL | 0.037* | 0.045* | 0.047* | — | 0.177* | 0.191* | 0.253* |
| | HER | 0.070* | 0.069* | 0.078* | 0.076* | — | 0.198* | 0.289* |
| | HEV | 0.060* | 0.068* | 0.076* | 0.075* | 0.095* | — | 0.277* |
| | LOZ | 0.087* | 0.086* | 0.100* | 0.100* | 0.118* | 0.093* | — |

**Table 3** Pairwise genetic differentiation and genetic distances between seven samples of *Gasterosteus aculeatus* calculated from 10 allozyme loci (top) and five microsatellite loci (bottom): pairwise $F_{ST}$ estimates (θ, above diagonal) calculated following Weir & Cockerham (1984) and pairwise genetic distances ($D_{CE}$, below diagonal) calculated following Cavalli-Sforza & Edwards (1967). Sample abbreviations as in Table 1

*Significant after Bonferroni correction.
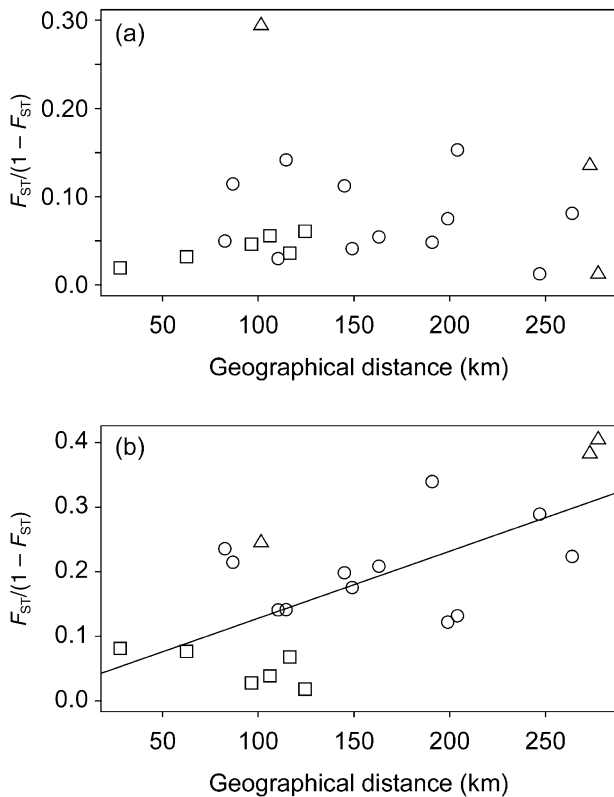
Fig. 2 Test for isolation by distance along watercourses between seven three-spined stickleback populations. (a) Linearized genetic distance ($F_{ST}$) based on 10 polymorphic allozyme loci vs. geographical distance ($r = -0.004$; $P = 0.45$). (b) Linearized genetic distance based on five polymorphic microsatellite loci vs. geographical distance ($r = 0.66$; $P = 0.044$). Squares, circles and triangles refer to anadromous population pairs, anadromous-freshwater population pairs and freshwater population pairs, respectively.

corresponding Mantel correlation was significantly positive ($r_M = 0.86$; $P = 0.016$), even after correction for geographical distance (partial Mantel test; $r_M = 0.77$, $P = 0.016$). The analyses based on $D_{CE}$ and $(d\mu)^2$ (not shown) showed similar but weaker trends. The Mantel correlation between geographical distances and hypothetical raceme model dissimilarities was positive but not significant ($r_M = 0.60$; $P = 0.0859$), confirming that the effect of geographical distance and the effect of the anadromous/freshwater categorization of population pairs are not equivalent.

A maximum likelihood tree based on microsatellite data (Ln likelihood = 825.02; Fig. 4) grouped freshwater and anadromous stickleback populations according to the two clades model. However, the branch splitting both ecotypes, was not supported by a high bootstrap value (44%). Moreover, any alternative raceme-like tree topology, e.g. [OUD,(YER,LOZ),(STJ,(DOEL,(HER,HEV)))], allowing for two independent freshwater origins (Ln likelihood = 821.95; $P > 0.05$) and [(OUD,LOZ),(YER,(HER,DOEL),(STJ,HEV)], allowing for three independent freshwater origins (Ln likelihood = 821.86; $P > 0.05$) was found to be worse than the maximum likelihood tree. Better insight was obtained with Bayesian assignment analyses; individual-based clustering considering no admixture and correlated allele frequencies revealed that the smallest number of groups capturing the major structure in the microsatellite data was $K = 4$. These four groups, respectively, accounted for 85.2%, 99.8%, 99.4% and 99.9% membership of the grouped anadromous populations and the three separate freshwater populations (Herzele, Heverlee and Lozen). Such results are highly congruent with the expectations of the raceme model; the strong and asymmetrical assignment success provides an indication of reliable population structure.
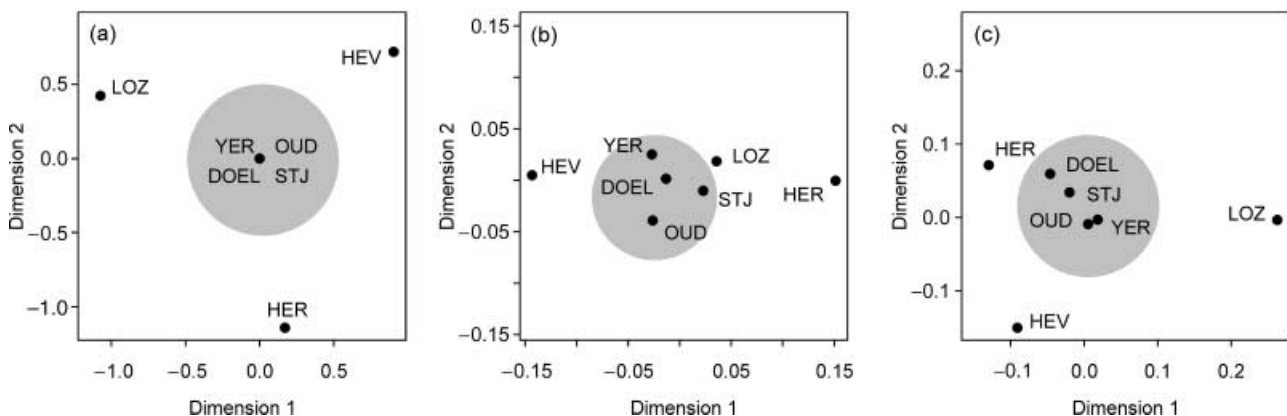


Fig. 3 Comparison of dissimilarities in the raceme model with linearized genetic distances among seven three-spined stickleback populations by classical metric multidimensional scaling (CMDS). (a) Visualization of hypothetical raceme model dissimilarities by CMDS (stress value: 0.13). (b) CMDS plot for linearized allozyme $F_{ST}$s (stress value: 0.28). (c) CMDS plot for linearized microsatellite $F_{ST}$s (stress value: 0.25). Shaded areas group anadromous populations. For abbreviations, see Table 1.
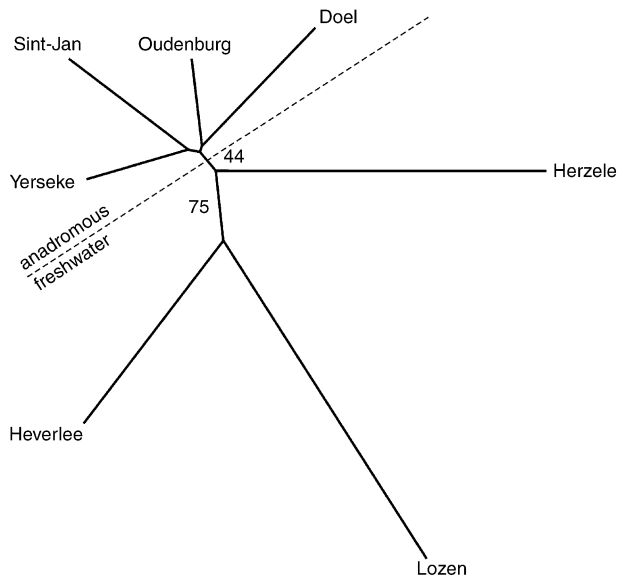
**Fig. 4** Unrooted maximum likelihood tree of the relationships among seven three-spined stickleback populations based on microsatellite data. Numbers along branches represent percentage support > 40% based on 1000 bootstrap replicate analyses of the allele frequency matrix.

Assignments to the predefined raceme structure were consistent with the above findings, as a very high degree of the freshwater individuals was allocated to the anadromous cluster if the population of origin was not included (Herzele 98%; Heverlee 100%; Lozen 94%). When all populations were included, 95.8% of the anadromous individuals and all freshwater individuals were assigned to the predefined cluster of origin. For 97.3% of the freshwater individuals, the anadromous cluster (rather than another freshwater population) represented the second best assignment destination. Allozyme data revealed no interpretable structure in any assignment analysis as no clusters were detected (*K* = 1). Assignment of freshwater individuals to predefined anadromous plate morph groups neither revealed meaningful results (41%, 31% and 28% were assigned to the complete, partial, and low-plated anadromous morphs, respectively).

## Discussion

### Divergence in lateral plate number

Heuts (1947) discovered the discontinuous distribution of morphological variation in Belgian stickleback populations. In western Belgium, coastal populations were anadromous and characterized by a high number of lateral plates. In the central part, populations were resident and carried a low plate number. Morphologically intermediate populations were rare. Our data confirm the general distribution pattern of plate number according to Heuts (1947), with the clear dominance of the low-plated morph in freshwater

populations. However, low-plated specimens were also common in anadromous populations. Consistent with Heuts' conclusion that the anadromous and freshwater ecotypes are genetically differentiated, Klepaker (1996) stated that plate morph variation in anadromous populations reflects true polymorphism and neither immigration of freshwater sticklebacks nor hybridization. This is confirmed by our results, as there was no genetic differentiation among anadromous plate morph groups, and freshwater individuals showed no affinities with any of these groups in particular.

Reduction in lateral plate number and armour loss is a common feature of freshwater populations (Klepaker 1995; Bergstrom 2002; Bell *et al.* 2004). However, the freshwater population of the Meuse basin (Lozen) showed only a moderate reduction in plate number, unlike the low-plated Scheldt populations. Heuts (1947) also discovered some populations in freshwater habitats that were not entirely low plated. Variation in plate number among basins and on a coastal–inland cline remains to be documented in the area, and is the focus of an ongoing project.

### Genetic variability and genetic structure

If the raceme model applies, postglacially derived freshwater populations should show reduced genetic diversity at neutral loci compared to their anadromous ancestors (Bell & Foster 1994). They also should contain a random sample of the alleles present in adjacent anadromous populations. Freshwater populations indeed exhibited lower (allozyme and microsatellite) heterozygosities and lower (microsatellite) allelic diversity than anadromous populations. Here, we assume that modern anadromous populations resemble ancestral genotypes more closely than the derived freshwater genotypes, which represents another prediction of the raceme model. Allelic allozyme and microsatellite diversities in freshwater populations represented subsets of the anadromous allelic diversity. Reduced genetic variation within founder (freshwater) populations relative to larger (anadromous) source populations has been observed in sticklebacks on several occasions (allozymes: Rafinski *et al.* 1989; mtDNA: Taylor & McPhail 1999; microsatellites: Taylor & McPhail 2000; Reusch *et al.* 2001).

Allozyme and microsatellite estimates of divergence for anadromous–freshwater pairs are typically low (Rafinski *et al.* 1989; Taylor & McPhail 2000), which is indicative of a recent post-Pleistocene origin (McKinnon & Rundle 2002). In this study, both allozyme and microsatellite markers differentiated our populations, but microsatellite differentiation was larger than allozyme differentiation. Pairwise $F_{ST}$s, especially for microsatellites, revealed a close relationship between anadromous populations and distant freshwater isolates. This is also consistent with the raceme model, although other explanations are possible (see succeeding discussion). Recent studies estimating stickleback

genetic divergence focused on benthic–limnetic (Taylor & McPhail 2000), lake–stream (Reusch *et al.* 2001; Hendry *et al.* 2002), and anadromous–freshwater species pairs (Taylor & McPhail 2000; Reusch *et al.* 2001; McKinnon *et al.* 2004). These studies used microsatellite markers to unveil recent processes, comparing multiple genetic distances and mutation models. In contrast, we question the timing of divergence, and judge two types of nuclear markers to be more relevant to cover the evolutionary time span.

### Isolation by distance, two clades or raceme-like divergence?

As discussed previously, genetic variability and divergence within the anadromous–freshwater stickleback system supported the raceme model, assuming independent derivation of freshwater populations from anadromous ones. However, obligate freshwater fish generally display greater levels of genetic differentiation and less genetic diversity than marine and anadromous fish, because of the isolating nature of river systems and small effective population size (DeWoody & Avise 2000). Marine populations are open to genetic exchange with conspecifics over larger areas, and gene flow between populations is on average higher. Hence, habitat-specific gene flow and genetic variability may obscure the ancestral relationships among differentiating ecotypes.

Although the microsatellite structure of stickleback populations reflects an isolation-by-distance (IBD) pattern, geographical isolation cannot be considered as the only explanation behind the variation in genetic distance. The significant effect of raceme model dissimilarities demonstrates that freshwater populations are more related to the anadromous cluster than to each other; genetic divergence is also larger between freshwater populations and lower between anadromous populations than expected from the IBD trend. Such additional variation may point to habitat-specific gene flow (which could arise in combination with any evolutionary model), or raceme-like differentiation, reflecting independent freshwater colonizations.

If only habitat-specific gene flow and not independent colonizations caused the raceme-like genetic structure, the freshwater ecotype could still represent a single evolutionary clade. As illustrated by the lack of discrimination between tree topologies by means of likelihood ratio tests, the distinction between the raceme model and such a two clades model is particularly subtle. First, both models assume common ancestry of freshwater populations, and it is crucial whether this ancestor was anadromous or not. Second, several allozyme loci and one microsatellite locus were less polymorphic. These loci started to drift towards fixation of the most common anadromous allele in each freshwater population. This is not unlikely under any model, but it creates a bias towards relatedness of freshwater populations, as expected under the two clades

model. Moreover, $D_{CE}$ values accumulate too little distance when loci are close to fixation (Felsenstein 1985). Hence, to evaluate the independence of freshwater populations relying on genetic differentiation, the level of polymorphism of the genetic markers should be high.

All together the raceme model is more likely than the two clades model, as assignment tests, allelic composition and microsatellite distances revealed strong evidence that freshwater populations are independently related to the anadromous cluster. On a worldwide scale, McKinnon *et al.* (2004) found genetic evidence for the raceme model, using samples from geographically distant and recently glaciated regions. In this study, independence of freshwater populations among close drainages flanking the late Pleistocene ice sheet was a priori much more unlikely. A principal effect of drainage has not yet been reported. Although more detail on divergence within drainages was included, Taylor & McPhail (2000) and Reusch *et al.* (2001) rather found effects of ecotype (benthic vs. limnetic and lake vs. river). However, different freshwater ecotypes do not exist in our study area. It remains unclear to what extent habitat-specific gene flow mimics or contributes to the pattern of independent colonizations. For instance, limited gene flow among freshwater habitats could explain the highly divergent but hydrographically dependent Scheldt populations. However, given the upstream distribution of anadromous populations in the Scheldt River (Heuts 1947; J.A.M. Raeymaekers, unpublished; Fig. 1), the hypothesis of independent colonizations is equally valid.

### Time since divergence

The marine core of a phylogenetic raceme is thought to persist during the ice ages (Bell & Andrews 1997), implying that independent freshwater lineages can diverge anytime. Most cases of anadromous–freshwater divergence have been attributed to postglacial differentiation, not older than 10 000–16 000 BP (McPhail 1994). However, highly divergent freshwater isolates of *Gasterosteus aculeatus* may be either much older, as they have existed for at least 10 million years (Myr), or much younger, as they may also evolve within decades (Bell *et al.* 2004). Although our data do not support a single origin of the freshwater ecotype, the divergence of a freshwater population older than the last deglaciation may be responsible for the present-day divergence. In the southern part of the Baltic floodplain (Flanders and the Netherlands), freshwater refugia have been detected for several fish taxa, such as the anadromous Atlantic salmon (*Salmo salar* L.) (Verspoor *et al.* 1999) and the freshwater bullhead (*Cottus gobio*) (Volckaert *et al.* 2002). If no such refugia existed for the three-spined stickleback, anadromous populations must have colonized the downstream valleys of the great rivers in the past 13 000 years (Van der Molen & de Swart 2001).

We evaluated a scenario where the observed level of genetic divergence between the three freshwater populations could have arisen postglacially (i.e. within at most 13 000 stickleback generations) from a common ancestral source. Using a sampling equation developed by Jin & Chakraborty (1995), we predicted the multilocus $F_{ST}$ under complete isolation as a function of time (generations), the effective population size $N_e$, the number of subpopulations ($n = 3$), and the heterozygosity ($H_O$ averaged across subpopulations). We applied this method for both allozymes and microsatellites with prediction of $F_{ST}$ values under the infinite alleles model (IAM). $N_e$ was calculated under the IAM according to the equation $N_e = (H_O/1 - H_O)/4\mu$ (Crow & Kimura 1970). Assuming a mutation rate $\mu = 10^{-3}$ for microsatellites (Ellegren 2000) and $\mu = 10^{-6}$ for allozymes (Nei 1975), $N_e$ amounted to 286 and 11 100 individuals, respectively, approaching previous freshwater stickleback $N_e$ estimates between 50 and 8300 (Reusch *et al.* 2001; Hendry *et al.* 2002). A comparison between the expected allozyme and microsatellite $F_{ST}$ as a function of generations shows that 4500 generations would be sufficient to reach an equilibrium allozyme $F_{ST} = 0.118$ in the absence of gene flow. For microsatellites, 440 generations would be sufficient to reach the observed $F_{ST}$ of 0.255. In other words, it is likely that the age of divergence within the anadromous–freshwater system is younger than the late Pleistocene glaciations.

Our study confirms the advantage of microsatellites to disentangle such recent divergence (see Reusch *et al.* 2001), in spite of traditional preferences for allozymes (McKinnon & Rundle 2002). For microsatellites, the levels of genetic diversity and divergence were consistent with the raceme model, whereas for allozymes only diversity levels were compatible. Hence, both marker sets show incomplete parallel trends, in part because the loci are presumably affected by population demographic factors operative on different historical time frames (DeWoody & Avise 2000). Both estimates of the minimal time needed for divergence also differ greatly. They strongly depend on the $N_e$ estimates, of which absolute values should be treated with caution. Hence, the allozyme estimate provides a useful replicate that increases the confidence on the Holocene origin, regardless of the models or methods behind the calculations. Our timing estimate based on two marker sets reaches a level of confidence comparable to the microsatellite study of Reusch *et al.* (2001), who also considered the stepwise mutation model (SMM). As they noted, population equilibrium under the SMM would be reached slower than under the IAM; this would inflate the timing estimate towards a level closer to the allozyme result.

The magnitude of the $F_{ST}$ for isolated populations is related to the amount of variation as determined by the mutation rate at a locus (Hedrick 1999). The maximum differentiation between pairs in freshwater populations (0.29 for microsatellites and 0.23 for allozymes; Table 3) does not approach the maximal expected values ($F_{ST} = 0.37$ and $F_{ST} = 0.94$, respectively). Maximum values of microsatellite differentiation between lake and river sticklebacks approached asymptotic values in Schleswig–Holstein (Reusch *et al.* 2001), which might be attributed to the higher average heterozygosity. Anyhow, $F_{ST}$ calculated from our microsatellite loci may also approach its limit within 13 000 possible postglacial generations. In contrast, the estimate calculated from allozymes remains far from saturated within this time frame.

### The balance between $N_e$, selection and time since divergence

Genetic markers have become favourite tools to disentangle processes of adaptive divergence as a consequence of ecological shifts (Bernatchez *et al.* 1999; Lemaire *et al.* 2000). Divergent natural and sexual selection are assumed to be the main forces driving the evolution of reproductive isolation in sticklebacks (Schluter 1996). In the anadromous–freshwater system, an important role for natural selection is suggested by the relationship between morphology and ecology for independently derived freshwater populations (McPhail 1994). Similar selection regimes may result in similar morphological outcomes (Foster *et al.* 1992), with traits evolving in parallel. As our results point to raceme-like differentiation, the reduction in plate number in freshwater sticklebacks from the Meuse and Scheldt basins probably occurred independently. This provides a strong indication that plate number in freshwater populations is adaptive (e.g. optimization of resources and fast-start-performance), or subject to relaxation of the selection conditions ruling the marine environment (e.g. reduced predation).

Armour loss seems an advantageous evolutionary pathway from anadromous to freshwater sticklebacks (Bell *et al.* 2004). However, phenotypes of different drainages did not diverge to the same extent. Several abiotic and ecological factors seem associated with variation in plate number among freshwater populations (Bergstrom 2002). Such variation has also been explained as a function of the divergence time of freshwater populations, on their way towards a low-plated end-stage, from their marine ancestors, dominated by completely plated morphs (Klepaker 1995). This means that time lags in postglacial relaxation among the west European great rivers could also explain differences in plate number. Hence, phenotypic polymorphism within the anadromous–freshwater system can fluctuate according to both selection and the time span during which selection has been active. Our microsatellite estimates of anadromous–freshwater differentiation should reflect such time span, given the absence of saturation. Although the Meuse population diverged most from the estuarine cluster, plate number reduced only moderately. A substantial

reduction in effective population size is a common characteristic of stickleback freshwater colonization (Taylor & McPhail 2000); its balance with selection (McKinnon & Rundle 2002) and time since divergence (Hedrick 1999) could explain discordant phenotypic and genetic evolution. Assuming a constant $N_e$, the discordance stresses the importance of differences among freshwater populations in selection pressure rather than in time since divergence.

## Acknowledgements

## References

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2002) GENETIX 4.04, Logiciel sous Windows TM Pour la Génétique des Populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Bell MA (1982) Differentiation of adjacent stream populations of threespine sticklebacks. *Evolution*, **36**, 189–199.

Bell MA, Andrews CA (1997) Evolutionary consequences of postglacial colonisation of fresh water by primitively anadromous fishes. In: *Evolutionary Ecology of Freshwater Animals* (eds Streit B, Städler T, Lively CM), pp. 323–363. Birkhäuser-Verlag, Basel.

Bell MA, Foster SA (1994) Introduction to the evolutionary biology of the threespine stickleback. In: *The Evolutionary Biology of the Threespine Stickleback* (eds Bell MA, Foster SA), pp. 1–26. Oxford University Press, Oxford.

Bell MA, Aguirre WE, Buck NJ (2004) Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution*, 58, 814–824.

Bergstrom CA (2002) Fast-start swimming performance and reduction in lateral plate number in threespine stickleback. *Canadian Journal of Zoology*, **80**, 207–213.

Bernatchez L, Chouinard A, Lu G (1999) Integrating molecular genetics and ecology in studies of adaptive radiation: whitefish, *Coregonus* sp., as a case study. *Biological Journal of the Linnean Society*, **68**, 173–194.

Black WC, Krafsur ES (1985) A FORTRAN program for the calculation and analysis of two-locus linkage disequilibrium coefficients. *Theoretical and Applied Genetics*, **70**, 491–496.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics*, **19**, 233–257.

Crow JF, Kimura M (1970) *An Introduction to Population Genetics Theory*. Harper & Row, New York.

DeWoody JA, Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology*, **56**, 461–473.

Ellegren H (2000) Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics*, **16**, 551–558.

Felsenstein J (1985) Phylogenies from gene frequencies: a statistical problem. *Systematic Zoology*, **34**, 300–311.

Felsenstein J (1995) *Phylogeny Inference Package*. PHYLIP 3.5. University of Washington, Seattle. Available at http://evolution.genetics.washington.edu/phylip.html.

Foster SA, Baker JA, Bell MA (1992) Phenotypic integration of life history and morphology: an example from three-spined stickleback, *Gasterosteus aculeatus* L. *Journal of Fish Biology*, **41** (Suppl. B), 21–35.

Goldstein DB, Linares AR, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics*, **139**, 463–471.

Hardy OJ, Vekemans X (2002) SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Hebert PDN, Beaton M (1989) *Methodologies for Cellulose Acetate Electrophoresis*. Helena Laboratories, Beaumont.

Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.

Hendry AP, Taylor EB, McPhail JD (2002) Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. *Evolution*, **56**, 1199–1216.

Heuts MJ (1947) The phenotypic variability of *Gasterosteus aculeatus* (L.) populations in Belgium. *Verhandelingen van de Koninklijke Vlaamse Academie voor Wetenschappen, Letteren en Schone Kunsten van België*, **9**, 1–63.

Jin L, Chakraborty R (1995) Population structure, stepwise mutations, and heterozygote deficiency and their implications in DNA forensics. *Heredity*, **74**, 274–285.

Klepaker T (1995) Postglacial evolution in lateral plate morphs in Norwegian freshwater populations of the threespine stickleback (*Gasterosteus aculeatus*). *Canadian Journal of Zoology*, **73**, 898–906.

Klepaker T (1996) Lateral plate polymorphism in marine and estuarine populations of the threespine stickleback (*Gasterosteus aculeatus*) along the coast of Norway. *Copeia*, **1996**, 832–838.

Legendre P (2000) Comparison of permutation methods for the partial correlation and partial Mantel tests. *Journal of Statistical Computation and Simulation*, **67**, 37–73.

Lemaire C, Allegrucci G, Naciri M *et al.* (2000) Do discrepancies between microsatellite and allozyme variation reveal differential selection between sea and lagoon in the sea bass (*Dicentrarchus labrax*)? *Molecular Ecology*, **9**, 457–467.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

McKinnon JS, Mori S, Blackman BK (2004) Evidence for ecology's role in speciation. *Nature*, **429**, 294–298.

McKinnon JS, Rundle HD (2002) Speciation in nature: the three-spine stickleback model systems. *Trends in Ecology and Evolution*, **17**, 480–488.

McPhail JD (1994) Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus*) of southwestern British Columbia. In: *The Evolutionary Biology of the Threespine Stickleback* (eds Bell MA, Foster SA), pp. 399–437. Oxford University Press, Oxford.

Nei M (1975) *Molecular Population Genetics and Evolution*. North Holland, Amsterdam.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first generation migrants detection. *Journal of Heredity* (accepted). Available at http://www.montpellier.inra.fr/URLB/.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Rafinski J, Banbura J, Przybylski M (1989) Genetic differentiation of freshwater and marine sticklebacks (*Gasterosteus aculeatus*) of Eastern Europe. *Zeitschrift für Zoologische Systematik und Evolutionsforschung*, **27**, 33–43.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): a population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Reusch TBH, Wegner KM, Kalbe M (2001) Rapid genetic divergence in postglacial populations of threespine stickleback (*Gasterosteus aculeatus*): the role of habitat type, drainage and geographical proximity. *Molecular Ecology*, **10**, 2435–2445.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Schluter D (1996) Ecological speciation in postglacial fishes. *Philosophical Transactions of the Royal Society London. Series B, Biological Sciences*, 351, 807–814.

Shaklee JB, Allendorf FW, Morizot DC, Whitt GS (1990) Gene nomenclature for protein-coding loci in fish. *Transactions of the American Fisheries Society*, **119**, 2–15.

Taylor EB (1998) Microsatellites isolated from the threespine stickleback *Gasterosteus aculeatus. Molecular Ecology*, **7**, 930–931.

Taylor EB, McPhail JD (2000) Historical contingency and ecological determinism interact to prime speciation in sticklebacks, *Gasterosteus. Proceedings of the Royal Society London. Series B, Biological Sciences*, **267**, 2375–2384.

Taylor EB, McPhail JD (1999) Evolutionary history of an adaptive radiation in species pairs of threespine sticklebacks (*Gasterosteus*): insights from mitochondrial DNA. *Biological Journal of the Linnean Society*, **66**, 271–291.

Taylor WR, Van Dyke GC (1985) Revised procedures for staining and clearing small fishes and other vertebrates for bone and cartilage study. *Cybium*, **9**, 107–119.

Van der Molen J, de Swart HE (2001) Holocene tidal conditions and tide-induced sand transport in the southern North Sea. *Journal of Geophysical Research–Oceans*, **106**, 9339–9362.

Verspoor E, McCarthy EM, Knox D (1999) The phylogeography of European Atlantic salmon (*Salmo salar* L.) based on RFLP analysis of the ND1/16sRNA region of the mtDNA. *Biological Journal of the Linnean Society*, **68**, 129–146.

Volckaert FAM, Hänfling B, Hellemans B, Carvalho GR (2002) Timing of the population dynamics of bullhead *Cottus gobio* (Teleostei: Cottidae) during the Pleistocene. *Journal of Evolutionary Biology*, **15**, 930–944.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Williams GC (1992) *Natural Selection: Domains, Levels, and Applications*. Oxford University Press, Oxford.

Ziuganov V (1983) Genetics of osteal plate polymorphism and microevolution of threespine stickleback (*Gasterosteus aculeatus* L.). *Theoretical and Applied Genetics*, **65**, 239–246.

## Appendix I

Number of individuals ($N$), expected (nonbiased) heterozygosity ($H_E$), observed heterozygosity ($H_O$) and number of alleles ($N_a$) for 10 polymorphic allozyme loci in seven samples of *Gasterosteus aculeatus*. For abbreviations see Table 1

| Locus | | OUD | YER | STJ | DOEL | HER | HEV | LOZ | Total |
|---|---|---|---|---|---|---|---|---|---|
| AAT* | $N$ | 49 | 50 | 50 | 50 | 49 | 50 | 50 | 348 |
| | $H_E$ | 0.0204 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0029 |
| | $H_O$ | 0.0204 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0029 |
| | $N_a$ | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| CK* | $N$ | 49 | 50 | 49 | 50 | 49 | 50 | 50 | 347 |
| | $H_E$ | 0.0000 | 0.0000 | 0.0204 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0029 |
| | $H_O$ | 0.0000 | 0.0000 | 0.0204 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0029 |
| | $N_a$ | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 |
| FH* | $N$ | 48 | 50 | 50 | 50 | 49 | 50 | 50 | 347 |
| | $H_E$ | 0.0414 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0058 |
| | $H_O$ | 0.0417 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0058 |
| | $N_a$ | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 3 |
| GPI-1* | $N$ | 49 | 45 | 50 | 49 | 49 | 50 | 50 | 342 |
| | $H_E$ | 0.0000 | 0.1820 | 0.0000 | 0.1340 | 0.0000 | 0.0200 | 0.0200 | 0.0513 |
| | $H_O$ | 0.0000 | 0.2000 | 0.0000 | 0.1429 | 0.0000 | 0.0200 | 0.0200 | 0.0526 |
| | $N_a$ | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 |
| IDH* | $N$ | 41 | 46 | 50 | 48 | 48 | 46 | 48 | 327 |
| | $H_E$ | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0638 | 0.0000 | 0.0091 |
| | $H_O$ | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0217 | 0.0000 | 0.0031 |
| | $N_a$ | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 |
| LDH* | $N$ | 49 | 50 | 50 | 50 | 49 | 50 | 50 | 348 |
| | $H_E$ | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0200 | 0.0000 | 0.0029 |
| | $H_O$ | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0200 | 0.0000 | 0.0029 |
| | $N_a$ | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 |
| MDH-1* | $N$ | 49 | 49 | 49 | 50 | 48 | 49 | 39 | 333 |
| | $H_E$ | 0.0000 | 0.0000 | 0.0600 | 0.0200 | 0.0000 | 0.0000 | 0.0000 | 0.0120 |
| | $H_O$ | 0.0000 | 0.0000 | 0.0612 | 0.0200 | 0.0000 | 0.0000 | 0.0000 | 0.0120 |
| | $N_a$ | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 2 |
| MDH-2* | $N$ | 44 | 49 | 49 | 45 | 47 | 50 | 39 | 323 |
| | $H_E$ | 0.3153 | 0.0000 | 0.2171 | 0.1998 | 0.0000 | 0.0000 | 0.0000 | 0.1160 |
| | $H_O$ | 0.2500 | 0.0000 | 0.1633 | 0.2222 | 0.0000 | 0.0000 | 0.0000 | 0.0898 |
| | $N_a$ | 2 | 1 | 2 | 2 | 1 | 1 | 1 | 4 |
| PGM-1* | $N$ | 49 | 50 | 50 | 50 | 49 | 50 | 50 | 348 |
| | $H_E$ | 0.0204 | 0.0960 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0171 |
| | $H_O$ | 0.0204 | 0.1000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0172 |
| | $N_a$ | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 3 |
| PGM-2* | $N$ | 45 | 50 | 29 | 49 | 44 | 43 | 50 | 310 |
| | $H_E$ | 0.3998 | 0.3578 | 0.4701 | 0.4563 | 0.5056 | 0.2430 | 0.4806 | 0.4375 |
| | $H_O$ | 0.3333 | 0.3800 | 0.5172 | 0.3469 | 0.5111 | 0.2326 | 0.4600 | 0.3903 |
| | $N_a$ | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 4 |

## Appendix II

Number of individuals ($N$), expected (nonbiased) heterozygosity ($H_E$), observed heterozygosity ($H_O$), number of alleles ($N_a$) and allele size range for five polymorphic microsatellite loci in seven samples of *Gasterosteus aculeatus*. For abbreviations see Table 1

| Locus | | OUD | YER | STJ | DOEL | HER | HEV | LOZ | Total |
|---|---|---|---|---|---|---|---|---|---|
| Gacμ4b | $N$ | 49 | 49 | 49 | 46 | 50 | 50 | 49 | 342 |
| | $H_E$ | 0.8414 | 0.8672 | 0.8197 | 0.8347 | 0.7133 | 0.5683 | 0.6960 | 0.8217 |
| | $H_O$ | 0.7959 | 0.8367 | 0.7959 | 0.8478 | 0.7400 | 0.6200 | 0.5918 | 0.7456 |
| | $N_a$ | 9 | 9 | 11 | 11 | 4 | 4 | 9 | 14 |
| | range | 114–134 | 114–134 | 114–136 | 114–140 | 120–132 | 122–132 | 102–138 | 102–140 |
| Gacμ7b | $N$ | 46 | 47 | 49 | 43 | 50 | 50 | 49 | 334 |
| | $H_E$ | 0.8294 | 0.8504 | 0.8290 | 0.8457 | 0.6832 | 0.6198 | 0.5754 | 0.8216 |
| | $H_O$ | 0.8478 | 0.7447 | 0.8776 | 0.7442 | 0.7000 | 0.5400 | 0.5510 | 0.7126 |
| | $N_a$ | 11 | 14 | 9 | 15 | 7 | 6 | 5 | 18 |
| | range | 95–141 | 95–139 | 109–139 | 95–141 | 113–139 | 121–135 | 107–135 | 95–141 |
| Gacμ9b | $N$ | 39 | 43 | 40 | 38 | 50 | 50 | 50 | 310 |
| | $H_E$ | 0.8735 | 0.9089 | 0.8908 | 0.8232 | 0.6006 | 0.6885 | 0.3077 | 0.9056 |
| | $H_O$ | 0.6410 | 0.5349 | 0.5750 | 0.5789 | 0.5400 | 0.6200 | 0.2200 | 0.5226 |
| | $N_a$ | 15 | 16 | 12 | 14 | 5 | 10 | 4 | 22 |
| | range | 166–204 | 168–208 | 164–200 | 168–204 | 168–200 | 170–200 | 188–198 | 164–208 |
| Gacμ10b | $N$ | 41 | 48 | 48 | 39 | 46 | 50 | 48 | 320 |
| | $H_E$ | 0.9572 | 0.8836 | 0.9544 | 0.9600 | 0.8516 | 0.9182 | 0.8785 | 0.9618 |
| | $H_O$ | 0.9024 | 0.9167 | 0.9583 | 0.9487 | 0.7609 | 0.9000 | 0.9375 | 0.9031 |
| | $N_a$ | 28 | 22 | 30 | 29 | 12 | 23 | 13 | 42 |
| | range | 192–272 | 190–258 | 182–270 | 194–260 | 198–242 | 204–266 | 216–264 | 182–272 |
| Gacμ14b | $N$ | 45 | 50 | 50 | 38 | 50 | 49 | 50 | 332 |
| | $H_E$ | 0.4851 | 0.4681 | 0.5610 | 0.5786 | 0.2481 | 0.0000 | 0.0000 | 0.4547 |
| | $H_O$ | 0.4000 | 0.4400 | 0.3600 | 0.5000 | 0.2800 | 0.0000 | 0.0000 | 0.2741 |
| | $N_a$ | 3 | 4 | 4 | 4 | 3 | 1 | 1 | 6 |
| | range | 109–113 | 99–113 | 109–115 | 107–113 | 109–113 | 109 | 109 | 99–115 |