

Signatures of selection in the three-spined stickleback along a small scale brackish water - freshwater transition zone

Journal:	<i>Ecology and Evolution</i>
Manuscript ID:	ECE-2015-03-00147.R1
Wiley - Manuscript type:	Original Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Konijnendijk, Nellie; KU Leuven, Laboratory of Aquatic Ecology Shikano, Takahito; University of Helsinki, Department of Biosciences Daneels, Dorien; KU Leuven, Laboratory of Aquatic Ecology Volckaert, Filip; KU Leuven, Laboratory of Aquatic Ecology Raeymaekers, Joost; KU Leuven, Laboratory of Aquatic Ecology
Search Terms:	Evolutionary Ecology, Molecular Ecology, Genomics
Abstract:	<p>Local adaptation is often obvious when gene flow is impeded, such as observed at large spatial scales and across strong ecological contrasts. However, it becomes less certain at small scales such as between adjacent populations or across weak ecological contrasts, when gene flow is strong. While studies on genomic adaptation tend to focus on the former, less is known about the genomic targets of natural selection in the latter situation. In this study we investigate genomic adaptation in populations of the three-spined stickleback <i>Gasterosteus aculeatus</i> L. across a small-scale ecological transition with salinities ranging from brackish to fresh.</p> <p>Adaptation to salinity has been repeatedly demonstrated in this species. A genome scan based on 87 microsatellite markers revealed only few signatures of selection, likely owing to the constraints that homogenizing gene flow puts on adaptive divergence. However, the detected loci appear repeatedly as targets of selection in similar studies of genomic adaptation in three-spined stickleback. We conclude that the signature of genomic selection in the face of strong gene flow is weak, yet detectable. We argue that the range of studies of genomic divergence should be extended to include more systems characterised by limited geographical and ecological isolation, which is often a realistic setting in nature.</p>

1 **Signatures of selection in the three-spined stickleback along a small scale brackish water -**
2 **freshwater transition zone**

3
4 Nellie Konijnendijk¹, Takahito Shikano², Dorien Daneels^{1,*}, Filip A.M. Volckaert¹, Joost A. M.
5 Raeymaekers¹

6
7 ¹ Laboratory of Biodiversity and Evolutionary Genomics, University of Leuven, Ch.
8 Deberiotstraat, 32, B-3000 Leuven, Belgium

9 ² Ecological Genetics Research Unit, Department of Biosciences, University of Helsinki, P.O.
10 Box 65, FI-000 14, Helsinki, Finland

11
12 * Current address: Centre for Medical Genetics, Reproduction and Genetics; Reproduction,
13 Genetics and Regenerative Medicine, Vrije Universiteit Brussel (VUB), Universitair ziekenhuis
14 Brussel (UZ Brussel), Laarbeeklaan 101, 1090 Brussels, Belgium.

15
16 **Running header:** Adaptation in three-spined stickleback

17
18 **Address for correspondence:**

19 Joost Raeymaekers, University of Leuven
20 Laboratory of Biodiversity and Evolutionary Genomics
21 Ch. Deberiotstraat, 32
22 B-3000 Leuven, Belgium
23 e-mail: joost.raeymaekers@bio.kuleuven.be
24 Phone : + 32 16 32 39 66, Fax : +32 16 32 45 75

25
26 **ABSTRACT**

27
28 Local adaptation is often obvious when gene flow is impeded, such as observed at large
29 spatial scales and across strong ecological contrasts. However, it becomes less certain at
30 small scales such as between adjacent populations or across weak ecological contrasts, when
31 gene flow is strong. While studies on genomic adaptation tend to focus on the former, less is
32 known about the genomic targets of natural selection in the latter situation. In this study we
33 investigate genomic adaptation in populations of the three-spined stickleback *Gasterosteus*
34 *aculeatus* L. across a small-scale ecological transition with salinities ranging from brackish to
35 fresh. Adaptation to salinity has been repeatedly demonstrated in this species. A genome
36 scan based on 87 microsatellite markers revealed only few signatures of selection, likely
37 owing to the constraints that homogenizing gene flow puts on adaptive divergence.
38 However, the detected loci appear repeatedly as targets of selection in similar studies of
39 genomic adaptation in three-spined stickleback. We conclude that the signature of genomic
40 selection in the face of strong gene flow is weak, yet detectable. We argue that the range of
41 studies of genomic divergence should be extended to include more systems characterised by
42 limited geographical and ecological isolation, which is often a realistic setting in nature.

43
44 **Keywords:** adaptation, candidate gene, evolution, fish, genome scan, population genomics,
45 selection

46 **INTRODUCTION**

47

48 Recent questions on the mechanisms of evolutionary biology revolve around the genomic
49 architecture of species and the effects that processes such as selection, drift, mutation and
50 gene flow have on the genome. Changes in the environment can now be linked to the
51 genetic signatures of these processes and hence promote the understanding of the genetic
52 basis of ecological adaptation. This allows us to understand the various components of the
53 mechanism of adaptation such as which genes are involved, how they are distributed in the
54 genome, and how often the same genes lead to certain adaptations both in the laboratory
55 (Becks *et al.* 2012; see Barrick & Lenski 2013 for other examples) and nature (Mitchell-Olds &
56 Schmitt 2006; Yang *et al.* 2011; Jones *et al.* 2012a; Orsini *et al.* 2012). Additionally we want
57 to understand which processes are essential for adaptation to succeed. One way to find
58 genes that are important for the process of adaptation is to identify genetic signatures of
59 adaptation in natural populations. These "genome scans" identify genes that are under
60 selection across contrasting environments and might thus be involved in the transition from
61 one ecological extreme to the other. When interpreted with care (de Villemereuil *et al.*
62 2014), this approach has shown to be a promising way to find links between genotype,
63 phenotype and fitness in natural populations (Storz 2005).

64

65 Genome scans have so far generated two major insights. First, selection may act on many
66 parts of the genome (Nosil 2009; Hohenlohe *et al.* 2010; Feder *et al.* 2012; Jones *et al.*
67 2012a; Jones *et al.* 2012b; Strasburg *et al.* 2012; Arnegard *et al.* 2014; Seehausen *et al.*
68 2014). Nosil (2009) estimated that 5-10 % of the genome is affected by natural selection.
69 Based on a genome-wide analysis of selection, Hohenlohe *et al.* (2012) showed that the
70 genome is much more structured and dynamic than expected from theory. More studies
71 have now contributed to these topics and show that the areas affected by natural selection
72 are typically patchily distributed ("hotspots") across the genome (Voight *et al.* 2006; Papa *et al.*
73 2008; Hohenlohe *et al.* 2010). Genomic inversions, deletions, repetitive elements and
74 other structural changes facilitate the rise of these islands of divergence, causing a physical
75 barrier to crossing over and thus adding to linkage disequilibrium (Hohenlohe *et al.* 2010).
76 On the other hand, most cases of adaptation in plants involve many genomic areas with a
77 wide distribution across the genome rather than clustered hotspots (Strasburg *et al.* 2012).

78

79 A second common finding of genome scans is that organisms may show parallel as well as
80 non-parallel genetic responses to environmental change. For example, Atlantic salmon
81 (*Salmo salar*) populations that undergo a parallel change in environment show non-parallel
82 adaptive divergence at the genomic level (Perrier *et al.* 2013). In cases of parallel phenotypic
83 evolution in the three-spined stickleback (*Gasterosteus aculeatus* L.), common (Hohenlohe
84 *et al.* 2010; Shimada *et al.* 2011; Jones *et al.* 2012b) as well as unique (Mäkinen *et al.* 2008;
85 Jones *et al.* 2012b; Roesti *et al.* 2012) genomic regions are targeted by selection. This has
86 been confirmed in experiments with stick insects (*Timema cristinae*) (Soria-Carrasco *et al.*
87 2014) and in genome scans of lake white fish (*Coregonus* sp.) (Bernatchez *et al.* 2010). In
88 contrast, parallel phenotypic evolution in other species, such as the nine-spined stickleback
89 (*Pungitius pungitius*) (Shikano *et al.* 2010a) or the rough periwinkle *Littorina saxatilis* (Butlin
90 *et al.* 2014), were mainly characterized by non-parallel genomic signatures of selection.

91

92 Adaptive divergence is thought to proceed as a balance between divergent selection and
93 homogenizing gene flow (Levene 1953; Hagen 1967; Endler 1973; Bell 1982) and hence may
94 reach various stages. This has long been recognised at the phenotypic level (Tregenza 2002;
95 Moore *et al.* 2007; Hendry 2009; Schluter 2009). More recently, theoretical and empirical
96 studies have improved our understanding of the genomic architecture at various stages of
97 adaptation as well (Pinho & Hey 2010; Yeaman & Otto 2011; Yeaman & Whitlock 2011).
98 Nevertheless, the focus of genome scans is often on systems where it is reasonable to
99 assume that population divergence has a strong adaptive component. To do so, genome
100 scans often target populations at large spatial scales and across strong ecological contrasts
101 (Hohenlohe *et al.* 2010; Poncet *et al.* 2010; Stapley *et al.* 2010; Zulliger *et al.* 2013). While
102 this has generated great insight in the genomic basis of adaptation and speciation,
103 knowledge might be biased towards stages where gene flow may be largely impeded, and
104 where adaptation is already largely achieved. Gene flow modifies the response to selection
105 by modulating the distribution of the genes that underlie ecologically relevant traits (Slatkin
106 1987). The study of adaptation with gene flow, especially in study systems where genetic
107 divergence is far from complete such as across a small-scale ecological transitions or
108 between highly connected populations, is therefore crucial to understand how populations
109 diverge to different ecological optima (Hansen *et al.* 2002; Storz 2005; Nielsen *et al.* 2009;
110 Coscia *et al.* 2011; DeFaveri *et al.* 2013; Vandamme *et al.* 2014).

111
112 The three-spined stickleback represents an excellent model for the study of adaptive
113 divergence, as phenotypic responses to several ecological changes are frequent and well-
114 documented (McPhail 1994; Foster *et al.* 1998; McKinnon & Rundle 2002; Boughman 2007).
115 A diverse number of populations can be found that occupy various stages of divergence from
116 panmixia to complete and irreversible reproductive isolation (Hendry *et al.* 2009). This
117 provides an excellent framework to investigate population divergence at the phenotypic and
118 genomic level. The genome scans that have been applied to three-spined stickleback have
119 contrasted marine-freshwater (Mäkinen *et al.* 2008; Hohenlohe *et al.* 2010; DeFaveri *et al.*
120 2011; Shimada *et al.* 2011; Jones *et al.* 2012a; Jones *et al.* 2012b; DeFaveri & Merilä 2013),
121 lake-stream (Deagle *et al.* 2012; Roesti *et al.* 2012), and benthic-limnetic (Olafsdottir &
122 Snorrason 2009; Jones *et al.* 2012a; Lucek *et al.* 2014) population pairs, as well as
123 populations from clean versus polluted water (Lind & Grahn 2011). A common finding is that
124 several genes or gene regions are repeatedly selected across populations and locations,
125 although population-specific regions do appear as well (Hohenlohe *et al.* 2010; DeFaveri *et al.*
126 2011; Shimada *et al.* 2011; Jones *et al.* 2012b). Others have found that most regions
127 under selection were highly specific to the location under study (Mäkinen *et al.* 2008; Deagle
128 *et al.* 2012; Roesti *et al.* 2012).

129
130 In this study we investigate genomic adaptation in three-spined stickleback populations from
131 the Belgian-Dutch lowlands. Populations in this area differ in various morphological traits,
132 which are often correlated with salinity and distance to the coast (Heuts 1947; Raeymaekers
133 *et al.* 2005; Raeymaekers *et al.* 2007; Van Dongen *et al.* 2009; Raeymaekers *et al.* 2012). At
134 the same time, gene flow between these populations is moderate to strong due to high
135 connectivity (Raeymaekers *et al.* 2014). This has the advantage that we can study adaptation
136 with ongoing gene flow, which is often a realistic setting in nature. All sites are part of an
137 interconnected landscape of canals and streams along an ecological transition (i.e. discrete
138 habitats with salinities ranging between brackish and freshwater) at a much smaller

139 geographical scale than other genome scan studies across salinity gradients in three-spined
140 stickleback (Table 1).

141

142 In order to explore the adaptive changes across this transition, we screened two populations
143 from each end of the transition for genomic signatures of selection. We hypothesized that (i)
144 the outcome of small-scale adaptation along ecological transitions may vary at the genomic
145 level and that (ii) adaptation at the genomic level may be influenced by gene flow.
146 Populations were screened for 87 microsatellite markers of which 41 are linked to genes
147 with a range of ecologically relevant functions (Shikano *et al.* 2010b; Shimada *et al.* 2011).
148 We expect that the relatively high gene flow among these populations might constrain local
149 adaptation, despite obvious differences in phenotype across these extremes. We determine
150 outlier loci and compare the results to previous studies in three-spined stickleback across
151 similar ecological contrasts.

152 MATERIAL AND METHODS

153 *Study area*

154

155 Three-spined stickleback (*Gasterosteus aculeatus* L.; Gasterosteidae) from the coastal
156 lowlands (polder) of Belgium and the Netherlands (Figure 1) reside in ponds, ditches,
157 streams, estuaries or polder creeks. They have an anadromous or landlocked life style (Heuts
158 1947; Wootton 1976; Raeymaekers *et al.* 2005; Raeymaekers *et al.* 2007). The polder and
159 surrounding areas contain diked brackish and freshwater habitats of Holocene origin with
160 varying levels of connectivity to adjacent estuaries and the open sea. Populations which live
161 in close proximity to the sea (< 10 km) reside in brackish water of which the salinity is
162 influenced by rainfall and water management. On a scale of less than 50 km further inland,
163 salinity drops to freshwater levels (Raeymaekers *et al.* 2014). Lateral plate number, an
164 important ecological trait, also decreases with distance to the coast, with population
165 averages from 20 to 5 (Heuts 1947; Raeymaekers *et al.* 2014). Populations bordering the
166 North Sea and the Baltic are typically polymorphic for lateral plate number (Heuts 1947;
167 Raeymaekers *et al.* 2014), so higher or lower population averages are rare. This range is
168 therefore representative for the phenotypic extremes we can find in this part of the
169 stickleback's distribution range.

170 *Field sampling*

171

172 Field sampling was conducted in spring 2009 in parallel with a multi-year study by
173 Raeymaekers *et al.* (2014), describing the distribution of lateral plate number in the study
174 area. Two brackish water creeks (L01 and L02) and two freshwater ditches (L12 and U01),
175 representing the sites with the most extreme values for salinity, were selected (Table 2).
176 Thirty fish from each population were collected with dip nets, immediately anaesthetized
177 with MS222 (2 g.l⁻¹) and frozen on dry ice. In the lab, sticklebacks were thawed on ice,
178 measured (standard length (SL); ± 0.1 cm), weighed (± 0.01 g), photographed and fin clipped
179 for a DNA sample.

180 *Morphological measurements*

181

182 In line with previous studies on phenotypic divergence between the stickleback populations
183 of the Belgian-Dutch lowlands (Heuts 1947; Raeymaekers *et al.* 2005; Raeymaekers *et al.*

184 2007; Van Dongen *et al.* 2009; Raeymaekers *et al.* 2012) we investigated variation at six
185 morphological traits. The left side of each fish was photographed from a standard angle, and
186 a ruler was placed in each photograph for scaling. Dorsal spine length, pelvic spine length
187 and pelvic plate length were measured from pictures using the software tpsDig 1.37 (Rohlf
188 2002). A subsample of the fish was rinsed with water for 72 h, bleached for 4 h (1% KOH
189 bleach solution) and stained with alizarin red (Taylor & Dyke 1985). Stained fish were used to
190 determine the number of lateral plates on the right side of the fish. The right part of the gills
191 was then dissected and the number and length of the large gill rakers were quantified under
192 a dissection scope.

193 **Marker selection**

194

195 A set of 110 microsatellite markers was selected, including a range of putatively neutral
196 markers, to set a proper neutral F_{ST} background. In our study area there is a salinity cline, but
197 various other factors may co-vary with this cline. Therefore we included 41 markers that are
198 known to be linked to functional genes in a range of ecologically relevant functions such as
199 salinity, growth and immunity (Shikano *et al.* 2010b; Shimada *et al.* 2011). Several of these
200 genes have been found to be under selection in the three-spined stickleback or nine-spined
201 stickleback of other salinity transitions (Hohenlohe *et al.* 2010; Shikano *et al.* 2010a;
202 DeFaveri *et al.* 2011; Shimada *et al.* 2011; Jones *et al.* 2012a; Jones *et al.* 2012b). This allows
203 us to compare these systems to some extent with ours and pinpoint parallel changes across
204 systems. The marker closely linked to the *Eda* gene (*Stn380*), a major effect gene underlying
205 variation in plate number (Colosimo *et al.*, 2005), was included as a reference gene that is
206 often under selection in freshwater-salt water comparisons (Raeymaekers *et al.* 2007;
207 Mäkinen *et al.* 2008; DeFaveri *et al.* 2011; DeFaveri & Merilä 2013).

208 **DNA extraction and genotyping**

209

210 Genomic DNA was extracted from fin clips using a proteinase K digestion step and the
211 Nucleospin 96 Tissue DNA Extraction kit (Macherey-Nagel). Individuals were genotyped at
212 110 microsatellite loci that were arranged in 21 multiplexes of 4 to 8 markers at a time, with
213 EST based markers and markers within or near genes with relevant functions included (Table
214 S1). Loci were amplified with the Qiagen® Multiplex PCR Kit (Qiagen, Venlo, the
215 Netherlands). The 10 μ l PCR cocktail contained 1-100 ng genomic DNA, 2 pmol each of
216 forward and reverse primers, 1x Qiagen Multiplex PCR master mix, 0.5x Q-solution, and
217 RNase-free water. The reaction consisted of an initial activation step of 15 min at 95 °C,
218 followed by 30 cycles of 30 s at 94 °C, 90 s at 53 °C and 60 s at 72 °C. A final elongation step
219 of 5 min at 60 °C was performed. Allele sizes were determined by means of an internal ET
220 ROX 550 size standard (Amersham Biosciences). Polymerase chain products were visualized
221 using a MegaBace 1000 automated sequencer (Amersham Biosciences). Alleles were scored
222 with the Fragment Profiler v1.2 software (Amersham Biosciences), using visual scoring and
223 manual corrections. Marker *Stn380*, linked to the *Eda* gene, was scored separately to
224 determine the frequency of the “low plated” allele in each population (Table 2).

225 **Data analysis**

226

227 *Phenotypic differentiation*

228 For each trait, analysis of variance (ANOVA) was performed to test for statistical differences
229 between the four populations. For traits which depend on size, standard length was included
230 as a covariate.

231

232 *Genetic diversity and genetic differentiation*

233

234 Genotypes were checked for scoring errors attributable to stutter-products, large allele
235 dropout or null-alleles, using MICRO-CHECKER v2.3 (van Oosterhout *et al.* 2004). Estimates of
236 allelic richness, genetic diversity (H_e , H_o), global and pairwise F_{ST} , with a significance
237 calculated with 1000 bootstraps over loci, were calculated using the GENETIX v4.05.02
238 software (Belkhir *et al.* 1996).

239

240 *Genomic signatures of selection*

241

242 We conducted global outlier tests to find outliers across all populations and used pairwise
243 comparisons to check whether the outliers found by global tests were due to habitat
244 differences. Loci that are under directional selection are expected to have lower
245 intrapopulation variability and larger interpopulation variability than neutral loci. Loci under
246 directional selection can thus be traced by patterns in heterozygosity, differences in F_{ST}
247 values or a combination of the two. Evaluation of several outlier detection methods has
248 shown that these methods differ in number of false positives and false negatives (Narum &
249 Hess 2011). Four methods were therefore compared: LOSITAN (Antao *et al.* 2008), the
250 outlier detection method implemented in ARLEQUIN v3.5.2.3 using hierarchical clustering
251 (Excoffier & Lischer 2010), BAYESCAN v2.01 (Foll & Gaggiotti 2008), and LnRH (Kauer *et al.*
252 2003). The first three methods were used to determine global outliers. Additionally we did
253 pairwise comparisons of all populations using LOSITAN and LnRH to detect specific
254 signatures of selection in freshwater-brackish water, freshwater-freshwater and brackish
255 water-brackish water populations pairs.

256

257 The four methods are based on different underlying assumptions. LOSITAN is based on an
258 island model that uses a coalescent F_{ST} -outlier method based on the distribution of F_{ST} as a
259 function of the heterozygosity. We used the function that first establishes a neutral F_{ST}
260 baseline by removing putative markers under selection outside the 95% interval with 10^5
261 simulations. The infinite allele model was used with a 95% and 99% confidence interval. We
262 ran 10^5 simulations as recommended by Antao (2008). The outlier detection software
263 implemented in ARLEQUIN uses the same island model, but adds on the option for
264 hierarchical clustering. In the presence of strong hierarchical population structure it reduces
265 false positives by a hierarchical analysis of genetic differentiation (Excoffier *et al.* 2009). We
266 clustered populations according to two scenarios: in two groups (brackish water populations
267 versus freshwater populations), and three groups (brackish water populations versus each
268 freshwater population separately). The first scenario simulates a common descent of the
269 two freshwater populations from the marine population, while the second one simulates a
270 separate split of the two freshwater populations from the marine population. We used the
271 standard settings of 20,000 simulations for each run and 100 demes per group. The method
272 executed by BAYESCAN uses a logistic regression model which explains the observed pattern
273 of diversity by dividing it in a locus- and a population-specific component (Beaumont and
274 Balding 2004). One benefit of this method is that it allows for different migration rates and

275 different effect sizes and thus can be used for scenarios that deviate from the island model.
276 We conducted 10 pilot runs of 5,000 iterations, followed by an additional 150,000 iterations
277 and a burn-in of 50,000 iterations. Outliers were appointed based on 90%, 95% and 99%
278 posterior probabilities. Finally, the LnRH method is designed especially for microsatellite
279 markers and determines the reduction of heterozygosity. This method is based on the
280 assumption that microsatellites linked to a gene under selection will show reduced levels of
281 diversity between two populations. After standardization of the LnRH estimates with a mean
282 of zero and a standard deviation of one, we determined outliers at 95% and 99% level. We
283 subtracted outliers that were found in pairwise comparisons within the same habitat type,
284 to limit the number of false positives.

285

286 All tests based on simulations were executed three times to test for robustness of the
287 results. Only outliers that were detected at least twice with each method were scored as a
288 putative outlier. We detected outlier loci that are under balancing and directional selection.
289 However, interpreting loci under balancing selection is difficult, as there are still limitations
290 for the identification of loci under balancing selection (Hansen *et al.* 2010; Narum & Hess
291 2011). Therefore we only discuss the loci under directional or positive selection. Loci under
292 balancing selection are provided in Supplementary Table S2.

293

294 **RESULTS**

295

296 ***Phenotypic differentiation***

297

298 Statistical analysis of variance (ANOVA) confirmed the presence of significant differences
299 between the four populations for standard length, lateral plate number, pelvic plate and
300 spine length, and large gill raker length (Figure 2). In particular, the brackish water
301 populations (L01 and L02) were larger in size than the freshwater populations (L12 and U01;
302 $F_{3,155} = 75.6$; $P < 0.0001$), and had more lateral plates ($F_{3,147} = 10.4$; $P < 0.0001$), longer pelvic
303 spines (especially population L01; $F_{3,93} = 5.51$; $P = 0.0016$), longer pelvic plates ($F_{3,93} = 8.60$; P
304 < 0.0001), and larger gill rakers ($F_{3,93} = 19.8$; $P < 0.0001$).

305

306 ***Genetic diversity and genetic differentiation***

307

308 After genotyping of the individuals with 110 microsatellite markers, we selected 87 markers
309 with good amplification quality to perform the data analysis. A total of 1116 alleles were
310 observed in four populations of 26 to 28 individuals each, with an average of 13 alleles per
311 locus and a range from 2 to 34. Observed heterozygosity ranged from 0.022 to 0.95 across
312 loci. Expected heterozygosity ranged from 0.05 to 0.92 across loci and from 0.57 to 0.69
313 across populations (Table 2). Allelic richness among populations varied from 6.51 to 9.47,
314 with the freshwater populations being less diverse than the brackish water populations.
315 Eight loci were possibly affected by null alleles or stutters. We therefore performed all final
316 analyses with and without these loci and specify when results differ. Differentiation among
317 populations was moderate with the global F_{ST} value being 0.059. Pairwise F_{ST} values were
318 significant, except between the two brackish water populations (Table 3).

319

320 ***Genomic signatures of selection***

321

322 Six of the 87 loci were assigned as outliers in at least one of the methods, with the different
323 methods identifying 4, 2, 2, and 4 outliers for Lositan, Bayescan, InRH and Arlequin
324 respectively (Table 4). The methodologically similar tests performed by Lositan and Arlequin
325 with hierarchical clustering resulted in the same set of outliers. In contrast, only two outliers
326 were shared across methodologically different outlier detection methods; our conservative
327 measure of outliers hence amounts to 2.3 % of the total number of loci. This value is low in
328 comparison with other genome scan studies of three-spined stickleback (Table 1).

329

330 The two outliers for directional selection that appear consistently across methodologically
331 different tests are marker *Stn46* (identified by Arlequin/Lositan and Bayescan) and *Ppgm44*
332 (identified by InRH and Bayescan). *Stn46* has been previously associated with the Rho
333 guanine nucleotide exchange factor 9 (Table S1). It was also assigned as an outlier in one
334 pairwise freshwater versus brackish water comparison. *Ppgm44* is a marker that is linked to
335 the gene *myostatin2*, which is associated with growth (Table S1). It was also assigned as an
336 outlier in all four pairwise freshwater versus brackish water comparisons. Among the four
337 other outliers that were detected (Table 4), one is associated with osmoregulation, namely
338 an alpha subunit of the Na⁺/K⁺ ATPase (ATP1A1). Other outliers are either linked to
339 functions such as thermal response (HSPA14) or were assumed to be neutral. Locus *Stn34*
340 was also assigned as an outlier, but since null alleles were detected at this locus, it was
341 excluded from Table 4. The *Eda* gene, with frequencies of the low-plated allele ranging
342 between 0.44-0.51 in the brackish populations and 0.74-0.93 in the freshwater populations
343 (Table 2), did not show up as an outlier in the analyses (Table 4).

344

345 DISCUSSION

346

347 We found a limited number of loci showing a signature of selection among three-spined
348 stickleback populations inhabiting the coastal Dutch-Belgian lowlands. Six outliers were
349 found when all outlier detection methods were considered and only two of those were
350 shared among methodologically different tests, despite differentiation among populations in
351 ecology and phenotype. Not only the number but also the proportion of outliers was low as
352 compared to other studies (Table 1). The function of one of the outlier loci could be directly
353 related to salinity, the most obvious ecological gradient in our study area. We here discuss
354 possible explanations for the limited amount of outliers, the putative function of the genes
355 linked to outliers, and the consequences for adaptive divergence in the face of gene flow.

356

357 The presence of a limited number and proportion of outliers might be attributed to a
358 number of technical aspects. First, we used relatively few markers. However, among the 87
359 markers selected, 41 markers are linked to ecologically relevant functions, of which several
360 have been shown to be under selection in systems with similar salinity clines (Mäkinen *et al.*
361 2008; Shimada *et al.* 2011; DeFaveri *et al.* 2013). Second, analysing relatively few
362 populations may have two drawbacks: 1) it might increase the risk of false negatives due to
363 the limited number of individuals sampled, and 2) it might increase the difficulty in
364 separating the signal of selection from the geographical and historical signal (Bierne *et al.*
365 2011; Bierne *et al.* 2013). The small overlap in outlier loci among detection methods might
366 be a symptom of this. However, in a larger study across 14 populations (2320 individuals)
367 describing the distribution of lateral plate number and the underlying *Eda* gene in our study
368 area (Raeymaekers *et al.* 2014), the signature of selection at the *Eda* gene was neither

369 significant. This bi-allelic gene is experiencing selection across various other salinity
370 transitions in three-spined stickleback (Raeymaekers *et al.* 2007; Mäkinen *et al.* 2008;
371 DeFaveri *et al.* 2011; DeFaveri & Merilä 2013). This suggests that increasing sample size does
372 not necessarily enhance the detectability of selection. We therefore do not expect that
373 technical issues are a major explanation for our findings.

374

375 An alternative explanation for the low proportion of outlier loci is that selection might be
376 weak, owing to weaker environmental contrasts in our study area (i.e., brackish to
377 freshwater) compared to other studies (Table 1). For instance, Raeymaekers *et al.* (2014)
378 found that shifts in *Eda* allele frequencies from one generation to the next were associated
379 with salinity, but that at the landscape level salinity did not correlate with *Eda* allele
380 frequencies. This suggests that selection is still acting, but might be too weak to contribute
381 to local adaptation and leave a signature of selection at this gene. In addition, strong gene
382 flow might confound the effect of selection by mixing adapted and non-adapted alleles in
383 the respective populations. Gene flow being moderate to high, we expect this to be another
384 explanation for why we find so few consistent outlier loci compared to other studies.
385 Accordingly, Raeymaekers *et al.* (2014) found that the spatial distribution of the *Eda* allele
386 frequency correlated with distance to the coast, a proxy for population connectivity.
387 Recurrent contact between freshwater and estuarine or marine populations might lead to
388 the exchange of maladaptive alleles, but at the same time it has been argued that gene flow
389 might cause an opposite effect: efficient flow of advantageous loci (Schluter & Conte 2009;
390 Hohenlohe *et al.* 2012; Bell & Aguirre 2013; Raeymaekers *et al.* 2014), thus facilitating
391 adaptation from standing variation. Whether gene flow in this case fuels or constrains
392 adaptation is hard to say, but the signal of selection might be more difficult to pick up in
393 genome scans due to mixing of the genomic background. We therefore expect that the low
394 number of outlier loci is rather explained by either weak selection, confounding effects of
395 gene flow or a combination of both and are thus inherent to the system we study.

396

397 Genes that have been found in previous studies to be under selection relate to biological
398 functions such as bone formation, osmoregulation, growth, thermal response, maturation,
399 pigmentation, scent detection, spiggin production and morphology (Hohenlohe *et al.* 2010;
400 DeFaveri *et al.* 2011; Shimada *et al.* 2011; Jones *et al.* 2012b). Genes that appeared as
401 outliers in this study include markers linked to osmoregulation (*ATP1A1*), thermal response
402 (*HSPA14*) and growth (*myostatin2*). The marker linked to *ATP1A1* has been identified as
403 outlier gene in multiple stickleback saltwater-freshwater transitions worldwide (Jones *et al.*
404 2006; McCairns & Bernatchez 2010; DeFaveri *et al.* 2011; Shimada *et al.* 2011; Jones *et al.*
405 2012a; Jones *et al.* 2012b). A major outlier in our study was marker *Stn46*, which is linked to
406 the gene for Rho guanine nucleotide exchange factor 9. It is a member of the gene family
407 coding for rho proteins, a subfamily of the guanine nucleotide exchange factors (GEFs),
408 which are multi-domain proteins involved in the activation of small GTPases (Rossman *et al.*
409 2005). The Rho family is involved in relaying signals from cell-surface receptors to the actin
410 cytoskeleton and elsewhere (Dvorsky & Ahmadian 2004). Its function can be associated with
411 juvenile growth; in zebrafish it has been specifically linked to angiogenesis (Garnaas *et al.*
412 2008) and striated muscle and neural development (Raeker *et al.* 2010). Interestingly, locus
413 *Stn46* was also under selection in populations of the nine-spined stickleback (Shikano *et al.*
414 2010a), indicating that the gene might be involved in local adaptation in multiple species.
415 Another major outlier in our study was the marker linked to *myostatin2*. This gene, a

416 member of the transforming growth factor-beta (TGF-beta) family, is known to function as a
417 negative regulator of skeletal muscle development and growth in mammals (Walsh &
418 Celeste 2005) and teleost fish (Radaelli *et al.* 2003). Blocking the expression of myostatin in
419 zebrafish has led to the development of a giant phenotype (Acosta *et al.* 2005), but
420 myostatin is produced in many other tissues than skeletal muscles and is expected to
421 influence many more functions (Radaelli *et al.* 2003). Understanding why these specific
422 genes are selected requires further study due to the broad range of functions these genes
423 might have.

424

425 The low number of outlier loci that we find contrasts with the differentiation across
426 populations in several morphological traits. Local differentiation in the number of lateral
427 plates, for instance, has been shown to significantly exceed the level of neutral
428 differentiation in our study area (Raeymaekers *et al.* 2014). Plate number and other
429 morphological traits such as spine length and gill raker length have reasonably high
430 heritability values exceeding 40 % (Schluter 1996; Peichel *et al.* 2001; Berner *et al.* 2014),
431 suggesting that phenotypes are largely determined by genetic rather than by plastic effects.
432 Yet, theory predicts that only functional loci with a relatively large effect size under strong
433 divergent selection will be able to surpass gene flow (Via 2009; Yeaman & Whitlock 2011).
434 The discrepancy between phenotypic and genomic signatures of selection might therefore
435 become particularly strong for traits that involve many genes of small effect. Arnegard *et al.*
436 (2014) have another explanation for the relatively large phenotypic divergence. They show
437 that niche differentiation in sticklebacks, even in early stages of differentiation, can involve
438 many different genes and that gene flow between divergent niche-adapted populations has
439 a bigger impact on the phenotype than just the traits that are directly targeted by selection.
440 It may be attributed to incompatibilities in hybrids that harbor a mix of genes of
441 differentially adapted genotypes. These effects imply that phenotypic changes might not
442 always be adaptive. McCairns & Bernatchez (2010) found indications that freshwater
443 populations might suffer from a loss of plasticity and it might be that epigenetic effects
444 further enhance this discrepancy. For instance, Chaturvedi *et al.* (2014) found that regulation
445 by miRNAs might make a significant contribution to freshwater adaptation in stickleback
446 populations.

447

448 **Conclusion**

449

450 We find that weak selection, high levels of gene flow or a combination of both can limit the
451 number of outliers in genome scans. Although genome scans targeting populations across
452 strong environmental contrasts are possibly more effective for pinpointing genes that are
453 involved in adaptation, the genes identified by these studies do not necessarily play an
454 important role at every stage of divergence. Many of the genes involved in salt-freshwater
455 transitions might be site-specific or might not be involved when gene flow is constantly
456 mixing the gene pool. The genes that we do find are likely to be those with a major effect
457 size and thus an underrepresentation of the total number of genes involved. In addition,
458 phenotypic adaptation is not necessarily genetic, but might be facilitated by plastic and
459 epigenetic effects. It remains a challenging task to find which genes and how many are truly
460 involved in local adaptation. We here showed that even with ample gene flow and across
461 weak ecological contrasts, interesting insights on the repeatability of genomic signatures of
462 selection can be obtained.

463

464 **ACKNOWLEDGEMENTS**

465

466 We would like to thank A. Chaturvedi, S. Geldof, B. Hellemans, L. Jaquemin, S. Ruyts, N.
467 Thiercelin, I. Verdonck and D. Verheyen for field support and technical assistance. We thank
468 I. Coscia and J. Merilä for insightful and inspiring comments and the latter for collaboration
469 on the whole project. We appreciated the insightful suggestions of two anonymous
470 reviewers. Research was sponsored by the Research Foundation - Flanders (post-doctoral
471 fellowship to J.A.M.R. and research project G.0142.03), by the KU Leuven (GOA/2006/06
472 “Evolutionary responses to strong selection pressures and their ecological consequences”;
473 PF/2010/07 “Eco- and socio-evolutionary dynamics”) and facilitated by the research
474 community W0.037.10N “Eco-evolutionary dynamics in natural and anthropogenic
475 communities” sponsored by the Research Foundation - Flanders. The study received
476 additional support from Academy of Finland grant (129662, 134728 and 218343) to Juha
477 Merilä.

478

479 **REFERENCES**

480

- 481 Acosta J, Carpio Y, Borroto I, González O, Estrada MP (2005) Myostatin gene silenced by RNAi
482 show a zebrafish giant phenotype. *Journal of Biotechnology* **119**, 324-331.
- 483 Antao T, Lopes A, Lopes R, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect
484 molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics* **9**, 323.
- 485 Arnegard, M. E., McGee, M. D., Matthews, B., Marchinko, K. B., Conte, G. L., Kabir, S.,
486 Bedford, N., Bergek, S., Chan, Y. F., Jones, F. C., Kingsley, D. M., Peichel, C. L. &
487 Schluter, D. 2014. Genetics of ecological divergence during speciation. *Nature* **511**:
488 307-311.
- 489 Barrick JE, Lenski RE (2013) Genome dynamics during experimental evolution. *Nature*
490 *Reviews: Genetics* **14**, 827-839.
- 491 Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations
492 from genome scans. *Molecular Ecology* **13**, 969-980.
- 493 Becks L, Ellner SP, Jones LE, Hairston NG (2012) The functional genomics of an eco-
494 evolutionary feedback loop: linking gene expression, trait evolution, and community
495 dynamics. *Ecology Letters* **15**, 492-501.
- 496 Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996) GENETIX 4.05, logiciel sous
497 Windows TM pour la génétique des populations, pp. 1996-2004. Laboratoire
498 génome, populations, interactions, CNRS UMR.
- 499 Bell MA (1982) Differentiation of adjacent stream populations of threespine sticklebacks.
500 *Evolution* **36**, 189-199.
- 501 Bell MA, Aguirre WE (2013) Contemporary evolution, allelic recycling, and adaptive radiation
502 of the threespine stickleback. *Evolutionary Ecology Research* **15**, 377-411.
- 503 Bernatchez L, Renaut S, Whiteley AR, et al. (2010) On the origin of species: insights from the
504 ecological genomics of lake whitefish. *Philosophical Transactions of the Royal Society*
505 *of London, Series B: Biological Sciences* **365**, 1783-1800.
- 506 Berner D, Moser D, Roesti M, Buescher H, Salzburger W (2014) Genetic architecture of
507 skeletal evolution in european lake and stream stickleback. *Evolution* **68**, 1792-1805.

- 508 Bierne N, Roze D, Welch JJ (2013) Pervasive selection or is it...? why are FST outliers
509 sometimes so frequent? *Molecular Ecology* **22**, 2061-2064.
- 510 Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why
511 genome scans may fail to map local adaptation genes. *Molecular Ecology* **20**, 2044-
512 2072.
- 513 Boughman JW (2007) Speciation in sticklebacks. In: *Biology of the Three-Spined Stickleback*
514 (eds. Östlund-Nilsson S, Mayer I, Huntingford F), pp. 83–126. CRC Press., Boca Raton,
515 FL.
- 516 Butlin RK, Saura M, Charrier G, *et al.* (2014) Parallel evolution of local adaptation and
517 reproductive isolation in the face of gene flow. *Evolution* **68**, 935-949.
- 518 Chaturvedi A, Raeymaekers JAM, Volckaert FAM (2014) Computational identification of
519 miRNAs, their targets and functions in three-spined stickleback (*Gasterosteus*
520 *aculeatus*). *Molecular Ecology Resources* **14**, 768-777.
- 521 Colosimo PF, Hosemann KE, Balabhadra S, *et al.* (2005) Widespread parallel evolution in
522 sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**, 1928-1933.
- 523 Coscia I, Vogiatzi E, Kotoulas G, Tsigenopoulos C, Mariani S (2011) Exploring neutral and
524 adaptive processes in expanding populations of gilthead sea bream, *Sparus aurata* L.,
525 in the North-East Atlantic. *Heredity* **108**, 537-546.
- 526 de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan methods
527 against more complex models: when and how much should we trust them?
528 *Molecular Ecology* **23**, 2006-2019.
- 529 Deagle BE, Jones FC, Chan YF, *et al.* (2012) Population genomics of parallel phenotypic
530 evolution in stickleback across stream-lake ecological transitions. *Proceedings of the*
531 *Royal Society of London, Series B: Biological Sciences* **279**, 1277-1286.
- 532 DeFaveri J, Jonsson PR, Merilä J (2013) Heterogeneous genomic differentiation in marine
533 threespine sticklebacks: Adaptation along an environmental gradient. *Evolution* **67**,
534 2530-2546.
- 535 DeFaveri J, Merilä J (2013) Evidence for adaptive phenotypic differentiation in Baltic Sea
536 sticklebacks. *Journal of Evolutionary Biology* **26**, 1700-1715.
- 537 DeFaveri J, Shikano T, Shimada Y, Goto A, Merilä J (2011) Global analysis of genes involved in
538 freshwater adaptation in threespine stickleback (*Gasterosteus aculeatus*). *Evolution*
539 **65**, 1800-1807.
- 540 Dvorsky R, Ahmadian MR (2004) Always look on the bright site of Rho: structural
541 implications for a conserved intermolecular interface. *EMBO reports* **5**, 1130-1136.
- 542 Endler JA (1973) Gene flow and population differentiation: Studies of clines suggest that
543 differentiation along environmental gradients may be independent of gene flow.
544 *Science* **179**, 243-250.
- 545 Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured
546 population. *Heredity* **103**, 285-298.
- 547 Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform
548 population genetics analyses under Linux and Windows. *Molecular Ecology Resources*
549 **10**, 564-567.
- 550 Feder JL, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends in*
551 *Genetics* **28**, 342-350.
- 552 Feulner PGD, Chain FJJ, Panchal M, *et al.* (2013) Genome-wide patterns of standing genetic
553 variation in a marine population of three-spined sticklebacks. *Molecular Ecology* **22**,
554 635-649.

- 555 Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for
556 both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977-
557 993.
- 558 Foster SA, Scott RJ, Cresko WA (1998) Nested biological variation and speciation.
559 *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*
560 **353**, 207-218.
- 561 Garnaas MK, Moodie KL, Liu M-I, *et al.* (2008) Syx, a RhoA guanine exchange factor, is
562 essential for angiogenesis in vivo. *Circulation Research* **103**, 710-716.
- 563 Hagen DW (1967) Isolating mechanisms in threespine sticklebacks (*Gasterosteus*). *Journal of*
564 *the Fisheries Research Board of Canada* **24**, 1637-&.
- 565 Hansen MM, Meier K, Mensberg K-LD (2010) Identifying footprints of selection in stocked
566 brown trout populations: a spatio-temporal approach. *Molecular Ecology* **19**, 1787-
567 1800.
- 568 Hansen MM, Ruzzante DE, Nielsen EE, Bekkevold D, Mensberg K-LD (2002) Long-term
569 effective population sizes, temporal stability of genetic composition and potential for
570 local adaptation in anadromous brown trout (*Salmo trutta*) populations. *Molecular*
571 *Ecology* **11**, 2523-2535.
- 572 Hendry AP (2009) Ecological speciation! Or the lack thereof? *Canadian Journal of Fisheries*
573 *and Aquatic Sciences* **66**, 1383-1398.
- 574 Hendry AP, Bolnick DI, Berner D, Peichel CL (2009) Along the speciation continuum in
575 sticklebacks. *Journal of Fish Biology* **75**, 2000-2036.
- 576 Heuts MJ (1947) The phenotypical variability of *Gasterosteus aculeatus* L. populations in
577 Belgium. *Verhandelingen van de Koninklijke Vlaamse Academie voor Wetenschappen,*
578 *Letteren en Schone Kunsten van België* **9**, 1-63.
- 579 Hohenlohe PA, Bassham S, Currey M, Cresko WA (2012) Extensive linkage disequilibrium and
580 parallel adaptive divergence across threespine stickleback genomes. *Philosophical*
581 *Transactions of the Royal Society of London, Series B: Biological Sciences* **367**, 395-
582 408.
- 583 Hohenlohe PA, Bassham S, Etter PD, *et al.* (2010) Population genomics of parallel adaptation
584 in threespine stickleback using sequenced RAD tags. *PLoS Genetics* **6**.
- 585 Jones F, Brown C, Pemberton J, Braithwaite V (2006) Reproductive isolation in a threespine
586 stickleback hybrid zone. *Journal of Evolutionary Biology* **19**, 1531-1544.
- 587 Jones FC, Chan YF, Schmutz J, *et al.* (2012a) A genome-wide SNP genotyping array reveals
588 patterns of global and repeated species-pair divergence in sticklebacks. *Current*
589 *Biology* **22**, 83-90.
- 590 Jones FC, Grabherr MG, Chan YF, *et al.* (2012b) The genomic basis of adaptive evolution in
591 threespine sticklebacks. *Nature* **484**, 55-61.
- 592 Kauer MO, Dieringer D, Schlotterer C (2003) A microsatellite variability screen for positive
593 selection associated with the "Out of Africa" habitat expansion of *Drosophila*
594 *melanogaster*. *Genetics* **165**, 1137-1148.
- 595 Levene H (1953) Genetic equilibrium when more than one ecological niche is available.
596 *American Naturalist* **87**, 331-333.
- 597 Lind EE, Grahn M (2011) Directional genetic selection by pulp mill effluent on multiple
598 natural populations of three-spined stickleback (*Gasterosteus aculeatus*).
599 *Ecotoxicology* **20**, 503-512.

- 600 Lucek K, Sivasundar A, Kristjánsson BK, Skúlason S, Seehausen O (2014) Quick divergence but
601 slow convergence during ecotype formation in lake and stream stickleback pairs of
602 variable age. *Journal of Evolutionary Biology* **27**, 1878–1892.
- 603 Mäkinen HS, Cano JM, Merilä J (2008) Identifying footprints of directional and balancing
604 selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*)
605 populations. *Molecular Ecology* **17**, 3565-3582.
- 606 McCairns RJS, Bernatchez L (2010) Adaptive divergence between freshwater and marine
607 sticklebacks: insights into the role of phenotypic plasticity from an integrated analysis
608 of candidate gene expression. *Evolution* **64**, 1029-1047.
- 609 McKinnon JS, Rundle HD (2002) Speciation in nature: the threespine stickleback model
610 systems. *Trends in Ecology & Evolution* **17**, 480-488.
- 611 McPhail J (1994) Speciation and the evolution of reproductive isolation in the sticklebacks
612 (*Gasterosteus*) of south-western British Columbia. In: *The evolutionary biology of the*
613 *threespine stickleback* (eds. Bell AM, Foster S), pp. 399-437. Oxford University Press
614 Inc., New York.
- 615 Mitchell-Olds T, Schmitt J (2006) Genetic mechanisms and evolutionary significance of
616 natural variation in *Arabidopsis*. *Nature* **441**, 947-952.
- 617 Moore J-S, Gow JL, Taylor EB, Hendry AP (2007) Quantifying the constraining influence of
618 gene flow on adaptive divergence in the lake-stream threespine stickleback System.
619 *Evolution* **61**, 2015-2026.
- 620 Narum SR, Hess JE (2011) Comparison of F-ST outlier tests for SNP loci under selection.
621 *Molecular Ecology Resources* **11**, 184-194.
- 622 Nielsen E, Hemmer-Hansen J, Poulsen N, *et al.* (2009) Genomic signatures of local directional
623 selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC*
624 *Evolutionary Biology* **9**, 276.
- 625 Nosil P (2009) Adaptive population divergence in cryptic color-pattern following a reduction
626 in gene flow. *Evolution* **63**, 1902-1912.
- 627 Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic
628 divergence. *Molecular Ecology* **18**, 375-402.
- 629 Olafsdottir GA, Snorrason SS (2009) Parallels, nonparallels, and plasticity in population
630 differentiation of threespine stickleback within a lake. *Biological Journal of the*
631 *Linnean Society* **98**, 803-813.
- 632 Orsini L, Spanier KI, De Meester LUC (2012) Genomic signature of natural and anthropogenic
633 stress in wild populations of the waterflea *Daphnia magna*: validation in space, time
634 and experimental evolution. *Molecular Ecology* **21**, 2160-2175.
- 635 Papa R, Martin A, Reed RD (2008) Genomic hotspots of adaptation in butterfly wing pattern
636 evolution. *Current Opinion in Genetics & Development* **18**, 559-564.
- 637 Peichel CL, Nereng KS, Ohgi KA, *et al.* (2001) The genetic architecture of divergence between
638 threespine stickleback species. *Nature* **414**, 901-905.
- 639 Perrier C, Bourret V, Kent MP, Bernatchez L (2013) Parallel and nonparallel genome-wide
640 divergence among replicate population pairs of freshwater and anadromous Atlantic
641 salmon. *Molecular Ecology* **22**, 5577-5593.
- 642 Pinho C, Hey J (2010) Divergence with gene flow: models and data. *Annual Review of*
643 *Ecology, Evolution and Systematics* **41**, 215-230.
- 644 Poncet BN, Herrmann D, Gugerli F, *et al.* (2010) Tracking genes of ecological relevance using
645 a genome scan in two independent regional population samples of *Arabidopsis alpina*.
646 *Molecular Ecology* **19**, 2896-2907.

- 647 Radaelli G, Rowlerson A, Mascarello F, Patrino M, Funkenstein B (2003) Myostatin precursor
648 is present in several tissues in teleost fish: a comparative immunolocalization study.
649 *Cell and Tissue Research* **311**, 239-250.
- 650 Raeker MÖ, Bieniek AN, Ryan AS, *et al.* (2010) Targeted deletion of the zebrafish obscurin A
651 RhoGEF domain affects heart, skeletal muscle and brain development.
652 *Developmental Biology* **337**, 432-443.
- 653 Raeymaekers JAM, Konijnendijk N, Larmuseau MHD, *et al.* (2014) A gene with major
654 phenotypic effects as a target for selection vs. homogenizing gene flow. *Molecular*
655 *Ecology* **23**, 162-181.
- 656 Raeymaekers JAM, Lens LUC, Van den Broeck F, Van Dongen S, Volckaert FAM (2012)
657 Quantifying population structure on short timescales. *Molecular Ecology* **21**, 3458-
658 3473.
- 659 Raeymaekers JAM, Maes GE, Audenaert E, Volckaert FAM (2005) Detecting Holocene
660 divergence in the anadromous-freshwater three-spined stickleback (*Gasterosteus*
661 *aculeatus*) system. *Molecular Ecology* **14**, 1001-1014.
- 662 Raeymaekers JAM, Van Houdt JKJ, Larmuseau MHD, Geldof S, Volckaert FAM (2007)
663 Divergent selection as revealed by P_{ST} and QTL-based F_{ST} in three-spined stickleback
664 (*Gasterosteus aculeatus*) populations along a coastal-inland gradient. *Molecular*
665 *Ecology* **16**, 891-905.
- 666 Roesti M, Hendry AP, Salzburger W, Berner D (2012) Genome divergence during evolutionary
667 diversification as revealed in replicate lake-stream stickleback population pairs.
668 *Molecular Ecology* **21**, 2852-2862.
- 669 Rohlf FJ (2002) tpsDig, digitize landmarks and outlines. Department of Ecology and Evolution,
670 State University of New York at Stony Brook., New York.
- 671 Rossman KL, Der CJ, Sondek J (2005) GEF means go: turning on RHO GTPases with guanine
672 nucleotide-exchange factors. *Nature Reviews: Molecular Cell Biology* **6**, 167-180.
- 673 Schluter D (1996) Adaptive radiation along genetic lines of least resistance. *Evolution*, 1766-
674 1774.
- 675 Schluter D (2009) Evidence for Ecological Speciation and Its Alternative. *Science* **323**, 737-
676 741.
- 677 Schluter D, Conte GL (2009) Genetics and ecological speciation. *Proceedings of the National*
678 *Academy of Sciences, USA* **106**, 9955-9962.
- 679 Seehausen O, Butlin RK, Keller I, *et al.* (2014) Genomics and the origin of species. *Nature*
680 *Reviews: Genetics* **15**, 176-192.
- 681 Shikano T, Ramadevi J, Merilä J (2010a) Identification of local-and habitat-dependent
682 selection: scanning functionally important genes in nine-spined sticklebacks
683 (*Pungitius pungitius*). *Molecular Biology and Evolution* **27**, 2775-2789.
- 684 Shikano T, Ramadevi J, Shimada Y, Merilä J (2010b) Utility of sequenced genomes for
685 microsatellite marker development in non-model organisms: a case study of
686 functionally important genes in nine-spined sticklebacks (*Pungitius pungitius*). *BMC*
687 *Genomics* **11**, 334.
- 688 Shimada Y, Shikano T, Merilä J (2011) A high incidence of selection on physiologically
689 important genes in the three-spined stickleback, *Gasterosteus aculeatus*. *Molecular*
690 *Biology and Evolution* **28**, 181-193.
- 691 Slatkin (1987). Gene flow and the geographic structure of natural populations. *Science*
692 **236**,787–792.

- 693 Soria-Carrasco V, Gompert Z, Comeault AA, *et al.* (2014) Stick insect genomes reveal natural
694 selection's role in parallel speciation. *Science* **344**, 738-742.
- 695 Stapley J, Reger J, Feulner PGD, *et al.* (2010) Adaptation genomics: the next generation.
696 *Trends in Ecology & Evolution* **25**, 705-712.
- 697 Storz JF (2005) Invited review: Using genome scans of DNA polymorphism to infer adaptive
698 population divergence. *Molecular Ecology* **14**, 671-688.
- 699 Strasburg JL, Sherman NA, Wright KM, *et al.* (2012) What can patterns of differentiation
700 across plant genomes tell us about adaptation and speciation? *Philosophical*
701 *Transactions of the Royal Society of London, Series B: Biological Sciences* **367**, 364-
702 373.
- 703 Taylor WR, Dyke GCV (1985) Revised procedures for staining and clearing small fishes and
704 other vertebrates for bone and cartilage study. *Cybium* **9**, 107-119.
- 705 Tregenza T (2002) Divergence and reproductive isolation in the early stages of speciation. In:
706 *Genetics of Mate Choice: From Sexual Selection to Sexual Isolation* (eds. Etges WJ,
707 Noor MAF), pp. 291-300. Springer Netherlands.
- 708 Van Dongen S, Lens L, Pape E, Volckaert FAM, Raeymaekers JAM (2009) Evolutionary history
709 shapes the association between developmental instability and population-level
710 genetic variation in three-spined sticklebacks. *Journal of Evolutionary Biology* **22**,
711 1695-1707.
- 712 van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: software for
713 identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology*
714 *notes* **4**, 535-538.
- 715 Vandamme S, Maes G, Raeymaekers J, *et al.* (2014) Regional environmental pressure
716 influences population differentiation in turbot (*Scophthalmus maximus*). *Molecular*
717 *Ecology* **23**, 618-636.
- 718 Via S (2009) Natural selection in action during speciation. *Proceedings of the National*
719 *Academy of Sciences, USA* **106**, 9939-9946.
- 720 Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the
721 human genome. *Plos Biology* **4**, e72.
- 722 Walsh F, Celeste A (2005) Myostatin: a modulator of skeletal-muscle stem cells. *Biochemical*
723 *Society Transactions* **33**, 1513-1517.
- 724 Wootton RJ (1976) *The Biology of the Sticklebacks*. Academic Press, London.
- 725 Yang L, Jelsbak L, Marvig RL, *et al.* (2011) Evolutionary dynamics of bacteria in a human host
726 environment. *Proceedings of the National Academy of Sciences, USA* **108**, 7481-7486.
- 727 Yeaman S, Otto SP (2011) Establishment and maintenance of adaptive genetic divergence
728 under migration, selection, and drift. *Evolution* **65**, 2123-2129.
- 729 Yeaman S, Whitlock MC (2011) The genetic architecture of adaptation under migration-
730 selection balance. *Evolution* **65**, 1897-1911.
- 731 Zulliger D, Schnyder E, Gugerli F (2013) Are adaptive loci transferable across genomes of
732 related species? Outlier and environmental association analyses in Alpine
733 Brassicaceae species. *Molecular Ecology* **22**, 1626-1639.
- 734

735 **TABLES**

736

737 **Table 1.** Comparison of microsatellite-based genome scan studies in three-spined stickleback
 738 across freshwater - brackish/salt water gradients, including spatial scale (from regional to
 739 global), percentage of conservative outliers, F_{ST} , F_{ST} at the *Eda* locus, and F_{ST} at the *ATP1A1*
 740 locus.

Study	Spatial scale	% outliers	F_{ST}	$F_{ST} Eda$	$F_{ST} ATP1A1$
This study	78 km	2.3	0.059	0.118	0.154
DeFaveri <i>et al.</i> 2013	Regional	12.5	0.008	-	0.016
Shimada <i>et al.</i> 2011	Continental	8.3	0.107	-	0.225
Mäkinen <i>et al.</i> 2008	Continental	3.0	0.166	0.653	-
DeFaveri <i>et al.</i> 2011	Global	33	0.119	0.405	-

For Review Only

741
 742 **Table 2.** Characteristics of the four three-spined stickleback populations used in this study. DTC:
 743 distance to coast; MPN: Mean plate number; Eda^L : frequency of *Eda* low plated allele; AR: allelic
 744 richness; H_e : expected heterozygosity; F_{IS} : inbreeding coefficient, with confidence interval (CI
 745 95%).
 746

Code	Coordinates (N, E)	Habitat	DTC (km)	Salinity(psu)	MPN	Eda^L	AR	H_e	F_{IS} (CI 95%)
L01	51°21'10.66, 3°26'01.83	Creek	3.94	2.04	17.4	0.44	9.47	0.69	0.032 (-0.017 - 0.033)
L02	51°21'56.33, 3°31'11.09	Creek	4.30	1.83	16.8	0.51	9.06	0.69	0.095 (0.048 - 0.096)
L12	51°10'29.03, 3°28'10.45	Stream	22.84	0.32	12.7	0.74	6.80	0.61	0.044 (-0.012 - 0.050)
U01	51°02'48.40, 3°33'01.60	Ditch	36.20	0.52	7.0	0.93	6.51	0.57	0.039 (-0.010 - 0.040)

747 **Table 3.** Pairwise F_{ST} values below the diagonal, and associated P-value above the diagonal. For
748 site codes see Table 2.
749

Code	L01	L02	L12	U01
L01	-	0.600	< 0.010	< 0.010
L02	0.005	-	< 0.010	< 0.010
L12	0.054	0.056	-	< 0.010
U01	0.094	0.097	0.047	-

750

For Review Only

751 **Table 4.** Global outlier loci that are putatively under directional selection detected in four
 752 populations of three-spined stickleback. The table includes the locus considered, its F_{ST} -value
 753 and either the q value for Bayescan or the P-value for LnRH, LOSITAN and Arlequin. Significance
 754 is marked as † = $q < 0.1$, †† = $q < 0.05$ and ††† = $q < 0.01$ or as * = $p < 0.05$, ** = $p < 0.01$ and ***
 755 = $p < 0.001$. Marker *Stn380*, linked to the *Eda* gene, was added as a reference.
 756

Locus	F_{ST}	Bayescan	LnRH	LOSITAN	Arlequin (2 groups)	Arlequin (3 groups)
<i>ATP1A1</i>	0.149			*	*	*
<i>HSPA14</i>	0.159			*	*	*
<i>Ppgm44</i>	0.120	†	***			
<i>Stn74</i>	0.083		*			
<i>Stn177</i>	0.189			**	**	**
<i>Stn46</i>	0.199	†††		**	**	**
<i>Stn380</i> (<i>Eda</i>)	0.118					

757

758 **Supplementary Table S1.** Characterisation of microsatellite loci of three-spined stickleback,
 759 including name, type of marker, gene assignment and putative function.

Locus	Type of marker	Gene	Putative function
<i>1097PBBE</i>	Random genomic	Novel gene	
<i>1125PBBE</i>	Random genomic	Novel gene	
<i>7033PBBE</i>	Random genomic	Integrin beta-3 precursor	
<i>7080PBBE</i>	Random genomic	Neurensin-1	
<i>ATP1A1</i>	Gene-based	Na ⁺ /K ⁺ -ATPase alpha-subunit isoform 1	Osmoregulation
<i>ATP4A</i>	Gene-based	H ⁺ /K ⁺ -ATPase alpha subunit	Osmoregulation
<i>CLCN7</i>	Gene-based	CLC chloride channel isoform 7	Osmoregulation
<i>DCT</i>	Gene-based	L-dopachrome tautomerase precursor	Pigmentation
<i>FGF18</i>	Gene-based	Fibroblast growth factor isoform 18	Growth
<i>FGF8</i>	Gene-based	Fibroblast growth factor isoform 8	Growth
<i>GAest11</i>	Random EST		
<i>GAest16</i>	Random EST	Novel gene	
<i>GAest3</i>	Random EST	Dihydropyrimidinase-related protein 2	
<i>GAest34</i>	Random EST	6-phosphofructokinase type C	
<i>GAest35</i>	Random EST	Rod outer segment membrane protein 1	
<i>GAest41</i>	Random EST	Tumor necrosis factor receptor superfamily member 19 precursor	
<i>GAest47</i>	Random EST		
<i>GAest50</i>	Random EST	Eukaryotic translation initiation factor 3 subunit K	
<i>GAest51</i>	Random EST	Protein kinase C delta-binding protein	
<i>GAest6</i>	Random EST	Ninjurin-2	
<i>GAest67</i>	Random EST	Novel gene	
<i>GAest7</i>	Random EST	Ras-related protein Rab-3A	
<i>GAest82</i>	Random EST	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	
<i>GAest84</i>	Random EST		
<i>GAest87</i>	Random EST		
<i>GS1</i>	Gene-based	Glycogen synthase subtype 1	Growth
<i>HSP70Ab</i>	Gene-based	Heat shock protein 70 kDa subtype 2	Thermal response
<i>HSP70Ad</i>	Gene-based	Heat shock protein 70 kDa subtype 4	Thermal response
<i>IGF-II</i>	Gene-based	Insulin-like growth factor II precursor	Growth
<i>KCNJ4b</i>	Gene-based	Inward rectifier potassium channel subfamily J member 4 subtype 2	Osmoregulation
<i>MYHb</i>	Gene-based	Myosin heavy chain subtype 2	Thermal response
<i>NHE2c</i>	Gene-based	Na ⁺ /H ⁺ exchanger subtype 3	Hypoxia
<i>NHE3</i>	Gene-based	Sodium/hydrogen exchanger isoform 3 (SLC9A3)	Osmoregulation
<i>NPY2Rb</i>	Gene-based	Neuropeptide Y receptor Y2 subtype 2	Growth
<i>PER1</i>	Gene-based	Period circadian protein homolog 1	Maturation
<i>Pp gm44</i>	Gene-based	Myostatin subtype 2	Growth
<i>Pp gm17</i>	Gene-based	Type I iodothyronine deiodinase	Morphology

<i>Ppgm16</i>	Gene-based	Cold shock protein 70 kDa subtype 2	Thermal response
<i>Ppgm24</i>	Gene-based	Growth hormone receptor 1	Growth
<i>Ppgm28</i>	Gene-based	General transcription factor IIB	Osmoregulation
<i>Ppgm58</i>	Gene-based	Tristetraprolin	Disease
<i>Ppgm21</i>	Gene-based	Fibroblast growth factor 18	Growth
<i>Ppgm47</i>	Gene-based	Na ⁺ /K ⁺ /2Cl ⁻ -cotransporter isoform 1 subtype 2	Osmoregulation
<i>Ppgm35</i>	Gene-based	Heat shock protein 70 kDa isoform beta	Thermal response
<i>Ppgm7</i>	Gene-based	Na ⁺ /K ⁺ -ATPase alpha-subunit isoform 2	Osmoregulation
<i>Ppgm56</i>	Gene-based	Amine-associated receptor	Taste
<i>Ppgm1</i>	Gene-based	Pituitary adenylate cyclase activating polypeptide receptor sub-type 1	Growth
<i>Ppgm2</i>	Gene-based	Pituitary adenylate cyclase activating polypeptide receptor sub-type 2	Growth
<i>Ppgm49</i>	Gene-based	Neuropeptide Y precursor	Growth
<i>Ppgm30</i>	Gene-based	Heat shock protein 25 kDa	Thermal response
<i>Ppgm33</i>	Gene-based	Heat shock protein 70 kDa isoform alpha subtype 2	Thermal response
<i>Ppgm50</i>	Gene-based	Parvalbumin subtype 2	Thermal response
<i>Ppgm6</i>	Gene-based	Na ⁺ /K ⁺ -ATPase alpha-subunit isoform 1	Osmoregulation
<i>Ppgm38</i>	Gene-based	Insulin-like growth factor I	Growth
<i>PRL-R</i>	Gene-based	Prolactin receptor a	Osmoregulation
<i>PVALBa</i>	Gene-based	Parvalbumin subtype 1	Growth
<i>RhBG</i>	Gene-based	Rhesus glycoprotein isoform 3	Toxic stress
<i>SLC4A10a</i>	Gene-based	Electroneutral Na-coupled HCO ₂ transporter isoform 10 subtype 1	Hypoxia
<i>Stn102</i>	Random genomic	EF-hand domain-containing family member A2	
<i>Stn108</i>	Random genomic	Bifunctional heparan sulfate N-deacetylase/N-sulfotransferase 4	
<i>Stn119</i>	Random EST		
<i>Stn122</i>	Random EST		
<i>Stn124</i>	Random EST	Brain-specific angiogenesis inhibitor 2 precursor	
<i>Stn163</i>	Random EST	Novel gene	
<i>Stn164</i>	Random EST		
<i>Stn167</i>	Random genomic		
<i>Stn177</i>	Random genomic		
<i>Stn194</i>	Random genomic	Nuclear-localized factor 1	
<i>Stn20</i>	Random EST		
<i>Stn222</i>	Random genomic	Ras-related protein Rab-18	
<i>Stn223</i>	Random genomic	Zinc finger homeobox protein 4	
<i>Stn235</i>	Random EST		
<i>Stn240</i>	Random EST		
<i>Stn257</i>	Random EST	Solute carrier family 35 member F2	
<i>Stn299</i>	Random genomic		
<i>Stn302</i>	Random genomic		
<i>Stn315</i>	Random genomic	Histone deacetylase 4	
<i>Stn328</i>	Random genomic	RING finger protein 220	
<i>Stn34</i>	Random EST		
<i>Stn380</i>	Gene-based	Eda	plate nr

<i>Stn46</i>	Random genomic	Rho guanine nucleotide exchange factor 9	
<i>Stn49</i>	Random genomic	Dedicator of cytokinesis protein 4	
<i>Stn71</i>	Random EST	Novel gene	
<i>Stn74</i>	Random EST		
<i>Stn79</i>	Random genomic	Reticulon-2	
<i>Stn89</i>	Random EST	Rho GTPase-activating protein SYDE2	
<i>HPX</i>	Gene-based	Hemopexin	Immune response

760

For Review Only

761 **Supplementary Table S2.** Global outlier loci that are putatively under balancing selection as
 762 detected with the methods of Bayescan, LOSITAN and Arlequin. Significance is marked as † = $q <$
 763 0.1, †† = $q < 0.05$ and ††† = $q < 0.01$ or * = $p < 0.05$, ** = $p < 0.01$ and *** = $p < 0.001$.

Locus	F _{ST}	Bayescan	LOSITAN	Arlequin (2 groups)	Arlequin (3 groups)
<i>FGF8</i>	0.022		*		
<i>GAest3</i>	0.021		*	*	
<i>GAest34</i>	0.015	††	***	*	*
<i>GAest47</i>	0.009		*	*	*
<i>IGF-II</i>	-0.007		*		
<i>NHE2c</i>	0.027		*	*	
<i>NHE3</i>	0.004		*	*	*
<i>Ppgm16</i>	-0.009		*		
<i>Ppgm2</i>	0.005			**	*
<i>Ppgm30</i>	0.013		*	*	*
<i>Ppgm6</i>	0.005			*	*
<i>Stn122</i>	0.014		***	**	*
<i>Stn20</i>	0.017		*		
<i>Stn223</i>	0.019	†	**	*	*
<i>Stn79</i>	-0.020		***	*	*
<i>Stn89</i>	0.028	†††			

764

765 **FIGURE LEGENDS**

766

767 **Figure 1.** Map with the four sampling locations of three-spined stickleback in northwestern Belgium and
768 southwestern Netherlands. River network and major towns are mapped.

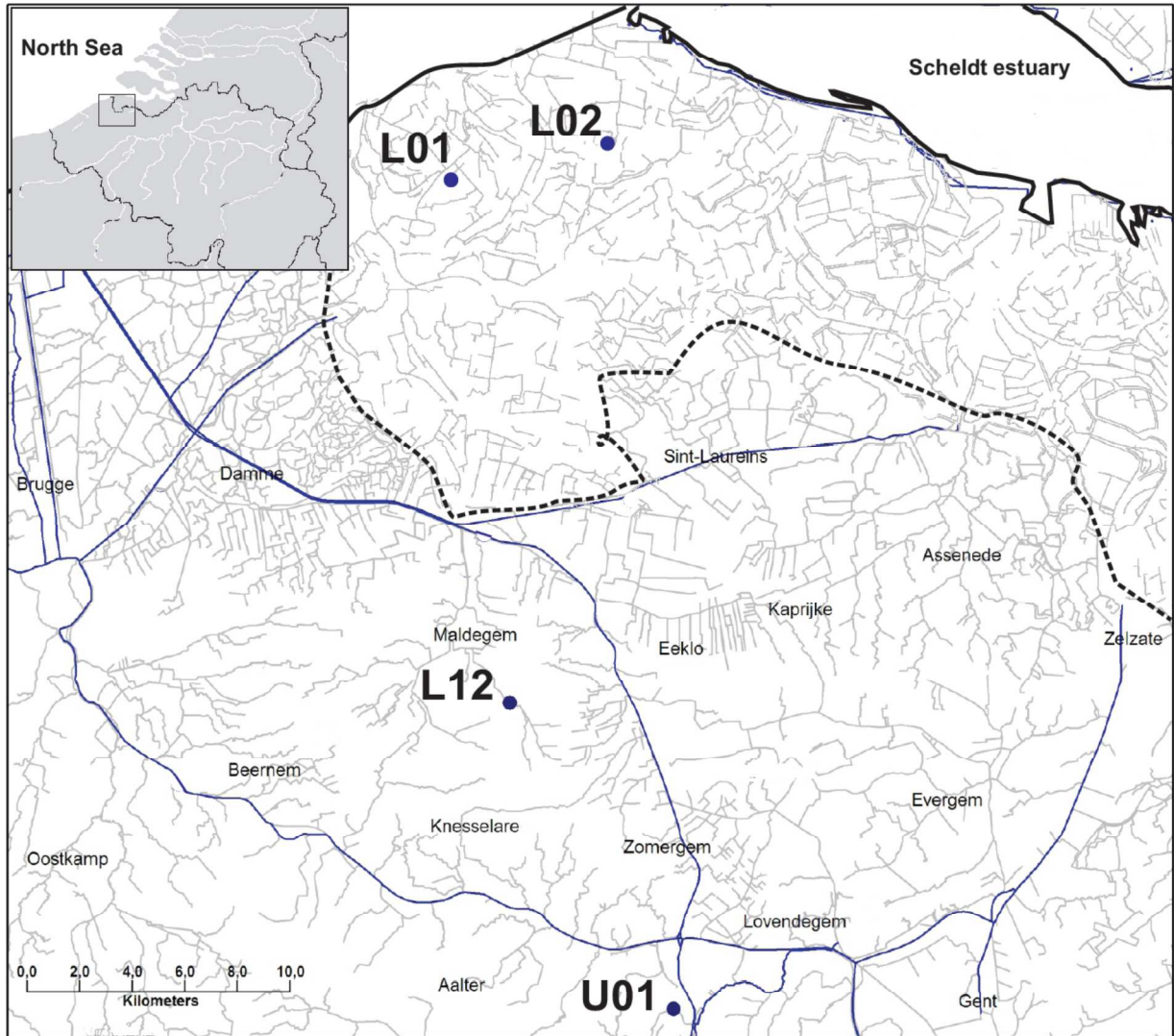
769

770 **Figure 2.** Average values for A) standard length, B) plate number, C) relative pelvic spine length, D)
771 relative pelvic plate length, and E) relative length of the second large gill raker in two brackish water
772 populations (L01, L02) and two freshwater populations (L12 and U01)) of three-spined stickleback.
773 Horizontal bars represent 95% confidence intervals.

774

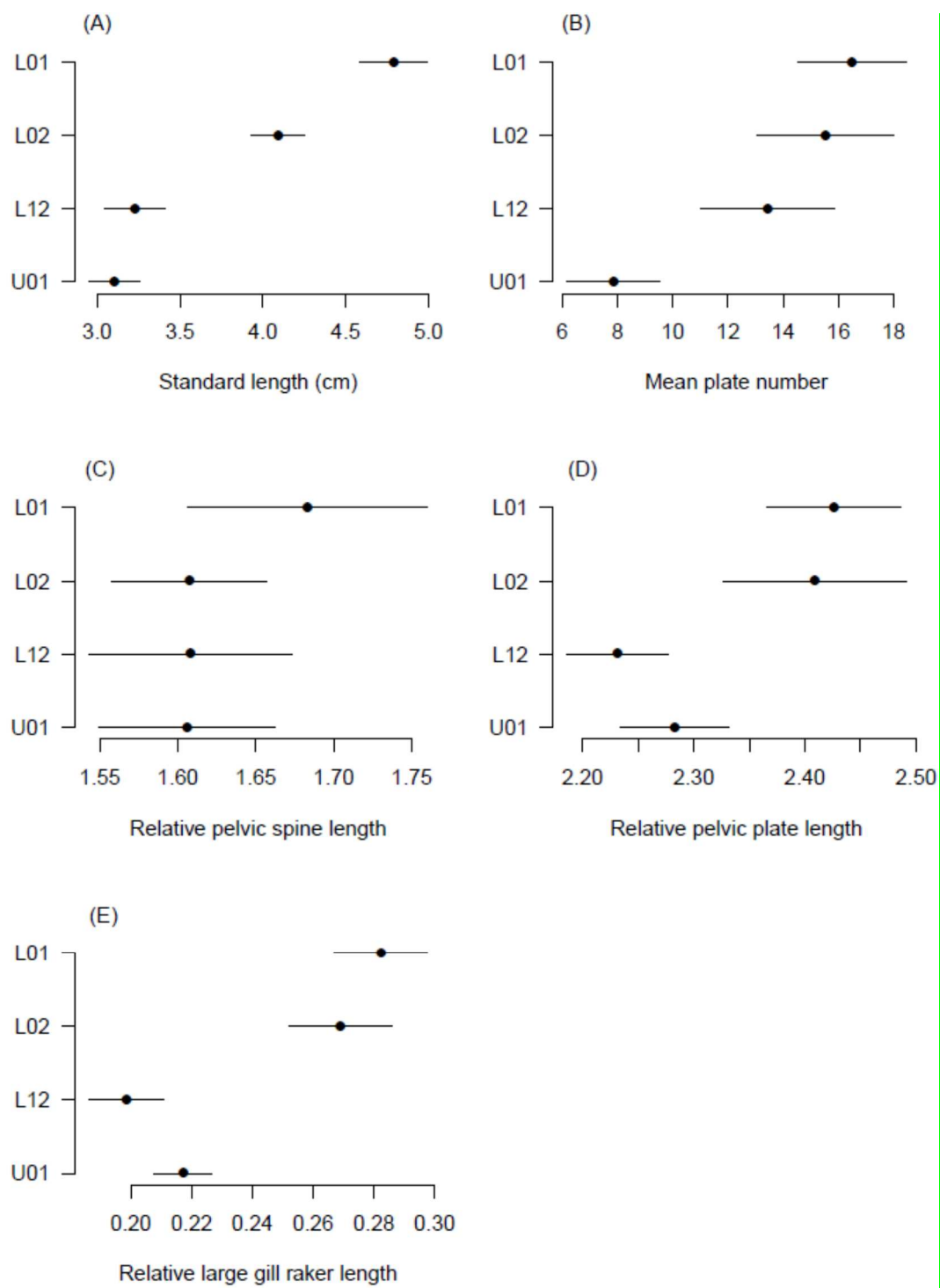
For Review Only

775 FIGURES
 776
 777 **Figure 1**
 778



779
 780
 781

782 **Figure 2**
783



784

1 | Signatures of selection in the three-spined stickleback along a small scale ~~salt~~brackish
2 | water-and- freshwater transition zone

3 |
4 | Nellie Konijnendijk¹, Takahito Shikano², Dorien Daneels^{1,*}, Filip A.M. Volckaert¹, Joost A. M.
5 | Raeymaekers¹

6 |
7 | ¹ Laboratory of Biodiversity and Evolutionary Genomics, University of Leuven, Ch.
8 | Deberiotstraat, 32, B-3000 Leuven, Belgium

9 | ² Ecological Genetics Research Unit, Department of Biosciences, University of Helsinki, P.O.
10 | Box 65, FI-000 14, Helsinki, Finland

11 | * Current address: Centre for Medical Genetics, Reproduction and Genetics; Reproduction,
12 | Genetics and Regenerative Medicine, Vrije Universiteit Brussel (VUB), Universitair ziekenhuis
13 | Brussel (UZ Brussel), Laarbeeklaan 101, 1090 Brussels, Belgium.

14 |
15 | **Running header:** Adaptation in three-spined stickleback
16 |
17 |
18 | **Address for correspondence:**

19 | Joost Raeymaekers, University of Leuven
20 | Laboratory of Biodiversity and Evolutionary Genomics
21 | Ch. Deberiotstraat, 32
22 | B-3000 Leuven, Belgium
23 | e-mail: joost.raeymaekers@bio.kuleuven.be
24 | Phone : + 32 16 32 39 66, Fax : +32 16 32 45 75

25 |

26 **ABSTRACT**

27 Local adaptation is often obvious when gene flow is impeded, such as observed at large
28 spatial scales and across strong ecological contrasts. However, but it becomes less certain at
29 small scales such as between adjacent populations or across weak ecological contrasts, when
30 gene flow is strong. While studies on Genomic-genomic adaptation tend to focus on the
31 former, is often investigated in populations that are either geographically isolated or that
32 occur across strong ecological contrasts. In both cases, gene flow might be strongly impeded,
33 and adaptation might have been largely achieved. Less is known about the genomic targets
34 of natural selection at the onset of divergence, in the latter situation. when gene flow is
35 strong. In this study we focus investigate genomic adaptation on in a system characterized
36 by very limited genetic isolation, which is often a realistic setting in nature. We performed a
37 genome scan with 87 microsatellite markers among populations of the three-spined
38 stickleback *Gasterosteus aculeatus* L. across a small-scale along an ecological transition with
39 salinities ranging from brackish to fresh water habitats. Adaptation to Salinity salinity is
40 known to be an important driver of adaptation has been repeatedly demonstrated in this
41 species and has been studied in stark contrasting environments several times. Adaptation to
42 this ecological variable is therefore likely. A genome scan based on 87 microsatellite markers
43 As expected, we detected revealed only few signatures of selection at this scale of population
44 divergence, likely owing to the constraints that homogenizing gene flow puts on adaptive
45 divergence. However, those that we did find the detected loci appear repeatedly as targets
46 of selection in similar studies of genomic adaptation in three-spined sticklebacks. We
47 conclude that the signature of genomic selection in the face of strong gene flow is weak, yet
48 detectable. We argue that the range of studies of genomic divergence should be extended to
49 include more the very onset of population divergencesystems characterised by limited

Formatted: Font: Italic

50 ~~geographical and ecological isolation, which is often a realistic setting in nature i.e. at the~~
51 ~~level where it has been hypothesized that adaptation is primarily achieved through non-~~
52 ~~genomic mechanisms such as phenotypic plasticity.~~

53

54 **Keywords:** adaptation, candidate gene, evolution, fish, genome scan, population genomics,
55 selection

56

For Review Only

57 INTRODUCTION

58

59 Recent questions on the mechanisms of evolutionary biology revolve around the genomic
60 architecture of species and the effects that processes such as selection, drift, mutation and
61 gene flow have on the genome. Changes in the environment can now be linked to the
62 genetic signatures of these processes and hence promote the understanding of the genetic
63 basis of ecological adaptation. This allows us to understand the various components of the
64 mechanism of adaptation such as which genes are involved, how they are distributed in the
65 genome, and how often the same genes lead to certain adaptations both in the laboratory
66 (Becks *et al.* 2012; ~~and~~ see Barrick & Lenski 2013 for other examples) and nature (Mitchell-
67 Olds & Schmitt 2006; Yang *et al.* 2011; Jones *et al.* 2012a; Orsini *et al.* 2012). Additionally we
68 want to understand which processes are essential for adaptation to succeed. One way to
69 find genes that are important for the process of adaptation is to identify genetic signatures
70 of adaptation in natural populations. These "genome scans" identify genes that are under
71 selection across contrasting environments and might thus be involved in the transition from
72 one ecological extreme to the other. When interpreted with care (de Villemereuil *et al.*
73 2014), this approach has shown to be a promising way to find links between genotype,
74 phenotype and fitness in natural populations (Storz 2005).

75

76 Genome scans have so far generated two major insights. First, selection may act on many
77 parts of the genome (Nosil 2009; Hohenlohe *et al.* 2010; Feder *et al.* 2012; Jones *et al.*
78 2012a; Jones *et al.* 2012b; Strasburg *et al.* 2012; Arnegard *et al.* 2014; Seehausen *et al.*
79 2014). Nosil (2009) estimated that 5-10 % of the genome is affected by natural selection
80 ~~(although such estimates should be interpreted cautiously as the authors suggested, see also~~

81 ~~Stapley et al. (2010) and Excoffier et al. (2009)~~. Based on a genome-wide analysis of
82 selection, Hohenlohe et al. (2012) showed that the genome is much more structured and
83 dynamic than expected from theory. More studies have now contributed to these topics and
84 show that the areas affected by natural selection are typically patchily distributed
85 (“hotspots”) across the genome (Voight et al. 2006; Papa et al. 2008; Hohenlohe et al. 2010).
86 ~~Hohenlohe et al. (2010) suggest that~~ genomic inversions, deletions, repetitive elements
87 and other structural changes facilitate the rise of these islands of divergence, causing a
88 physical barrier to crossing over and thus adding to linkage disequilibrium (Hohenlohe et al.
89 ~~2010~~). On the other hand, ~~a review dedicated to genome scans in plants (Strasburg et al.~~
90 ~~2012) describes that~~ most cases of adaptation ~~across environments in plants~~ involve many
91 genomic areas with a wide distribution across the genome rather than clustered hotspots
92 (Strasburg et al. 2012).

93
94 A second common finding of genome scans is that ~~organisms may show parallel as well as~~
95 ~~non-parallel there is large variation across studies in the degree of parallelism for~~ genetic
96 responses to environmental change. For example, Atlantic salmon (*Salmo salar*) populations
97 that undergo a parallel change in environment show non-parallel adaptive divergence ~~within~~
98 ~~the genome at the genomic level~~ (Perrier et al. 2013). In cases of parallel phenotypic
99 evolution in the three-spined stickleback (*Gasterosteus aculeatus* L.), common (Hohenlohe
100 et al. 2010; Shimada et al. 2011; Jones et al. 2012b) as well as unique (Mäkinen et al. 2008;
101 Jones et al. 2012b; Roesti et al. 2012) genomic regions are targeted by selection. This has
102 been confirmed in experiments with stick insects (*Timema cristinae*) (Soria-Carrasco et al.
103 2014) and in genome scans of lake white fish (*Coregonus* sp.) (Bernatchez et al. 2010). In
104 contrast, parallel phenotypic evolution in ~~a~~ other species, ~~like such as~~ the nine-spined

105 stickleback (*Pungitius pungitius*) (Shikano *et al.* 2010a) or the rough periwinkle *Littorina*
106 *saxatilis* (Butlin *et al.* 2014), were mainly characterized by non-parallel genomic signatures of
107 selection.

108

109 Adaptive divergence is thought to proceed as a balance between divergent selection and
110 homogenizing gene flow (Levene 1953; Hagen 1967; Endler 1973; Bell 1982) and hence may
111 reach various stages. This has long been recognised at the phenotypic level (Tregenza 2002;
112 Moore *et al.* 2007; Hendry 2009; Schluter 2009). More recently, theoretical and empirical
113 studies have improved our understanding of the genomic architecture at various stages of
114 adaptation as well (Pinho & Hey 2010; Yeaman & Otto 2011; Yeaman & Whitlock 2011).

115 Nevertheless, the focus of genome scans is often on systems where it is reasonable to
116 assume that population divergence has a strong adaptive component. To do so, genome
117 scans often target populations ~~from highly contrasting environments~~ at large spatial scales
118 and across strong ecological contrasts (Hohenlohe *et al.* 2010; Poncet *et al.* 2010; Stapley *et*
119 *al.* 2010; Zulliger *et al.* 2013). While this has generated great insight in the genomic basis of

120 adaptation and speciation, knowledge might be biased towards stages ~~of population~~
121 ~~divergence with very little gene flow~~ where gene flow may be largely impeded, and where
122 adaptation is already largely achieved. Gene flow modifies the response to selection by
123 modulating the distribution of the genes that underlie ecologically relevant traits (Slatkin
124 1987). The study of adaptation with gene flow, especially in study systems where genetic
125 divergence is far from complete such as across a small-scale ecological transitions or
126 between highly connected populations, is therefore crucial to understand how populations
127 diverge to different ecological optima ~~Less is known about the genomic targets of natural~~
128 ~~selection at the onset of divergence, when gene flow is strong_ (Feder *et al.* 2012). To~~

129 ~~understand how populations diverge to different ecological optima it is crucial to study~~
130 ~~systems where isolation is far from complete. Populations with recurrent gene flow can~~
131 ~~teach us about the generality and constraints of adaptive divergence, and the importance of~~
132 ~~genomic processes at the onset of population divergence~~ (Hansen *et al.* 2002; Storz 2005;
133 Nielsen *et al.* 2009a; Coscia *et al.* 2011; [DeFaveri *et al.* 2013](#); Vandamme *et al.* 2014).

134
135 The three-spined stickleback represents an excellent model for the study of adaptive
136 divergence, as phenotypic responses to several ecological changes are frequent and well-
137 documented (McPhail 1994; Foster *et al.* 1998; McKinnon & Rundle 2002; Boughman 2007).

138 A diverse number of populations can be found that occupy ~~different positions along a~~
139 ~~“speciation continuum” from state 1 where there is continuous gene flow in a panmictic~~
140 ~~population to state 4 where there is various stages of divergence from panmixia to complete~~
141 and irreversible reproductive isolation (Hendry *et al.* 2009; ~~Feder *et al.* 2012~~). This provides
142 an excellent framework to investigate ~~progress towards ecological speciation~~population
143 divergence at the phenotypic and genomic level. The genome scans that have been applied
144 to three-spined stickleback ~~populations cover a broad diversity of habitat pairs across a wide~~
145 ~~geographical range. The habitat pairs have been divided in~~have contrasted marine-fresh
146 water (Mäkinen *et al.* 2008; Hohenlohe *et al.* 2010; DeFaveri *et al.* 2011; Shimada *et al.*
147 2011; Jones *et al.* 2012a; Jones *et al.* 2012b; DeFaveri & Merilä 2013), lake-stream (Deagle
148 *et al.* 2012; Roesti *et al.* 2012), and benthic-limnetic (Olafsdottir & Snorrason 2009; Jones *et*
149 *al.* 2012a; Lucek *et al.* 2014) population pairs, as well as populations from ~~and~~ clean versus
150 polluted water (Lind & Grahn 2011). ~~In fresh-salt water comparisons, the studies only cover~~
151 ~~a limited part of the speciation continuum. They have largely been carried out on~~
152 ~~populations in stark contrasting environments with limited gene flow (Table 1), although~~

153 ~~some global studies have added some populations with recurrent gene flow, (Jones *et al.*~~
154 ~~2012a), like the Scottish river system that harbours a very narrow range of hybrid zone~~
155 ~~(Jones *et al.* 2006) or with a limited set of markers in the Baltic Sea (DeFaveri *et al.* 2013).~~
156 ~~Several studies cover a global or regional scale (Mäkinen *et al.* 2008; DeFaveri *et al.* 2011;~~
157 ~~Jones *et al.* 2012a; Jones *et al.* 2012b; DeFaveri *et al.* 2013) and a larger area than what~~
158 ~~would be considered the dispersal rates typical of any of these populations of three-spined~~
159 ~~sticklebacks. A common finding is that several genes or gene regions are repeatedly selected~~
160 ~~across populations and locations, although population-specific regions do appear as well~~
161 ~~(Hohenlohe *et al.* 2010; DeFaveri *et al.* 2011; Shimada *et al.* 2011; Jones *et al.* 2012b). Others~~
162 ~~have found that most regions under selection were highly specific to the location under~~
163 ~~study (Mäkinen *et al.* 2008; Deagle *et al.* 2012; Roesti *et al.* 2012). ~~Whether any of these~~~~
164 ~~patterns can be attributed to the stage in the speciation continuum was not considered.~~

165

166 In this study we investigate genomic adaptation in three-spined stickleback populations from
167 the Belgian-Dutch lowlands. Populations in this area differ in various morphological traits,
168 which are often correlated with salinity and distance to the coast (Heuts 1947; Raeymaekers
169 *et al.* 2005; Raeymaekers *et al.* 2007; Van Dongen *et al.* 2009; Raeymaekers *et al.* 2012). At
170 the same time, gene flow between these populations is ~~recurrent moderate to strong~~ due to
171 high connectivity (Raeymaekers *et al.* 2014). This has the advantage that we can study
172 adaptation with ongoing gene flow, which is often a realistic setting in nature. All sites are
173 part of an interconnected landscape of canals and streams along an ecological ~~gradient~~
174 ~~transition~~ (i.e. discrete habitats with salinities ranging between brackish and freshwater) at a
175 ~~much comparatively smaller~~ geographical scale ~~than other genome scan studies across~~
176 ~~salinity gradients in three-spined stickleback (Table 1).~~

Formatted: Indent: First line: 0"

177
178 ~~We can thus use this system to deduce which genes are the targets of selection at the~~
179 ~~very onset of population divergence.~~

180 Formatted: Indent: First line: 0.49"

181 ~~In order to explore the adaptive changes across this transition, we screened two populations~~ Formatted: Indent: First line: 0"

182 ~~from each end of the transition for genomic signatures of selection. We hypothesized~~~~We~~

183 ~~performed a genome scan along this ecological transition hypothesising~~ that (i) the outcome

184 of small-scale adaptation along ecological transitions may vary at the genomic level and that

185 (ii) adaptation at the genomic level may be influenced by gene flow. ~~Four populations were~~

186 ~~selected along a salinity cline and screened with~~Populations were screened for 87

187 microsatellite markers of which 41 are linked to genes with a range of ecologically relevant

188 functions (Shikano *et al.* 2010b; Shimada *et al.* 2011). We expect that the relatively high

189 gene flow among these populations might constrain local adaptation, despite obvious

190 differences in phenotype across these extremes. We determine outlier loci and compare the

191 results to previous studies in three-spined stickleback across similar ecological contrasts.

192

193 MATERIAL AND METHODS

194 *Study area*

195 Three-spined stickleback (*Gasterosteus aculeatus* L.; Gasterosteidae) from the coastal

196 lowlands (polder) of Belgium and the Netherlands (Figure 1) reside in ponds, ditches,

197 streams, estuaries or polder creeks. They have an anadromous or landlocked life style (Heuts

198 1947; Wootton 1976; Raeymaekers *et al.* 2005; Raeymaekers *et al.* 2007). The polder and

199 surrounding areas contain diked brackish and freshwater habitats of Holocene origin with

200 varying levels of connectivity to adjacent estuaries and the open sea. Populations which live
201 in close proximity to the sea (< 10 km) reside in brackish water of which the salinity is
202 influenced by rainfall and water management. On a scale of less than 50 km further inland,
203 salinity and associated environmental variables change with increasing distance to the
204 coast to freshwater levels (Raeymaekers *et al.* 2014). ~~Populations which live in close~~
205 ~~proximity to the sea (< 10 km) reside in brackish water of which the salinity is influenced by~~
206 ~~rainfall and water management.~~ Lateral plate number, an important ecological trait, also
207 decreases with distance to the coast, with population averages from 20 to 5 (Heuts 1947;
208 Raeymaekers *et al.* 2014). Populations bordering the North Sea and the Baltic are typically
209 polymorphic for lateral plate number (Heuts 1947; Raeymaekers *et al.* 2014), so higher or
210 lower population averages are rare. This range is therefore representative for the
211 phenotypic extremes we can find in this part of the stickleback's distribution range.

212

213 *Field sampling*

214 Field sampling was conducted in spring 2009 in parallel with a multi-year study by
215 Raeymaekers *et al.* (2014), describing the distribution of lateral plate number in the study
216 area. Fifteen sites inhabited by the three-spined stickleback were monitored on a regular
217 basis during two years for various environmental variables including salinity, pH and water
218 depth (Raeymaekers *et al.* 2014). Two brackish water creeks (L01 and L02) and two
219 freshwater ditches (L12 and U01), representing the sites with the most extreme values for
220 salinity, were selected (Table 2). Thirty fish from each population were collected with dip
221 nets, immediately anaesthetized with MS222 (2 g.l⁻¹) and frozen on dry ice. In the lab,
222 sticklebacks were thawed on ice, measured (standard length (SL); ± 0.1 cm), weighed (± 0.01
223 g), photographed and fin clipped for a DNA sample. ~~Field measurements included~~

224 ~~monitoring the water for pH and conductivity ($\mu\text{S}\cdot\text{cm}^{-1}$; converted to salinity in psu for~~
225 ~~analysis) using a Hach probe (Hach, Loveland, Co, USA). An index for the turbidity of the~~
226 ~~water was obtained with a Sneller tube and water depth was calculated as the median depth~~
227 ~~(cm) of five equidistant points along a stretch of 100 m of the water body. All macro-~~
228 ~~invertebrate predators of three-spined stickleback, including backswimmers (*Notonecta*~~
229 ~~*glauca*), dragonfly larvae (*Anax* sp. and *Aeschna* sp.) and great diving beetles (*Dytiscus*~~
230 ~~*marginalis*) obtained in the nets were counted.~~

231

232 *Morphological measurements*

233 In line with previous studies on phenotypic divergence between the stickleback populations
234 of the Belgian-Dutch lowlands (Heuts 1947; Raeymaekers *et al.* 2005; Raeymaekers *et al.*
235 2007; Van Dongen *et al.* 2009; Raeymaekers *et al.* 2012) we investigated variation at six
236 morphological traits. The left side of each fish was photographed from a standard angle, and
237 a ruler was placed in each photograph for scaling. Dorsal spine length, pelvic spine length
238 and pelvic plate length were measured from pictures using the software tpsDig 1.37 (Rohlf
239 2002). A subsample of the fish was rinsed with water for 72 h, bleached for 4 h (1% KOH
240 bleach solution) and stained with alizarin red (Taylor & Dyke 1985). Stained fish were used to
241 determine the number of lateral plates on the right side of the fish. The right part of the gills
242 was then dissected and the number and length of the large gill rakers were quantified under
243 a dissection scope. ~~Statistical analysis of variance (ANOVA) confirmed the presence of~~
244 ~~significant differences between the four populations for standard length, lateral plate~~
245 ~~number, pelvic plate and spine length and large gill raker length (Figure 2). The most studied~~
246 ~~trait in these populations is the plate morphology which can either be described as low~~
247 ~~plated or high plated. These terms refer to the number of lateral plates the individual has~~

248 ~~with high plated individuals having 20 or more plates, while low plated individual have less~~
249 ~~than 10. It is well established that this trait correlates with the salinity of the water within~~
250 ~~the habitat and that this trait is regulated by a major effect gene called *Eda* and several~~
251 ~~minor effect genes (Colosimo *et al.* 2005). All populations in this study system are~~
252 ~~polymorphic for this trait, but the frequency of the different morphs change considerably~~
253 ~~(Raeymaekers *et al.* 2014).~~

254

255 *Marker selection*

256 A set of 110 microsatellite markers was selected, including a range of putatively neutral
257 markers, to set a proper neutral F_{ST} background. In our study area there is a salinity cline, but
258 ~~numerous various~~ other factors may co-vary with this cline. Therefore we included 41
259 markers that are known to be linked to functional genes in a range of ecologically relevant
260 functions such as salinity, growth and immunity (Shikano *et al.* 2010b; Shimada *et al.* 2011).
261 Several of these genes have been found to be under selection in the three-spined stickleback
262 or nine-spined stickleback of other ~~fresh salt water salinity systems transitions~~ (Hohenlohe *et*
263 *al.* 2010; Shikano *et al.* 2010a; DeFaveri *et al.* 2011; Shimada *et al.* 2011; Jones *et al.* 2012a;
264 Jones *et al.* 2012b). This allows us to compare these systems to some extent with ours and
265 pinpoint parallel changes across systems. The marker closely linked to the *Eda* gene
266 (*Stn380*), a major effect gene underlying variation in plate number (Colosimo *et al.*, 2005),
267 was included as a reference gene that is often under selection in freshwater-salt water
268 comparisons (Raeymaekers *et al.* 2007; Mäkinen *et al.* 2008; Barrett *et al.* 2009; Van Dongen
269 *et al.* 2009; DeFaveri *et al.* 2011; DeFaveri & Merilä 2013).

270

271

Formatted: Dutch (Belgium)

272 *DNA extraction and genotyping*

273 Genomic DNA was extracted from fin clips using a proteinase K digestion step and the
274 Nucleospin 96 Tissue DNA Extraction kit (Macherey-Nagel). Individuals were genotyped at
275 110 microsatellite loci that were arranged in 21 multiplexes of 4 to 8 markers at a time, with
276 EST based markers and markers within or near genes with relevant functions included (Table
277 S1). Loci were amplified with the Qiagen® Multiplex PCR Kit (Qiagen, Venlo, the
278 Netherlands). The 10 µl PCR cocktail contained 1-100 ng genomic DNA, 2 pmol each of
279 forward and reverse primers, 1x Qiagen Multiplex PCR master mix, 0.5x Q-solution, and
280 RNase-free water. The reaction consisted of an initial activation step of 15 min at 95 °C,
281 followed by 30 cycles of 30 s at 94 °C, 90 s at 53 °C and 60 s at 72 °C. A final elongation step
282 of 5 min at 60 °C was performed. Allele sizes were determined by means of an internal ET
283 ROX 550 size standard (Amersham Biosciences). Polymerase chain products were visualized
284 using a MegaBace 1000 automated sequencer (Amersham Biosciences). Alleles were scored
285 with the Fragment Profiler v1.2 software (Amersham Biosciences), using visual scoring and
286 manual corrections. Marker Stn380, linked to the *Eda* gene, was scored separately to
287 determine the frequency of the “low plated” allele in each population (Table 2).

288

289 *Data analysis*290 *Phenotypic differentiation*

291

292 For each trait, analysis of variance (ANOVA) was performed to test for statistical differences
293 between the four populations. For traits which depend on size, standard length was included
294 as a covariate.

295 |
296 *Genetic diversity and genetic differentiation*

297 Genotypes were checked for scoring errors attributable to stutter-products, large allele
298 dropout or null-alleles, using MICRO-CHECKER v2.3 (van Oosterhout *et al.* 2004). Estimates of
299 allelic richness, genetic diversity (H_e , H_o), global and pairwise F_{ST} , with a significance
300 calculated with 1000 bootstraps over loci, were calculated using the GENETIX v4.05.02
301 software (Belkhir *et al.* 1996).

302 *Genomic signatures of selection*

303 | We conducted global outlier tests ~~using our entire dataset~~ to find outliers across all
304 populations and used pairwise comparisons to check whether the outliers found by global
305 tests ~~could be confirmed and~~ were due to habitat differences. Loci that are under directional
306 selection are expected to have lower intrapopulation variability and larger interpopulation
307 variability than neutral loci. Loci under directional selection can thus be traced by patterns in
308 heterozygosity, differences in F_{ST} values or a combination of the two. ~~We used several~~
309 ~~outlier detection approaches as e~~ Evaluation of several outlier detection methods has shown
310 that the se methods differ in number of false positives and false negatives (Narum & Hess
311 2011). ~~Therefore we implemented and compared the results from four different outlier~~
312 ~~detection tests: LOSITAN (Antao *et al.* 2008), the outlier detection method within ARLEQUIN~~
313 ~~v3.5.2.3 (Excoffier & Lischer 2010), hierarchical clustering and BAYESCAN v2.01 (Foll &~~
314 ~~Gaggiotti 2008) LnRH (Kauer *et al.* 2003). Four methods were therefore compared: LOSITAN~~
315 ~~(Antao *et al.* 2008), the outlier detection method implemented in ARLEQUIN v3.5.2.3 using~~
316 ~~hierarchical clustering (Excoffier & Lischer 2010), BAYESCAN v2.01 (Foll & Gaggiotti 2008),~~
317 ~~and LnRH (Kauer *et al.* 2003).~~ The first three methods ~~were used~~ were used to determine
318 global outliers. Additionally we did pairwise comparisons of all populations using LOSITAN
319 and LnRH to detect specific ~~outliers signatures of selection~~ in fresh water-salt-brackish water

320 ~~population pairs~~, fresh~~water~~-freshwater and ~~saltbrackish water-salt-brackish~~ water
321 populations pairs.

322 The four methods are based on different underlying assumptions. LOSITAN is based on an
323 island model that uses a coalescent F_{ST} -outlier method based on the distribution of F_{ST} as a
324 function of the heterozygosity. We used the function that first establishes a neutral F_{ST}
325 baseline by removing putative markers under selection outside the 95% interval with 10^5
326 simulations. The infinite allele model was used with a 95% and 99% confidence interval. We
327 ran 10^5 simulations as recommended by Antao (2008). The outlier detection software
328 implemented in ARLEQUIN uses the same island model, but adds on the option for
329 hierarchical clustering. In the presence of strong hierarchical population structure it reduces
330 false positives by a hierarchical analysis of genetic differentiation (Excoffier *et al.* 2009). We
331 clustered populations ~~for according to two probable scenarios~~: in two groups (brackish
332 water populations versus freshwater populations), and three groups (brackish water
333 populations versus each freshwater population separately). The first scenario simulates a
334 common descent of the two freshwater populations from the marine population, while the
335 second one simulates a separate split of the two freshwater populations from the marine
336 population. We used the standard settings of 20,000 simulations for each run and 100
337 demes per group. The method executed by BAYESCAN uses a logistic regression model
338 ~~described by Beaumont and Balding (2004)~~, which explains the observed pattern of diversity
339 by dividing it in a locus- and a population-specific component— ~~(Beaumont and Balding~~
340 2004). One benefit of this method is that it allows for different migration rates and different
341 effect sizes and thus can be used for scenarios that deviate from the island model. We
342 conducted 10 pilot runs of 5,000 iterations, followed by an additional 150,000 iterations and
343 a burn-in of 50,000 iterations. Outliers were appointed based on 90%, 95% and 99%

344 | posterior probabilities. ~~Finally, The the~~ LnRH method is designed especially for microsatellite
345 | markers and determines the reduction of heterozygosity. This method is based on the
346 | assumption that microsatellites linked to a gene under selection will show reduced levels of
347 | diversity between two populations. After standardization of the LnRH estimates with a mean
348 | of zero and a standard deviation of one, we determined outliers at 95% and 99% level. We
349 | subtracted outliers that were found in pairwise comparisons within the same habitat type,
350 | to limit the number of false positives.

351 | All tests based on simulations were executed three times to test for robustness of the
352 | results. Only outliers that were detected ~~in two out of three trials within each method at~~
353 | least twice with each method were scored as a putative outlier. We detected outlier loci that
354 | are under balancing and directional selection. However, interpreting loci under balancing
355 | selection is difficult, as there are still limitations for the identification of loci under balancing
356 | selection (Hansen *et al.* 2010; Narum & Hess 2011). Therefore we only discuss ~~only~~ the loci
357 | under directional or positive selection. Loci under balancing selection are provided in
358 | Supplementary Table S2.

359

360 | RESULTS

361 | Phenotypic differentiation

362 | Statistical analysis of variance (ANOVA) confirmed the presence of significant differences
363 | between the four populations for standard length, lateral plate number, pelvic plate and
364 | spine length, and large gill raker length (Figure 2). In particular, the brackish water
365 | populations (L01 and L02) were larger in size than the freshwater populations (L12 and U01;
366 | $F_{3,155} = 75.6$; $P < 0.0001$), and had more lateral plates ($F_{3,147} = 10.4$; $P < 0.0001$), longer pelvic

367 spines (especially population L01; $F_{3,93} = 5.51$; $P = 0.0016$), longer pelvic plates ($F_{3,93} = 8.60$; P
368 < 0.0001), and larger gill rakers ($F_{3,93} = 19.8$; $P < 0.0001$).

369

370 *Genetic diversity and genetic differentiation*

371 After genotyping of the individuals with 110 microsatellite markers, we selected 87 markers
372 with good amplification quality to perform the data analysis. A total of 1116 alleles were
373 observed in four populations of 26 to 28 individuals each, with an average of 13 alleles per
374 locus and a range from 2 to 34. Observed heterozygosity ranged from 0.022 to 0.95 across
375 loci. ~~and e~~Expected heterozygosity ranged from 0.05 to 0.92 across loci and from 0.57 to
376 0.69 across populations (Table 2). Allelic richness among populations varied from 6.51 to
377 9.47, with the freshwater populations being less diverse than the brackish water
378 populations. Eight loci were possibly affected by null alleles or stutters. We therefore
379 performed all final analyses with and without these loci and specify when results differ.
380 Differentiation among populations was moderate with the global F_{ST} value being 0.059.
381 Pairwise F_{ST} values were significant, except between the two brackish water populations
382 (Table 3).

383

384 *Genomic signatures of selection*

385 Six of the 87 loci were assigned as outliers in at least one of the methods, with the different
386 methods identifying 4, 2, 2, and 4 outliers for Lositan, Bayescan, InRH and Arlequin ~~(with~~
387 ~~hierarchical nesting)~~ respectively (Table 4). The methodologically similar tests performed by
388 Lositan and Arlequin with hierarchical ~~grouping~~ clustering resulted in the same set of
389 outliers. In contrast, only two outliers were shared across methodologically different outlier
390 detection methods; our conservative measure of outliers hence amounts to ~~2.2~~ 3 % of the

391 total number of loci. This value is low in comparison with other ~~outlier-genome scan~~ studies
392 ~~of~~ of three-spined stickleback ~~populations~~ (Table 1).

393

394 The two outliers for directional selection that appear consistently across methodologically
395 different tests are marker *Stn46* (identified by Arlequin/Lositan and Bayescan) and *Ppgm44*
396 (identified by InRH and Bayescan). *Stn46* has been previously associated with the Rho
397 guanine nucleotide exchange factor 9 (Table S1). ~~In pairwise comparisons it is was also~~
398 assigned as an outlier in ~~only one of the pairwise~~ freshwater versus salt-brackish water
399 comparisons. ~~In InRH it is not assigned as an outlier.~~ *Ppgm44* is a marker that is linked to the
400 gene *myostatin2MSTNb*, which is associated with growth (Table S1). ~~In pairwise~~
401 ~~comparisons of InRH, PPGM44 is found in all the~~ was also assigned as an outlier in all four
402 pairwise ~~fresh-water~~ versus salt-brackish water comparisons. ~~In LOSITAN the same was~~
403 ~~found in pairwise comparisons, yet PPGM44 was not assigned as outlier in the global~~
404 ~~analysis. Of Among~~ the four other ~~markers-outliers~~ that ~~are-were~~ detected (Table 4), one is
405 associated with osmoregulation, namely an alpha subunit of the Na⁺/K⁺ ATPase (ATP1A1).

406 Other outliers are either linked to functions ~~like-such as~~ thermal response (HSPA14) or ~~are~~
407 ~~were~~ assumed to be neutral. ~~The outlierLocus~~ *Stn34* was also assigned as an outlier, but ~~this~~
408 ~~since might be caused by~~ null alleles were detected at this locus, it was and was deleted
409 excluded from Table 4. The *Eda* gene, with frequencies of the low-plated allele ranging
410 between 0.44-0.51 in the brackish populations and 0.74-0.93 in the freshwater populations
411 (Table 2), did not show up as an outlier in the analyses (Table 4).

412

413 DISCUSSION

414 We found ~~that~~ a limited ~~proportion number~~ of loci ~~show showing~~ a signature of selection
415 among three-spined stickleback populations inhabiting the coastal Dutch-Belgian lowlands
416 ~~(Table 4)~~. Six outliers were found when all outlier detection methods were considered and
417 only two ~~out of those six~~ ~~of those~~ were shared among methodologically different tests,
418 despite differentiation among populations in ecology and phenotype. Not only the number
419 but also the proportion of outliers was low as compared to other studies (Table 1). The
420 function of one of the outlier loci ~~can be~~ could be directly related to salinity, the most
421 obvious ecological gradient in our study area. We here discuss possible explanations for the
422 limited amount of outliers, the putative function of the genes linked to outliers, and ~~what~~
423 ~~this tells us about the consequences for genetic differentiation in populations with~~
424 ~~concurrent adaptive divergence in the face of~~ gene flow.

425
426 The presence of a limited ~~amount number and proportion~~ of outliers might be attributed to
427 ~~several a number of~~ technical aspects. First, we used relatively few markers. However,
428 among the 87 markers selected, 41 markers are linked to ecologically relevant functions, of
429 which several have been shown to be under selection ~~in other fresh salt water in~~ systems
430 with similar salinity clines (Mäkinen *et al.* 2008; Shimada *et al.* 2011; DeFaveri *et al.* 2013).
431 ~~Additionally, not only the number, but also the proportion of loci under selection was low~~
432 ~~compared to other studies, indicating that the number of markers did not cause this limited~~
433 ~~amount of outliers~~. Second, analysing relatively few populations may have two drawbacks:
434 1) it might increase the risk of ~~not detecting all possible genes under selection false~~
435 negatives due to the limited number of individuals sampled, and 2) it might increase the
436 difficulty in separating the signal of selection from the geographical and historical signal

437 (Bierne *et al.* 2011; Bierne *et al.* 2013) ~~as we miss the possibility to confirm outliers across~~
438 ~~more than two population pairs.~~ The small overlap in outlier loci among detection methods
439 might be a symptom of this. However, in a larger ~~set study of across~~ 14 populations ~~(and~~
440 2320 individuals), ~~describing the distribution of lateral plate number and the underlying *Eda*~~
441 ~~gene in from~~ our study area (Raeymaekers *et al.* 2014), the signature of selection at the *Eda*
442 gene was ~~equally weak neither or significant either.~~ This bi-allelic gene is experiencing
443 ~~selection across various other salinity transitions in three-spined stickleback (Raeymaekers~~
444 ~~et al. 2007; Mäkinen et al. 2008; DeFaveri et al. 2011; DeFaveri & Merilä 2013) (Raeymaekers~~
445 ~~et al. 2014).~~ ~~This it shows~~ This suggests that increasing sample size does not necessarily
446 enhance ~~detecting the signature the detectability~~ of selection. We therefore do not expect
447 that technical issues are a major explanation for our findings.

448
449 ~~There might be several other reasons~~ An alternative explanation for the low proportion of
450 outlier loci, ~~despite the established knowledge from other salt-freshwater comparisons in~~
451 ~~the three-spined stickleback that several genes can be under natural selection in these~~
452 ~~contrasting environments (Hohenlohe et al. 2010; DeFaveri et al. 2011; Shimada et al. 2011;~~
453 ~~Jones et al. 2012a; Jones et al. 2012b).~~ ~~S is that~~ selection might be weak, owing to the
454 ~~relatively weak weaker~~ environmental contrasts in our study area (i.e., brackish to
455 ~~freshwater) compared to other studies (Table 1). For instance, Raeymaekers et al. (2014)~~
456 ~~found that Raeymaekers et al. (2014) identified only a weak signature of selection in our~~
457 ~~study area on *Eda*, which is a bi-allelic gene underlying variation in lateral plate number and~~
458 ~~experiencing selection in many other systems. In line with this we find no evidence for~~
459 ~~selection on the marker that is linked to *Eda*. Nevertheless, the frequency of the *Eda* allele~~
460 ~~accounting for low plate number showed a significant increase with distance from the coast,~~

461 ~~and~~ shifts in *Eda* allele frequencies from one generation to the next were ~~linked-associated~~
462 with salinity ~~(Raeymaekers et al. 2014), but that at the landscape level salinity did not~~
463 ~~correlate with *Eda* allele frequencies~~. This suggests that selection is still acting, but might be
464 too weak to contribute to local adaptation and leave a signature of selection at this gene.
465 ~~Additionally~~In addition, strong gene flow might confound the effect of selection by mixing
466 adapted and non-adapted alleles in the respective populations. Gene flow being moderate
467 to high, we expect this to be ~~one-another~~ explanation for why we find so few consistent
468 outlier loci compared to other studies ~~(see Table 1)~~. Accordingly, Raeymaekers et al. (2014)
469 found that the spatial distribution of the *Eda* allele frequency correlated with distance to the
470 coast, a proxy for population connectivity. Recurrent contact between freshwater and
471 estuarine or marine populations might lead to the exchange of maladaptive alleles, but at
472 the same time it has been argued that gene flow might cause an opposite effect: efficient
473 flow of advantageous loci (Schluter & Conte 2009; Hohenlohe et al. 2012; Bell & Aguirre
474 2013; Raeymaekers et al. 2014), thus facilitating adaptation from standing variation.
475 Whether gene flow in this case fuels or constrains adaptation is hard to say, but the signal of
476 selection might be more difficult to pick up in genome scans due to mixing of the genomic
477 background. ~~Finally, population differentiation among the brackish water populations was~~
478 ~~very low, which could arguably lead to the conclusion that these are in fact one population.~~
479 ~~However, this is a feature of many marine populations where connectivity is high and~~
480 ~~populations structure thus low, and has been shown neither to prevent the detection of~~
481 ~~genomic regions under selection (Nielsen et al. 2009b; DeFaveri et al. 2013; Feulner et al.~~
482 ~~2013; Milano et al. 2014; Vandamme et al. 2014) nor can it be assumed a priori that these~~
483 ~~populations have not diverged in adaptive loci (Pinho & Hey 2010).~~ We therefore expect that

484 the low number of outlier loci is rather explained by either weak selection, confounding
485 effects of gene flow or a combination of both and are thus inherent to the system we study.

486

487 Genes that have been found in previous studies to be under selection relate to biological
488 functions such as bone formation, osmoregulation, growth, thermal response, maturation,
489 pigmentation, scent detection, spiggin production and morphology (Hohenlohe *et al.* 2010;
490 DeFaveri *et al.* 2011; Shimada *et al.* 2011; Jones *et al.* 2012b). Genes that appeared as
491 outliers in this study include markers linked to osmoregulation (*ATP1A1*), thermal response

492 (*HSPA14*) and growth (*myostatin2MSTNb*). The marker linked to *ATP1A1* has been identified
493 as outlier gene in multiple stickleback saltwater-freshwater transitions worldwide (Jones *et al.*
494 *et al.* 2006; McCairns & Bernatchez 2010; DeFaveri *et al.* 2011; Shimada *et al.* 2011; Jones *et al.*

495 2012a; Jones *et al.* 2012b). ~~One of the two markers pinpointed as outliers in more than one~~
496 ~~detection method~~ A major outlier in our study was marker *Stn46*, which is linked to the gene
497 for Rho guanine nucleotide exchange factor 9. It is a member of the gene family coding for

498 rho proteins, a subfamily of the guanine nucleotide exchange factors (GEFs), which are
499 multi-domain proteins involved in the activation of small GTPases (Rossman *et al.* 2005). The
500 Rho family is involved in relaying signals from cell-surface receptors to the actin cytoskeleton

501 and elsewhere (Dvorsky & Ahmadian 2004). Its function can be associated with juvenile
502 growth; in zebrafish it has been specifically linked to angiogenesis (Garnaas *et al.* 2008) and

503 striated muscle and neural development (Raeker *et al.* 2010). Interestingly, locus marker
504 *Stn46* was also under selection in populations of the nine-spined stickleback (Shikano *et al.*
505 2010a), indicating that the gene might be involved in local adaptation in multiple species.

506 ~~The other outlier detected in more than one method is~~ Another major outlier in our study
507 was the marker linked to ~~the gene *myostatin2MSTNb*. *Myostatin-b*,~~ This gene, a member of

508 the transforming growth factor-beta (TGF-beta) family, is known to function as a negative
509 regulator of skeletal muscle development and growth in mammals (Walsh & Celeste 2005)
510 and teleost fish (Radaelli *et al.* 2003). Blocking the expression of myostatin in zebrafish has
511 led to the development of a giant phenotype (Acosta *et al.* 2005), but myostatin is
512 produced in many other tissues than skeletal muscles and is expected to influence many
513 more functions (Radaelli *et al.* 2003). Understanding why these specific genes are selected
514 requires further study due to the broad range of functions these genes might have.

515 ~~Interestingly, marker Stn46 was also under selection in populations of the nine-spined~~
516 ~~stickleback (Shikano *et al.* 2010a), indicating that the gene might be involved in local~~
517 ~~adaptation in multiple species.~~

518

519 The low number of outlier loci that we find contrasts with the differentiation across
520 populations in several morphological traits. Local differentiation in the number of lateral
521 plates, for instance, has been shown to significantly exceed the level of neutral
522 differentiation in our study area (Raeymaekers *et al.* 2014). Plate number and other
523 morphological traits such as spine length and gill raker length have reasonably high
524 heritability values exceeding 40 % (Schluter 1996; Peichel *et al.* 2001; Berner *et al.* 2014),
525 suggesting that phenotypes are largely determined by genetic rather than by plastic effects.

526 Yet, theory predicts that only functional loci with a relatively large effect size under strong
527 divergent selection will be able to surpass gene flow (Via 2009; Yeaman & Whitlock 2011).

528 The discrepancy between phenotypic and genomic signatures of selection might therefore

529 become particularly strong for traits that involve many genes of small effect. ~~That the~~
530 ~~signature of selection on a major effect locus such as the *Eda* gene is not detected in our~~
531 ~~study is unexpected, but might be partially due to the limited power of outlier detection~~

532 | ~~tests.~~ Arnegard *et al.* (2014) have another ~~possible~~ explanation for the relatively large
533 phenotypic divergence. They show that niche differentiation in sticklebacks, even in early
534 stages of differentiation, can involve many different genes and that gene flow between
535 divergent niche-adapted populations has a bigger impact on the phenotype than just the
536 traits that are directly targeted by selection. It may be attributed to incompatibilities in
537 hybrids that harbor a mix of genes of differentially adapted genotypes. These effects imply
538 that phenotypic changes might not always be adaptive. McCairns & Bernatchez (2010) found
539 indications that freshwater populations might suffer from a loss of plasticity and it might be
540 that epigenetic effects further enhance this discrepancy. For instance, Chaturvedi *et al.*
541 | (2014) found that regulation by miRNAs might ~~have~~ make a significant contribution to an
542 important function in the salt to freshwater adaptation in stickleback populations.

543

544 | ~~When taking the different states of ecological speciation into account, we expect a~~
545 ~~large range of possible outcomes and degrees of differentiation in both strength of genetic~~
546 ~~and morphological divergence and associated genes. In the three-spined stickleback, this has~~
547 ~~already been shown in lake-river habitats (Berner *et al.* 2009; Berner *et al.* 2010). Seehausen~~
548 ~~*et al.* (2014) pointed out that studying population differentiation and estimating the effect of~~
549 ~~gene flow will teach us a lot about the order in which reproductive barriers emerge.~~
550 ~~Furthermore, patterns of divergence might be substantially different from study systems~~
551 ~~that experience recurrent gene flow. Other genes might be selected due to divergence~~
552 ~~hitchhiking (Via *et al.* 2012), resulting in a mosaic pattern of genomic divergence. Genome~~
553 ~~scans might seem useful to discover these genes, but our study shows that it is challenging~~
554 ~~to detect the outliers in the first stages of divergence due to the inherent nature of these~~

555 ~~systems. Yet, others have shown it is possible (Rogers & Bernatchez 2007; Egan et al. 2008;~~
556 ~~Michel et al. 2010) and the relevance for understanding the mechanism of ecological~~
557 ~~speciation is high.~~

558 Conclusion

559 ~~As expected, w~~We find that weak selection, high levels of gene flow or a combination of
560 both can limit the number of outliers in genome scans. Although genome scans targeting
561 populations across ~~starker~~ strong environmental contrasts are possibly more effective for
562 pinpointing genes that are involved in adaptation, the genes identified by these studies do
563 not necessarily play an important role at every stage of divergence. Many of the genes
564 involved in salt-freshwater transitions might be site-specific or might not be involved ~~in early~~
565 ~~stages of divergence~~ when gene flow is constantly mixing the gene pool. The genes that we
566 do find are likely to be those with a major effect size and thus an underrepresentation of the
567 total number of genes involved. In addition, phenotypic adaptation is not necessarily
568 genetic, but might be facilitated by plastic and epigenetic effects. It remains a challenging
569 task to find which genes and how many are truly involved ~~at the onset of population~~
570 ~~divergence~~ in local adaptation. We here showed that even with ample gene flow and across
571 weak ecological contrasts, ~~but we show that these systems reveal~~ interesting insights on
572 the repeatability in the of genomic signatures of adaptation selection across systems, even
573 at the very onset of divergence can be obtained.

574

575 ACKNOWLEDGEMENTS

576

577 We would like to thank A. Chaturvedi, S. Geldof, B. Hellemans, L. Jaquemin, S. Ruyts, N.
578 Thiercelin, I. Verdonck and D. Verheyen for field support and technical assistance. We thank
579 I. Coscia and J. Merilä for insightful and inspiring comments and the latter for collaboration
580 on the whole project. We appreciated the insightful suggestions of two anonymous
581 reviewers. Research was sponsored by the Research Foundation - Flanders (post-doctoral
582 fellowship to J.A.M.R. and research project G.0142.03), by the KU Leuven (GOA/2006/06
583 “Evolutionary responses to strong selection pressures and their ecological consequences”;
584 PF/2010/07 “Eco- and socio-evolutionary dynamics”) and facilitated by the research
585 community W0.037.10N “Eco-evolutionary dynamics in natural and anthropogenic
586 communities” sponsored by the Research Foundation - Flanders. The study received
587 additional support from Academy of Finland grant (129662, 134728 and 218343) to Juha
588 Merilä.

589 TABLES

590 **Table 1.** Comparison of microsatellite-based genome scan studies in three-spined stickleback
 591 across freshwater - brackish/salt water gradients, including spatial scale (from regional to
 592 global), percentage of conservative outliers, F_{ST} , F_{ST} at the *Eda* locus, and F_{ST} at the *ATP1A1*
 593 locus.
 594

Study	Spatial scale	% outliers	F_{ST}	F_{ST} <i>Eda</i>	F_{ST} <i>ATP1A1</i>
This study	78 km	2.3	0.059	0.118	0.154
DeFaveri <i>et al.</i> 2013	Regional	12.5	0.008	-	0.016
Shimada <i>et al.</i> 2011	Continental	8.3	0.107	-	0.225
Mäkinen <i>et al.</i> 2008	Continental	3.0	0.166	0.653	-
DeFaveri <i>et al.</i> 2011	Global	33	0.119	0.405	-

595

596

597 **Table 2.** Characteristics of the four three-spined stickleback populations used in this study. DTC: distance to coast; MPN: Mean plate
598 number; *Eda^L*: frequency of *Eda* low plated allele; AR: allelic richness; *H_e*: expected heterozygosity; *F_{IS}*: inbreeding coefficient, with
599 confidence interval (CI 95%).
600

Code	Coordinates (N, E)	Habitat	DTC (km)	Salinity (psu)	MPN	<i>Eda^L</i>	AR	<i>H_e</i>	<i>F_{IS}</i> (CI 95%)
L01	51°21'10.66, 3°26'01.83	Creek	3.94	2.04	17.4	0.44	9.47	0.69	0.032 (-0.017 - 0.033)
L02	51°21'56.33, 3°31'11.09	Creek	4.30	1.83	16.8	0.51	9.06	0.69	0.095 (0.048 - 0.096)
L12	51°10'29.03, 3°28'10.45	Stream	22.84	0.32	12.7	0.74	6.80	0.61	0.044 (-0.012 - 0.050)
U01	51°02'48.40, 3°33'01.60	Ditch	36.20	0.52	7.0	0.93	6.51	0.57	0.039 (-0.010 - 0.040)

601 **Table 3.** Pairwise F_{ST} values below the diagonal, and associated P-value above the diagonal. For
602 site codes see Table 2.

Code	L01	L02	L12	U01
L01	-	0.600	< 0.010	< 0.010
L02	0.005	-	< 0.010	< 0.010
L12	0.054	0.056	-	< 0.010
U01	0.094	0.097	0.047	-

603

604 **Table 4.** Global outlier loci that are putatively under directional selection detected in four
 605 populations of three-spined stickleback. The table includes the locus considered, its F_{ST} -value
 606 and either the q value for Bayescan or the P-value for LnRH, LOSITAN and Arlequin. Significance
 607 is marked as † = $q < 0.1$, †† = $q < 0.05$ and ††† = $q < 0.01$ or as * = $p < 0.05$, ** = $p < 0.01$ and ***
 608 = $p < 0.001$. Marker *Stn380*, linked to the *Eda* gene, was added as a reference.

Locus	F_{ST}	Bayescan	LnRH	LOSITAN	Arlequin (2 groups)	Arlequin (3 groups)
<i>ATP1A1</i>	0.149			*	*	*
<i>HSPA14</i>	0.159			*	*	*
<i>Ppgm44</i>	0.120	†	***			
<i>Stn74</i>	0.083		*			
<i>Stn177</i>	0.189			**	**	**
<i>Stn46</i>	0.199	†††		**	**	**
<i>Stn380</i> (<i>Eda</i>)	0.118					

609

610 **Supplementary Table S1.** Characterisation of microsatellite loci of three-spined stickleback,
 611 including name, type of marker, gene assignment and putative function.

Locus	Type of marker	Gene	Putative function
<i>1097PBBE</i>	Random genomic	Novel gene	
<i>1125PBBE</i>	Random genomic	Novel gene	
<i>7033PBBE</i>	Random genomic	Integrin beta-3 precursor	
<i>7080PBBE</i>	Random genomic	Neurensin-1	
<i>ATP1A1</i>	Gene-based	Na ⁺ /K ⁺ -ATPase alpha-subunit isoform 1	Osmoregulation
<i>ATP4A</i>	Gene-based	H ⁺ /K ⁺ -ATPase alpha subunit	Osmoregulation
<i>CLCN7</i>	Gene-based	CLC chloride channel isoform 7	Osmoregulation
<i>DCT</i>	Gene-based	L-dopachrome tautomerase precursor	Pigmentation
<i>FGF18</i>	Gene-based	Fibroblast growth factor isoform 18	Growth
<i>FGF8</i>	Gene-based	Fibroblast growth factor isoform 8	Growth
<i>GAest11</i>	Random EST		
<i>GAest16</i>	Random EST	Novel gene	
<i>GAest3</i>	Random EST	Dihydropyrimidinase-related protein 2	
<i>GAest34</i>	Random EST	6-phosphofructokinase type C	
<i>GAest35</i>	Random EST	Rod outer segment membrane protein 1	
<i>GAest41</i>	Random EST	Tumor necrosis factor receptor superfamily member 19 precursor	
<i>GAest47</i>	Random EST		
<i>GAest50</i>	Random EST	Eukaryotic translation initiation factor 3 subunit K	
<i>GAest51</i>	Random EST	Protein kinase C delta-binding protein	
<i>GAest6</i>	Random EST	Ninjurin-2	
<i>GAest67</i>	Random EST	Novel gene	
<i>GAest7</i>	Random EST	Ras-related protein Rab-3A	
<i>GAest82</i>	Random EST	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	
<i>GAest84</i>	Random EST		
<i>GAest87</i>	Random EST		
<i>GS1</i>	Gene-based	Glycogen synthase subtype 1	Growth
<i>HSP70Ab</i>	Gene-based	Heat shock protein 70 kDa subtype 2	Thermal response
<i>HSP70Ad</i>	Gene-based	Heat shock protein 70 kDa subtype 4	Thermal response

<i>IGF-II</i>	Gene-based	Insulin-like growth factor II precursor	Growth
<i>KCNJ4b</i>	Gene-based	Inward rectifier potassium channel subfamily J member 4 subtype 2	Osmoregulation
<i>MYHb</i>	Gene-based	Myosin heavy chain subtype 2	Thermal response
<i>NHE2c</i>	Gene-based	Na ⁺ /H ⁺ exchanger subtype 3	Hypoxia
<i>NHE3</i>	Gene-based	Sodium/hydrogen exchanger isoform 3 (SLC9A3)	Osmoregulation
<i>NPY2Rb</i>	Gene-based	Neuropeptide Y receptor Y2 subtype 2	Growth
<i>PER1</i>	Gene-based	Period circadian protein homolog 1	Maturation
<i>Ppgm44</i>	Gene-based	Myostatin subtype 2	Growth
<i>Ppgm17</i>	Gene-based	Type I iodothyronine deiodinase	Morphology
<i>Ppgm16</i>	Gene-based	Cold shock protein 70 kDa subtype 2	Thermal response
<i>Ppgm24</i>	Gene-based	Growth hormone receptor 1	Growth
<i>Ppgm28</i>	Gene-based	General transcription factor IIB	Osmoregulation
<i>Ppgm58</i>	Gene-based	Tristetraprolin	Disease
<i>Ppgm21</i>	Gene-based	Fibroblast growth factor 18	Growth
<i>Ppgm47</i>	Gene-based	Na ⁺ /K ⁺ /2Cl ⁻ -cotransporter isoform 1 subtype 2	Osmoregulation
<i>Ppgm35</i>	Gene-based	Heat shock protein 70 kDa isoform beta	Thermal response
<i>Ppgm7</i>	Gene-based	Na ⁺ /K ⁺ -ATPase alpha-subunit isoform 2	Osmoregulation
<i>Ppgm56</i>	Gene-based	Amine-associated receptor	Taste
<i>Ppgm1</i>	Gene-based	Pituitary adenylate cyclase activating polypeptide receptor subtype 1	Growth
<i>Ppgm2</i>	Gene-based	Pituitary adenylate cyclase activating polypeptide receptor subtype 2	Growth
<i>Ppgm49</i>	Gene-based	Neuropeptide Y precursor	Growth
<i>Ppgm30</i>	Gene-based	Heat shock protein 25 kDa	Thermal response
<i>Ppgm33</i>	Gene-based	Heat shock protein 70 kDa isoform alpha subtype 2	Thermal response
<i>Ppgm50</i>	Gene-based	Parvalbumin subtype 2	Thermal response
<i>Ppgm6</i>	Gene-based	Na ⁺ /K ⁺ -ATPase alpha-subunit isoform 1	Osmoregulation
<i>Ppgm38</i>	Gene-based	Insulin-like growth factor I	Growth
<i>PRL-R</i>	Gene-based	Prolactin receptor a	Osmoregulation
<i>PVALBa</i>	Gene-based	Parvalbumin subtype 1	Growth
<i>RhBG</i>	Gene-based	Rhesus glycoprotein isoform 3	Toxic stress
<i>SLC4A10a</i>	Gene-based	Electroneutral Na-coupled HCO ₂ transporter isoform 10 subtype 1	Hypoxia
<i>Stn102</i>	Random genomic	EF-hand domain-containing family member A2	
<i>Stn108</i>	Random genomic	Bifunctional heparan sulfate N-deacetylase/N-sulfotransferase 4	

<i>Stn119</i>	Random EST		
<i>Stn122</i>	Random EST		
<i>Stn124</i>	Random EST	Brain-specific angiogenesis inhibitor 2 precursor	
<i>Stn163</i>	Random EST	Novel gene	
<i>Stn164</i>	Random EST		
<i>Stn167</i>	Random genomic		
<i>Stn177</i>	Random genomic		
<i>Stn194</i>	Random genomic	Nuclear-localized factor 1	
<i>Stn20</i>	Random EST		
<i>Stn222</i>	Random genomic	Ras-related protein Rab-18	
<i>Stn223</i>	Random genomic	Zinc finger homeobox protein 4	
<i>Stn235</i>	Random EST		
<i>Stn240</i>	Random EST		
<i>Stn257</i>	Random EST	Solute carrier family 35 member F2	
<i>Stn299</i>	Random genomic		
<i>Stn302</i>	Random genomic		
<i>Stn315</i>	Random genomic	Histone deacetylase 4	
<i>Stn328</i>	Random genomic	RING finger protein 220	
<i>Stn34</i>	Random EST		
<i>Stn380</i>	Gene-based	Eda	plate nr
<i>Stn46</i>	Random genomic	Rho guanine nucleotide exchange factor 9	
<i>Stn49</i>	Random genomic	Dedicator of cytokinesis protein 4	
<i>Stn71</i>	Random EST	Novel gene	
<i>Stn74</i>	Random EST		
<i>Stn79</i>	Random genomic	Reticulon-2	
<i>Stn89</i>	Random EST	Rho GTPase-activating protein SYDE2	
<i>HPX</i>	Gene-based	Hemopexin	Immune response

612

613

614 **Supplementary Table S2.** Global outlier loci that are putatively under balancing selection as
 615 detected with the methods of Bayescan, LOSITAN and Arlequin. Significance is marked as † = $q <$
 616 0.1, †† = $q < 0.05$ and ††† = $q < 0.01$ or * = $p < 0.05$, ** = $p < 0.01$ and *** = $p < 0.001$.

Locus	F_{ST}	Bayescan	LOSITAN	Arlequin (2 groups)	Arlequin (3 groups)
<i>FGF8</i>	0.022		*		
<i>GAest3</i>	0.021		*	*	
<i>GAest34</i>	0.015	††	***	*	*
<i>GAest47</i>	0.009		*	*	*
<i>IGF-II</i>	-0.007		*		
<i>NHE2c</i>	0.027		*	*	
<i>NHE3</i>	0.004		*	*	*
<i>Ppgm16</i>	-0.009		*		
<i>Ppgm2</i>	0.005			**	*
<i>Ppgm30</i>	0.013		*	*	*
<i>Ppgm6</i>	0.005			*	*
<i>Stn122</i>	0.014		***	**	*
<i>Stn20</i>	0.017		*		
<i>Stn223</i>	0.019	†	**	*	*
<i>Stn79</i>	-0.020		***	*	*
<i>Stn89</i>	0.028	†††			

617

618 REFERENCES

- 619 Acosta J, Carpio Y, Borroto I, González O, Estrada MP (2005) Myostatin gene silenced by RNAi show a
620 zebrafish giant phenotype. *Journal of Biotechnology* **119**, 324-331.
- 621 Antao T, Lopes A, Lopes R, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular
622 adaptation based on a Fst-outlier method. *BMC Bioinformatics* **9**, 323.
- 623 Arnegard, M. E., McGee, M. D., Matthews, B., Marchinko, K. B., Conte, G. L., Kabir, S., Bedford, N.,
624 Bergek, S., Chan, Y. F., Jones, F. C., Kingsley, D. M., Peichel, C. L. & Schluter, D. 2014. Genetics of
625 ecological divergence during speciation. *Nature* **511**: 307-311.
- 626 ~~Barrett, R. D. H., Rogers, S. M. & Schluter, D. 2009. Environment specific pleiotropy facilitates divergence
627 at the Ectodysplasin locus in threespine stickleback. *Evolution* **63**: 2831-2837.~~
- 628 Barrick JE, Lenski RE (2013) Genome dynamics during experimental evolution. *Nature Reviews: Genetics*
629 **14**, 827-839.
- 630 Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from
631 genome scans. *Molecular Ecology* **13**, 969-980.
- 632 Becks L, Ellner SP, Jones LE, Hairston NG (2012) The functional genomics of an eco-evolutionary feedback
633 loop: linking gene expression, trait evolution, and community dynamics. *Ecology Letters* **15**, 492-
634 501.
- 635 Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996) GENETIX 4.05, logiciel sous Windows TM
636 pour la génétique des populations, pp. 1996-2004. Laboratoire génome, populations,
637 interactions, CNRS UMR.
- 638 Bell MA (1982) Differentiation of adjacent stream populations of threespine sticklebacks. *Evolution* **36**,
639 189-199.
- 640 Bell MA, Aguirre WE (2013) Contemporary evolution, allelic recycling, and adaptive radiation of the
641 threespine stickleback. *Evolutionary Ecology Research* **15**, 377-411.
- 642 Bernatchez L, Renaut S, Whiteley AR, et al. (2010) On the origin of species: insights from the ecological
643 genomics of lake whitefish. *Philosophical Transactions of the Royal Society of London, Series B:*
644 *Biological Sciences* **365**, 1783-1800.
- 645 ~~Berner D, Grandchamp A-C, Hendry AP (2009) Variable Progress toward Ecological Speciation in
646 Parapatry: Stickleback across Eight Lake-Stream Transitions. *Evolution* **63**, 1740-1753.~~
- 647 Berner D, Moser D, Roesti M, Buescher H, Salzburger W (2014) Genetic architecture of skeletal evolution
648 in european lake and stream stickleback. *Evolution* **68**, 1792-1805.
- 649 ~~Berner D, Roesti M, Hendry AP, Salzburger W (2010) Constraints on speciation suggested by comparing
650 lake-stream stickleback divergence across two continents. *Molecular Ecology* **19**, 4963-4978.~~
- 651 Bierne N, Roze D, Welch JJ (2013) Pervasive selection or is it...? why are FST outliers sometimes so
652 frequent? *Molecular Ecology* **22**, 2061-2064.
- 653 Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans
654 may fail to map local adaptation genes. *Molecular Ecology* **20**, 2044-2072.
- 655 Boughman JW (2007) Speciation in sticklebacks. In: *Biology of the Three-Spined Stickleback* (eds.
656 Östlund-Nilsson S, Mayer I, Huntingford F), pp. 83-126. CRC Press., Boca Raton, FL.
- 657 Butlin RK, Saura M, Charrier G, et al. (2014) Parallel evolution of local adaptation and reproductive
658 isolation in the face of gene flow. *Evolution* **68**, 935-949.
- 659 Chaturvedi A, Raeymaekers JAM, Volckaert FAM (2014) Computational identification of miRNAs, their
660 targets and functions in three-spined stickleback (*Gasterosteus aculeatus*). *Molecular Ecology*
661 *Resources* **14**, 768-777.
- 662 Colosimo PF, Hosemann KE, Balabhadra S, et al. (2005) Widespread parallel evolution in sticklebacks by
663 repeated fixation of ectodysplasin alleles. *Science* **307**, 1928-1933.

- 664 Coscia I, Vogiatzi E, Kotoulas G, Tsigenopoulos C, Mariani S (2011) Exploring neutral and adaptive
665 processes in expanding populations of gilthead sea bream, *Sparus aurata* L., in the North-East
666 Atlantic. *Heredity* **108**, 537-546.
- 667 de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan methods against more
668 complex models: when and how much should we trust them? *Molecular Ecology* **23**, 2006-2019.
- 669 Deagle BE, Jones FC, Chan YF, *et al.* (2012) Population genomics of parallel phenotypic evolution in
670 stickleback across stream-lake ecological transitions. *Proceedings of the Royal Society of London,
671 Series B: Biological Sciences* **279**, 1277-1286.
- 672 DeFaveri J, Jonsson PR, Merilä J (2013) Heterogeneous genomic differentiation in marine threespine
673 sticklebacks: Adaptation along an environmental gradient. *Evolution* **67**, 2530-2546.
- 674 DeFaveri J, Merilä J (2013) Evidence for adaptive phenotypic differentiation in Baltic Sea sticklebacks.
675 *Journal of Evolutionary Biology* **26**, 1700-1715.
- 676 DeFaveri J, Shikano T, Shimada Y, Goto A, Merilä J (2011) Global analysis of genes involved in freshwater
677 adaptation in threespine stickleback (*Gasterosteus aculeatus*). *Evolution* **65**, 1800-1807.
- 678 Dvorky R, Ahmadian MR (2004) Always look on the bright site of Rho: structural implications for a
679 conserved intermolecular interface. *EMBO reports* **5**, 1130-1136.
- 680 ~~Egan SP, Nosil P, Funk DJ (2008) Selection and genomic differentiation during ecological speciation:
681 isolating the contributions of host association via a comparative genome scan of neochlamisus
682 bebbianae leaf beetles. *Evolution* **62**, 1162-1181.~~
- 683 Endler JA (1973) Gene flow and population differentiation: Studies of clines suggest that differentiation
684 along environmental gradients may be independent of gene flow. *Science* **179**, 243-250.
- 685 Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population.
686 *Heredity* **103**, 285-298.
- 687 Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population
688 genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567.
- 689 Feder JL, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends in Genetics* **28**, 342-
690 350.
- 691 Feulner PGD, Chain FJJ, Panchal M, *et al.* (2013) Genome-wide patterns of standing genetic variation in a
692 marine population of three-spined sticklebacks. *Molecular Ecology* **22**, 635-649.
- 693 Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both
694 dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977-993.
- 695 Foster SA, Scott RJ, Cresko WA (1998) Nested biological variation and speciation. *Philosophical
696 Transactions of the Royal Society of London, Series B: Biological Sciences* **353**, 207-218.
- 697 Garnaas MK, Moodie KL, Liu M-l, *et al.* (2008) Syx, a RhoA guanine exchange factor, is essential for
698 angiogenesis in vivo. *Circulation Research* **103**, 710-716.
- 699 Hagen DW (1967) Isolating mechanisms in threespine sticklebacks (*Gasterosteus*). *Journal of the Fisheries
700 Research Board of Canada* **24**, 1637-&.
- 701 Hansen MM, Meier K, Mensberg K-LD (2010) Identifying footprints of selection in stocked brown trout
702 populations: a spatio-temporal approach. *Molecular Ecology* **19**, 1787-1800.
- 703 Hansen MM, Ruzzante DE, Nielsen EE, Bekkevold D, Mensberg K-LD (2002) Long-term effective
704 population sizes, temporal stability of genetic composition and potential for local adaptation in
705 anadromous brown trout (*Salmo trutta*) populations. *Molecular Ecology* **11**, 2523-2535.
- 706 Hendry AP (2009) Ecological speciation! Or the lack thereof? *Canadian Journal of Fisheries and Aquatic
707 Sciences* **66**, 1383-1398.
- 708 Hendry AP, Bolnick DI, Berner D, Peichel CL (2009) Along the speciation continuum in sticklebacks.
709 *Journal of Fish Biology* **75**, 2000-2036.

- 710 Heuts MJ (1947) The phenotypical variability of *Gasterosteus aculeatus* L. populations in Belgium.
711 *Verhandelingen van de Koninklijke Vlaamse Academie voor Wetenschappen, Letteren en Schone*
712 *Kunsten van België* **9**, 1-63.
- 713 Hohenlohe PA, Bassham S, Currey M, Cresko WA (2012) Extensive linkage disequilibrium and parallel
714 adaptive divergence across threespine stickleback genomes. *Philosophical Transactions of the*
715 *Royal Society of London, Series B: Biological Sciences* **367**, 395-408.
- 716 Hohenlohe PA, Bassham S, Etter PD, et al. (2010) Population genomics of parallel adaptation in
717 threespine stickleback using sequenced RAD tags. *PLoS Genetics* **6**.
- 718 Jones F, Brown C, Pemberton J, Braithwaite V (2006) Reproductive isolation in a threespine stickleback
719 hybrid zone. *Journal of Evolutionary Biology* **19**, 1531-1544.
- 720 Jones FC, Chan YF, Schmutz J, et al. (2012a) A genome-wide SNP genotyping array reveals patterns of
721 global and repeated species-pair divergence in sticklebacks. *Current Biology* **22**, 83-90.
- 722 Jones FC, Grabherr MG, Chan YF, et al. (2012b) The genomic basis of adaptive evolution in threespine
723 sticklebacks. *Nature* **484**, 55-61.
- 724 Kauer MO, Dieringer D, Schlotterer C (2003) A microsatellite variability screen for positive selection
725 associated with the "Out of Africa" habitat expansion of *Drosophila melanogaster*. *Genetics* **165**,
726 1137-1148.
- 727 Levene H (1953) Genetic equilibrium when more than one ecological niche is available. *American*
728 *Naturalist* **87**, 331-333.
- 729 Lind EE, Grahn M (2011) Directional genetic selection by pulp mill effluent on multiple natural
730 populations of three-spined stickleback (*Gasterosteus aculeatus*). *Ecotoxicology* **20**, 503-512.
- 731 Lucek K, Sivasundar A, Kristjánsson BK, Skúlason S, Seehausen O (2014) Quick divergence but slow
732 convergence during ecotype formation in lake and stream stickleback pairs of variable age.
733 *Journal of Evolutionary Biology* **27**, 1878-1892.
- 734 Mäkinen HS, Cano JM, Merilä J (2008) Identifying footprints of directional and balancing selection in
735 marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. *Molecular*
736 *Ecology* **17**, 3565-3582.
- 737 McCairns RJS, Bernatchez L (2010) Adaptive divergence between freshwater and marine sticklebacks:
738 insights into the role of phenotypic plasticity from an integrated analysis of candidate gene
739 expression. *Evolution* **64**, 1029-1047.
- 740 McKinnon JS, Rundle HD (2002) Speciation in nature: the threespine stickleback model systems. *Trends in*
741 *Ecology & Evolution* **17**, 480-488.
- 742 McPhail J (1994) Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus*)
743 of south-western British Columbia. In: *The evolutionary biology of the threespine stickleback*
744 (eds. Bell AM, Foster S), pp. 399-437. Oxford University Press Inc., New York.
- 745 ~~Michel AP, Sim S, Powell THQ, et al. (2010) Widespread genomic divergence during sympatric speciation.~~
746 ~~*Proceedings of the Royal Society of London, Series B: Biological Sciences* **107**, 9724-9729.~~
- 747 ~~Milano I, Babbucci M, Cariani A, et al. (2014) Outlier SNP markers reveal fine-scale genetic structuring~~
748 ~~across European hake populations (*Merluccius merluccius*). *Molecular Ecology* **23**, 118-135.~~
- 749 Mitchell-Olds T, Schmitt J (2006) Genetic mechanisms and evolutionary significance of natural variation
750 in *Arabidopsis*. *Nature* **441**, 947-952.
- 751 Moore J-S, Gow JL, Taylor EB, Hendry AP (2007) Quantifying the constraining influence of gene flow on
752 adaptive divergence in the lake-stream threespine stickleback System. *Evolution* **61**, 2015-2026.
- 753 Narum SR, Hess JE (2011) Comparison of F-ST outlier tests for SNP loci under selection. *Molecular*
754 *Ecology Resources* **11**, 184-194.
- 755 Nielsen E, Hemmer-Hansen J, Poulsen N, et al. (2009a) Genomic signatures of local directional selection
756 in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology*
757 **9**, 276.

- 758 | ~~Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D (2009b) Population genomics of marine fishes:~~
 759 | ~~identifying adaptive variation in space and time. *Molecular Ecology* **18**, 3128-3150.~~
- 760 | Nosil P (2009) Adaptive population divergence in cryptic color-pattern following a reduction in gene flow.
 761 | *Evolution* **63**, 1902-1912.
- 762 | Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence.
 763 | *Molecular Ecology* **18**, 375-402.
- 764 | Olafsdottir GA, Snorrason SS (2009) Parallels, nonparallels, and plasticity in population differentiation of
 765 | threespine stickleback within a lake. *Biological Journal of the Linnean Society* **98**, 803-813.
- 766 | Orsini L, Spanier KI, De Meester LUC (2012) Genomic signature of natural and anthropogenic stress in
 767 | wild populations of the waterflea *Daphnia magna*: validation in space, time and experimental
 768 | evolution. *Molecular Ecology* **21**, 2160-2175.
- 769 | Papa R, Martin A, Reed RD (2008) Genomic hotspots of adaptation in butterfly wing pattern evolution.
 770 | *Current Opinion in Genetics & Development* **18**, 559-564.
- 771 | Peichel CL, Nereng KS, Ohgi KA, *et al.* (2001) The genetic architecture of divergence between threespine
 772 | stickleback species. *Nature* **414**, 901-905.
- 773 | Perrier C, Bourret V, Kent MP, Bernatchez L (2013) Parallel and nonparallel genome-wide divergence
 774 | among replicate population pairs of freshwater and anadromous Atlantic salmon. *Molecular*
 775 | *Ecology* **22**, 5577-5593.
- 776 | Pinho C, Hey J (2010) Divergence with gene flow: models and data. *Annual Review of Ecology, Evolution*
 777 | *and Systematics* **41**, 215-230.
- 778 | Poncet BN, Herrmann D, Gugerli F, *et al.* (2010) Tracking genes of ecological relevance using a genome
 779 | scan in two independent regional population samples of *Arabidopsis thaliana*. *Molecular Ecology* **19**,
 780 | 2896-2907.
- 781 | Radaelli G, Rowlerson A, Mascarello F, Patruno M, Funkenstein B (2003) Myostatin precursor is present
 782 | in several tissues in teleost fish: a comparative immunolocalization study. *Cell and Tissue*
 783 | *Research* **311**, 239-250.
- 784 | Raeker MÖ, Bieniek AN, Ryan AS, *et al.* (2010) Targeted deletion of the zebrafish obscurin A RhoGEF
 785 | domain affects heart, skeletal muscle and brain development. *Developmental Biology* **337**, 432-
 786 | 443.
- 787 | Raeymaekers JAM, Konijnendijk N, Larmuseau MHD, *et al.* (2014) A gene with major phenotypic effects
 788 | as a target for selection vs. homogenizing gene flow. *Molecular Ecology* **23**, 162-181.
- 789 | Raeymaekers JAM, Lens LUC, Van den Broeck F, Van Dongen S, Volckaert FAM (2012) Quantifying
 790 | population structure on short timescales. *Molecular Ecology* **21**, 3458-3473.
- 791 | Raeymaekers JAM, Maes GE, Audenaert E, Volckaert FAM (2005) Detecting Holocene divergence in the
 792 | anadromous-freshwater three-spined stickleback (*Gasterosteus aculeatus*) system. *Molecular*
 793 | *Ecology* **14**, 1001-1014.
- 794 | Raeymaekers JAM, Van Houdt JKJ, Larmuseau MHD, Geldof S, Volckaert FAM (2007) Divergent selection
 795 | as revealed by P_{ST} and QTL-based F_{ST} in three-spined stickleback (*Gasterosteus aculeatus*)
 796 | populations along a coastal-inland gradient. *Molecular Ecology* **16**, 891-905.
- 797 | Roesti M, Hendry AP, Salzburger W, Berner D (2012) Genome divergence during evolutionary
 798 | diversification as revealed in replicate lake-stream stickleback population pairs. *Molecular*
 799 | *Ecology* **21**, 2852-2862.
- 800 | ~~Rogers S, Bernatchez L (2007) The genetic architecture of ecological speciation and the association with~~
 801 | ~~signatures of selection in natural lake whitefish (*Coregonus* sp. Salmonidae) species pairs.~~
 802 | ~~*Molecular Biology and Evolution* **24**, 1423-1438.~~
- 803 | Rohlf FJ (2002) tpsDig, digitize landmarks and outlines. Department of Ecology and Evolution, State
 804 | University of New York at Stony Brook., New York.

- 805 Rossman KL, Der CJ, Sondek J (2005) GEF means go: turning on RHO GTPases with guanine nucleotide-
806 exchange factors. *Nature Reviews: Molecular Cell Biology* **6**, 167-180.
- 807 Schluter D (1996) Adaptive radiation along genetic lines of least resistance. *Evolution*, 1766-1774.
- 808 Schluter D (2009) Evidence for Ecological Speciation and Its Alternative. *Science* **323**, 737-741.
- 809 Schluter D, Conte GL (2009) Genetics and ecological speciation. *Proceedings of the National Academy of*
810 *Sciences, USA* **106**, 9955-9962.
- 811 Seehausen O, Butlin RK, Keller I, *et al.* (2014) Genomics and the origin of species. *Nature Reviews:*
812 *Genetics* **15**, 176-192.
- 813 Shikano T, Ramadevi J, Merilä J (2010a) Identification of local-and habitat-dependent selection: scanning
814 functionally important genes in nine-spined sticklebacks (*Pungitius pungitius*). *Molecular Biology*
815 *and Evolution* **27**, 2775-2789.
- 816 Shikano T, Ramadevi J, Shimada Y, Merilä J (2010b) Utility of sequenced genomes for microsatellite
817 marker development in non-model organisms: a case study of functionally important genes in
818 nine-spined sticklebacks (*Pungitius pungitius*). *BMC Genomics* **11**, 334.
- 819 Shimada Y, Shikano T, Merilä J (2011) A high incidence of selection on physiologically important genes in
820 the three-spined stickleback, *Gasterosteus aculeatus*. *Molecular Biology and Evolution* **28**, 181-
821 193.
- 822 [Slatkin \(1987\). Gene flow and the geographic structure of natural populations. *Science* **236**,787–792.](#)
- 823 Soria-Carrasco V, Gompert Z, Comeault AA, *et al.* (2014) Stick insect genomes reveal natural selection's
824 role in parallel speciation. *Science* **344**, 738-742.
- 825 Stapley J, Reger J, Feulner PGD, *et al.* (2010) Adaptation genomics: the next generation. *Trends in Ecology*
826 *& Evolution* **25**, 705-712.
- 827 Storz JF (2005) Invited review: Using genome scans of DNA polymorphism to infer adaptive population
828 divergence. *Molecular Ecology* **14**, 671-688.
- 829 Strasburg JL, Sherman NA, Wright KM, *et al.* (2012) What can patterns of differentiation across plant
830 genomes tell us about adaptation and speciation? *Philosophical Transactions of the Royal Society*
831 *of London, Series B: Biological Sciences* **367**, 364-373.
- 832 Taylor WR, Dyke GCV (1985) Revised procedures for staining and clearing small fishes and other
833 vertebrates for bone and cartilage study. *Cybiurn* **9**, 107-119.
- 834 Tregenza T (2002) Divergence and reproductive isolation in the early stages of speciation. In: *Genetics of*
835 *Mate Choice: From Sexual Selection to Sexual Isolation* (eds. Etges WJ, Noor MAF), pp. 291-300.
836 Springer Netherlands.
- 837 Van Dongen S, Lens L, Pape E, Volckaert FAM, Raeymaekers JAM (2009) Evolutionary history shapes the
838 association between developmental instability and population-level genetic variation in three-
839 spined sticklebacks. *Journal of Evolutionary Biology* **22**, 1695-1707.
- 840 van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: software for identifying
841 and correcting genotyping errors in microsatellite data. *Molecular Ecology notes* **4**, 535-538.
- 842 Vandamme S, Maes G, Raeymaekers J, *et al.* (2014) Regional environmental pressure influences
843 population differentiation in turbot (*Scophthalmus maximus*). *Molecular Ecology* **23**, 618-636.
- 844 Via S (2009) Natural selection in action during speciation. *Proceedings of the National Academy of*
845 *Sciences, USA* **106**, 9939-9946.
- 846 [Via S, Conte G, Mason-Foley C, Mills K \(2012\) Localizing FST outliers on a QTL map reveals evidence for](#)
847 [large genomic regions of reduced gene exchange during speciation with gene flow. *Molecular*](#)
848 [Ecology](#) **21**, 5546-5560.
- 849 Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human
850 genome. *Plos Biology* **4**, e72.
- 851 Walsh F, Celeste A (2005) Myostatin: a modulator of skeletal-muscle stem cells. *Biochemical Society*
852 *Transactions* **33**, 1513-1517.

- 853 Wootton RJ (1976) *The Biology of the Sticklebacks*. Academic Press, London.
- 854 Yang L, Jelsbak L, Marvig RL, *et al.* (2011) Evolutionary dynamics of bacteria in a human host
855 environment. *Proceedings of the National Academy of Sciences, USA* **108**, 7481-7486.
- 856 Yeaman S, Otto SP (2011) Establishment and maintenance of adaptive genetic divergence under
857 migration, selection, and drift. *Evolution* **65**, 2123-2129.
- 858 Yeaman S, Whitlock MC (2011) The genetic architecture of adaptation under migration–selection
859 balance. *Evolution* **65**, 1897-1911.
- 860 Zulliger D, Schnyder E, Gugerli F (2013) Are adaptive loci transferable across genomes of related species?
861 Outlier and environmental association analyses in Alpine Brassicaceae species. *Molecular*
862 *Ecology* **22**, 1626-1639.

For Review Only

863 **FIGURE LEGENDS**

864

865 **Figure 1.** Map with the four sampling locations of three-spined stickleback in northwestern Belgium and
866 southwestern Netherlands. River network and major towns are mapped.

867

868 **Figure 2.** Average values for A) standard length, B) plate number, C) relative pelvic spine length, D)
869 relative pelvic plate length, and E) relative length of the second large gill raker in two brackish water
870 populations (L01, L02) and two freshwater populations (L12 and U01)) of three-spined stickleback.
871 Horizontal bars represent 95% confidence intervals.

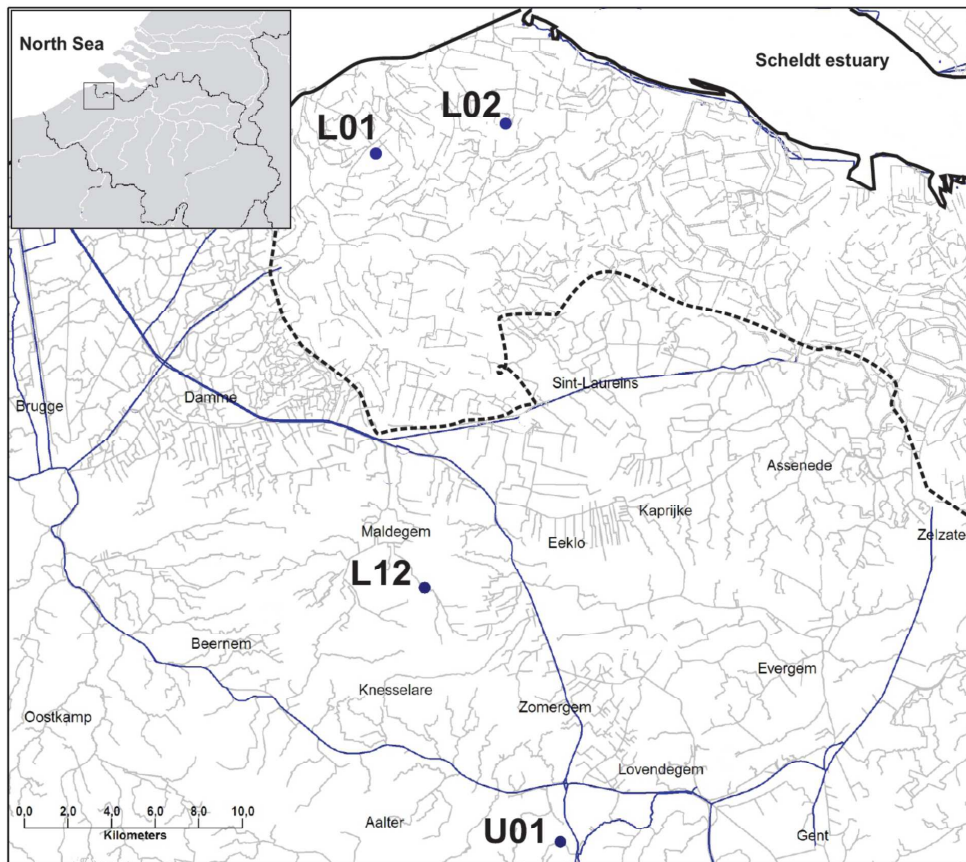
872

873 FIGURES

874

875 Figure 1

876

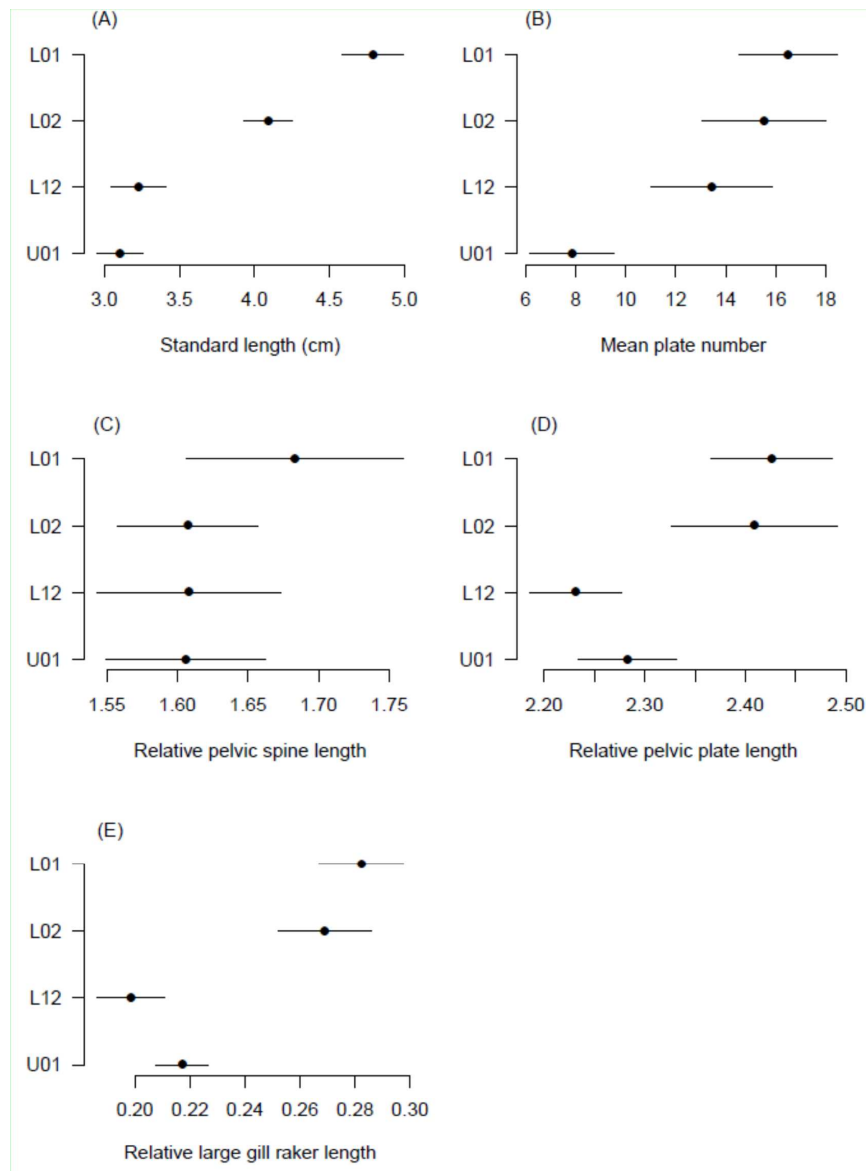


877

878

879

880 **Figure 2**
881



882
883

884

885

