

Adaptive Genetic Divergence along Narrow Environmental Gradients in Four Stream Insects

Kozo Watanabe^{1,2*}, So Kazama³, Tatsuo Omura⁴, Michael T. Monaghan¹

1 Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Müggelseedamm 301, Berlin, Germany, **2** Present address: Department of Civil and Environmental Engineering, Ehime University, Bunkyo-cho 3, Matsuyama, Japan, **3** Department of Civil and Environmental Engineering, Tohoku University, Aoba-yama 6-6-06, Sendai, Japan, **4** New Industry Creation Hatchery Center (NICHe), Tohoku University, Aoba-yama 6-6-04, Sendai, Japan

Abstract

A central question linking ecology with evolutionary biology is how environmental heterogeneity can drive adaptive genetic divergence among populations. We examined adaptive divergence of four stream insects from six adjacent catchments in Japan by combining field measures of habitat and resource components with genome scans of non-neutral Amplified Fragment Length Polymorphism (AFLP) loci. Neutral genetic variation was used to measure gene flow and non-neutral genetic variation was used to test for adaptive divergence. We identified the environmental characteristics contributing to divergence by comparing genetic distances at non-neutral loci between sites with Euclidean distances for each of 15 environmental variables. Comparisons were made using partial Mantel tests to control for geographic distance. In all four species, we found strong evidence for non-neutral divergence along environmental gradients at between 6 and 21 loci per species. The relative contribution of these environmental variables to each species' ecological niche was quantified as the specialization index, S , based on ecological data. In each species, the variable most significantly correlated with genetic distance at non-neutral loci was the same variable along which each species was most narrowly distributed (i.e., highest S). These were gradients of elevation (two species), chlorophyll- a , and ammonia-nitrogen. This adaptive divergence occurred in the face of ongoing gene flow ($F_{st}=0.01-0.04$), indicating that selection was strong enough to overcome homogenization at the landscape scale. Our results suggest that adaptive divergence is pronounced, occurs along different environmental gradients for different species, and may consistently occur along the narrowest components of species' niche.

Citation: Watanabe K, Kazama S, Omura T, Monaghan MT (2014) Adaptive Genetic Divergence along Narrow Environmental Gradients in Four Stream Insects. PLoS ONE 9(3): e93055. doi:10.1371/journal.pone.0093055

Editor: Norman Johnson, University of Massachusetts, United States of America

Received: July 3, 2013; **Accepted:** March 2, 2014; **Published:** March 28, 2014

Copyright: © 2014 Watanabe et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Japan Society for the Promotion of Science (JSPS) (grant numbers 22360192, 25289172, 24254003 and 25303020), the Ministry of Land, Infrastructure and Transport Japan (grant number 07052442), and the European Commission (grant number 237026). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: watanabe_kozo@cee.ehime-u.ac.jp

Introduction

Adaptive genetic divergence among populations within species is an evolutionary process on the way towards ecological speciation [1] and is thought to be an important driver of the formation and maintenance of biodiversity [2]. As the potential causes of adaptive divergence, evolutionary biologists pointed to complex mechanisms that involve multiple drivers such as divergent natural selection against migrants ([3]), secondary contact following allopatric isolation (e.g., [4]), interspecific competition [5], and genetic linkage between quantitative trait loci (QTLs) (e.g., [6]).

There is experimental and descriptive evidence for adaptive divergence across a range of taxa and environments (e.g., [7–9]). Studies that investigate the genetic basis of adaptation often use candidate gene and QTL approaches [10–13]. Unfortunately these molecular genetic tools are largely limited to model organisms and well characterized genes, making it difficult to apply the approach to natural populations or to look for general patterns across multiple species. Genome scans allow for the study of non-model organisms and a number of studies have reported evidence for genetic adaptation, where changes in allele frequencies at putatively adaptive loci are correlated with environmental

changes across the species' range (e.g., [14–17]). Such studies provide strong evidence for adaptive divergence among populations at one or more parts of the genome but are often limited to single species and a few environmental parameters. The result is that it is difficult to make any generalizations across taxa or environmental gradients in natural populations. Natural populations are subject to many selective forces simultaneously, and several environmental factors may be responsible for the observed genetic divergence.

An alternative approach is to study several co-occurring species in nature and multiple components of their environment and to look for general patterns of adaptive divergence (i.e. along similar environmental gradients). By screening large numbers of candidate loci and individuals [18–22], statistical methods can be used to identify loci that are either under direct selection or are linked to loci under selection (“outlier” loci; reviewed by [15]). Compared to loci that are neutral, loci influenced by directional selection should show more genetic differentiation, and loci subject to balancing selection should show less genetic differentiation. In practice, these statistical methods identify loci with F_{ST} values that are significantly different from those expected under neutral conditions.

Here we examined adaptive divergence in four species of aquatic insect from six adjacent stream catchments in northeastern Honshu Island, Japan. The species differ in their local distribution and thus are likely to have different ecological niches. Because most aquatic insect dispersal occurs in the adult stage, the distribution of aquatic immatures is primarily determined by local environmental characteristics [23]. There were three objectives in this study. The first was to measure the extent of adaptive divergence in natural populations. We genotyped individuals using Amplified Fragment Length Polymorphism (AFLP) loci and identified loci under selection (“outlier loci”) using locus-specific tests of genetic differentiation among sampling sites. The second objective was to examine the link between environmental heterogeneity and any observed non-neutral genetic divergence. We first measured 15 habitat variables at each location, and identified those characteristics most limiting species distributions in the study area using the ecological specialization index (S) [24]. We also identified environmental variables that are most likely to contribute to adaptive divergence for each species using Mantel tests, and compared the variables with those identified by the specialization index analyses. The third objective was to test whether the patterns across the four species were consistent.

Methods

Ethics statement

We did not require ethical approval to conduct this study as we did not handle or collect samples of any vertebrates. We received permissions to access study sites and to collect samples from the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) of Japan (#07052442).

Study area and species

The study was carried out in six adjacent stream catchments in Miyagi Prefecture, Japan (*ca* 120 sq km) (Fig. 1). Streams in the area are characterized by high environmental heterogeneity along short and steep corridors. The region is largely mountainous or hilly and forested with beech (*Fagus*) and cedar (*Cryptomeria*). Lowland areas are used for agriculture (13% of total study area, primarily rice paddy) or are a mix of residential and commercial areas (11%). Study sites ($n = 62$) were chosen to represent the wide range of physical and chemical habitat characteristics in the region.

The four species examined are typical of high-gradient, fast-flowing streams in Japan. Three were caddisflies (Trichoptera), *Hydropsyche orientalis* (Ho), *Stenopsyche marmorata* (Sm), and *Hydropsyche albicephala* (Ha); and one was a mayfly (Ephemeroptera), *Ephemera japonica* (Ej). Individuals of each species live in the water for up to one year before emergence, reproduction, and aerial dispersal. As aquatic larvae, they feed primarily on fine (<1 mm diameter) particulate organic matter that is either transported in the water column (Ho, Sm and Ha) or deposited on the stream bed (Ej). Sampling was conducted in July and November (summer and autumn) 2006. At each site, two person-hours were spent collecting larval individuals of as many of the target species as possible using a kick net (250 μ m mesh) along a 300-m stream reach. If more than one individual was found in either July or November the species was considered present at the site. Samples were preserved in 99% ethanol in the field, returned to the laboratory and identified to species under a stereomicroscope (150 \times).

Environmental variables

We measured 31 biological, physical, and chemical characteristics of the habitat known to influence the distribution of larval insects in streams [23]. These included both resource (e.g., organic matter transported in the water column) and habitat (e.g., stream width) variables. Data were collected from each study site using standard methods in stream ecological surveys (e.g., [25,26]). Using average values of July and November measurements, 15 of these variables were chosen for further analysis by eliminating variables that were correlated (Table S1 in File S1).

Of the 15 variables, altitude was measured using an altimeter at the center of each site. Stream order was determined using a 1:25000 map. Width of the stream channel was measured at 10 randomly selected cross-sections using a tape measure. Water velocity was measured at 20 randomly selected points in riffles at $0.2 \times$ water depth from the surface using an electromagnetic current meter. Substrate (gravel) size was calculated as the mean of the longest diameter of 36 pieces of gravel at grid points located in a haphazardly selected 1-sq m area of the stream bottom in a riffle. Benthic (on the stream bottom) coarse particulate organic matter (BCPOM; particles >1.0 mm), mostly composed of leaves and branches, was collected using a Hess sampler (0.09 m²). Attached algae was measured by scrubbing a 9-sq cm surface of randomly selected gravels and collecting the supernatant on glass fiber filters (pore size 1 μ m, Whatman GF/B). Water samples were taken from the surface of river, poured through a sieve (1 mm) and then filtered onto glass fiber filters (1 μ m). These filtered samples were used to quantify suspended fine particulate organic matter (SFPOM: 1 μ m–1.0 mm), suspended solids (SS: 1 μ m–1.0 mm) and chlorophyll-*a* (Chl-*a*) concentrations. BOM and POM were quantified by measuring ash free dry mass (AFDM; the decrease in mass of a sample after drying samples at 60°C for 24 h and combusting samples at 300°C for 30 min). Using the water samples, Biochemical Oxygen Demand (BOD), the amount of dissolved oxygen needed to biologically break down organic materials in the water sample, was measured as an index of organic pollution. Chl-*a* on the filters from the epilithon and water samples was quantified using a fluorescence spectrophotometer (HITACHI F-4500) following 99% methanol extraction (10°C, 24 h). All of the above biological variables were calculated as the mean of three replicate samples per site. Nutrient concentration in the water sample (dissolved inorganic nitrogen: NO_x-N and NH₄-N; soluble reactive phosphorus: PO₄-P) were measured using a TRAACS 800 auto-analyzer (BLTEC, Japan) ($n = 1$).

Species' niche and the specialization index (S)

The relative importance of each environmental variable in determining local habitat distribution was quantified using the specialization index (S) of Hirzel et al. [24]. S was calculated as the ratio of the standard deviation of an environmental variable across all sites (σ_g) to the standard deviation of the variable at sites where the species is present (σ_s) ($S = \sigma_g / \sigma_s$). When a species has an optimum for a given environmental variable, the variable is expected to show a narrow and nonrandom distribution at the sites where the species is present compared to the whole set of study sites. A randomly chosen set of sites is expected to have $S = 1$, with $S > 1$ indicating some form of specialization of a species to that particular variable [24]. A high value of S indicates that a species occurs along only a narrow range of the total gradient available among all habitats.

DNA extraction and AFLP fingerprinting

We genotyped 1793 individuals from 4 species (Table 1; 18.3 individuals/sites/species in average) using AFLP markers [27].

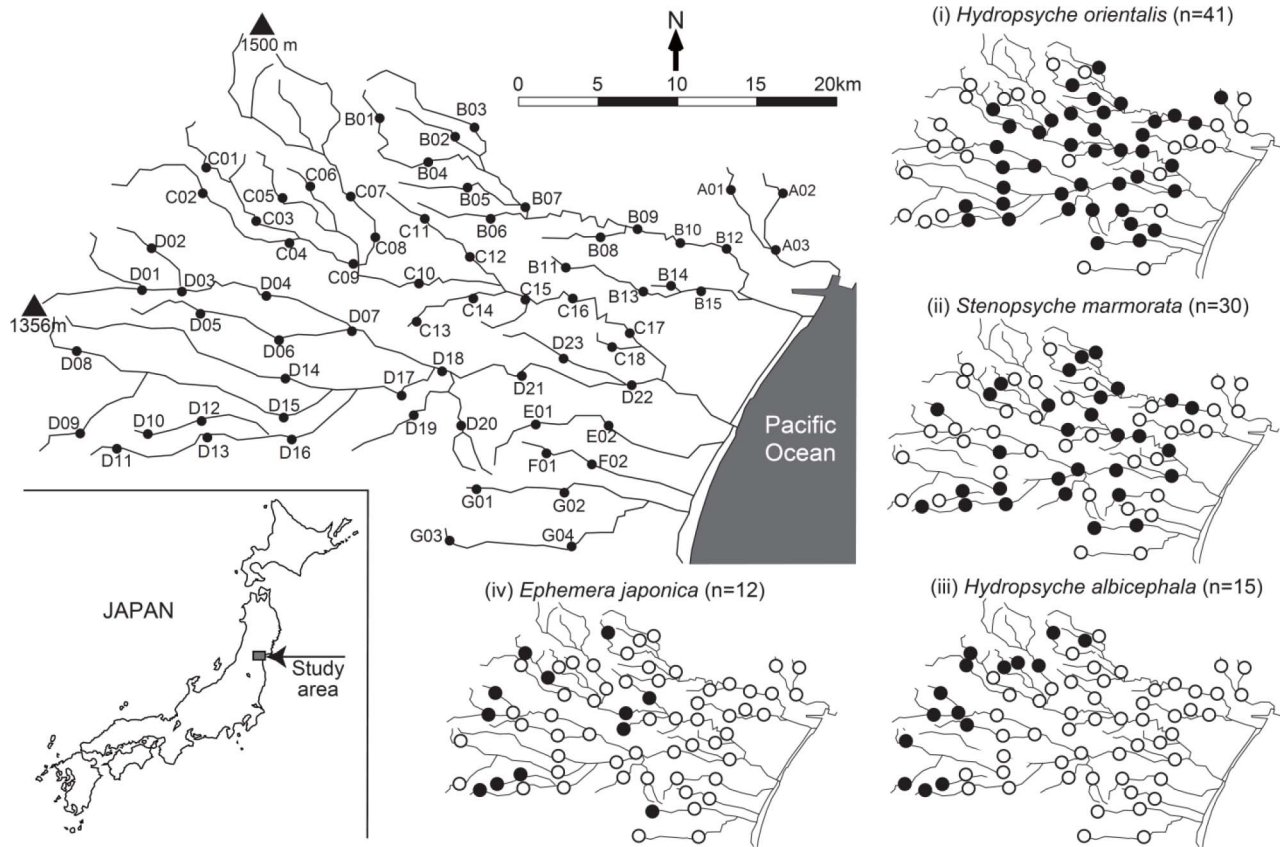


Figure 1. Map of the 62 study sites in Miyagi Prefecture, northeast Honshu, Japan. The four small panels (i–iv) show the presence (filled) and absence (open) of each species based on ecological surveys in summer (July) and autumn (November) 2006. The codes reflect the 6 different catchments.

doi:10.1371/journal.pone.0093055.g001

The digestive tract and epidermis were removed and DNA was isolated from remaining tissue using DNeasy 96 Blood and Tissue Kits (Qiagen, Tokyo). Genotyping was performed using the AFLP Plant Mapping Kit (Applied Biosystems, Tokyo) and the manufacturer's protocol with the following modifications: restriction was performed at 37°C for 3 h using 300 ng DNA with 1 U MseI, 5 U EcoRI, 5 µg BSA, 1 µL MseI buffer, 1 µL EcoRI buffer, and 13 µL sterilized dH₂O. Ligation was then performed by adding 1 U T4 DNA ligase, 1 µL MseI adapter, 1 µL EcoRI adapter, 4 µL T4 DNA ligase buffer (10 x), and 14 µL sterilized dH₂O and incubating at 16°C for 10 h. Restriction/ligation products were then diluted 1:5 with TE buffer (10 mM Tris-HCl; 0.5 M EDTA) prior to pre-selective amplification using standard MseI and EcoRI oligonucleotide primers.

For the selective amplification reaction, 64 primer pairs contained in the AFLP Plant Mapping Kit were initially used to survey polymorphisms using three individuals of each species. The three primer pairs that generated the highest variability (Table S2 in File S1) were then used for all individuals. The three primer combinations yielded between 128 and 473 polymorphic loci in the four species (Table 1). Fluorescently labeled AFLP products (1 µl) were mixed with 0.1 µL ROX Size Standard (50–500 bp; Applied Biosystems) and 10 µL formamide and size-separated on an Applied Biosystems 3130 xl automated sequencer. GeneMapper v 2.1 (Applied Biosystems) was used to prepare a data matrix of the calculated sizes for all fragments from each individual. Peak intensity >200 RFU was used to determine the presence or absence of a fragment in the generated electropherogram. Scored

fragments used for analysis ranged from 50–450 bp in length. Ten haphazardly chosen individuals of each species were analyzed twice (i.e., restriction, ligation, pre-amplification, selective amplification, size-separation, scoring) in order to test the repeatability of our AFLP methods. In these 40 individuals, 41 differences were observed out of 14670 scored fragments, indicating an error rate of 0.28%.

Detection of outlier loci

Two different genome scan methods of Dfdist and BayeScan were applied to identify outlier loci. Dfdist is an extension of Fdist [28] allowing the use of dominant markers. This method uses coalescent simulations to generate thousands of loci evolving under a neutral model of symmetrical islands with a mean global F_{ST} among sampling sites close to the observed mean global F_{ST} . Assuming that loci influenced by directional or balancing selection exhibit a larger or smaller genetic differentiation than neutral loci, the methods identify outlier loci that present observed mean global F_{ST} are significantly different from those expected under the neutral variation. Dfdist calculates observed global F_{ST} for dominant DNA markers at each locus using null-allele frequencies estimated by the Bayesian approach of Zhivotovskiy [29]. Mean global F_{ST} was calculated using the default method of first excluding 30% of the highest and lowest observed values. Empirical loci with F_{ST} values significantly greater ($p < 0.05$) than the simulated distribution (generated with 50,000 loci) were considered to be outliers.

Table 1. Number of locations where species were present from among the 62 study sites, number of individuals used for AFLP analysis (n), number of outlier loci detected with the two methods applied (Dfdist with 95% significance level; BayeScan at posterior probability >0.95) and number of loci detected by both methods, and the number of neutral loci.

Species	Sites	n	Total loci	Outlier loci			Neutral loci
				BayeScan	Dfdist	both	
<i>Hydropsyche orientalis</i> (Ho)	41	753	129	31	9	9 (7%)	98
<i>Stenopsyche marmorata</i> (Sm)	30	571	220	56	23	21 (10%)	164
<i>Hydropsyche albicephala</i> (Ha)	15	251	128	16	7	6 (5%)	111
<i>Ephemera japonica</i> (Ej)	12	218	473	23	7	7 (1%)	449

doi:10.1371/journal.pone.0093055.t001

An alternative approach, BayeScan, is a hierarchical Bayesian model-based method first described in [19,22] for dominant markers. The Bayesian method is based on the idea that F_{ST} values reflect contributions from locus-specific effects such as selection, and population-specific effects such as local effective size and immigration rates. The main advantage of this approach is to allow different demographic scenarios and different amounts of genetic drift in each population. Using a reversible jump Markov chain Monte Carlo approach, the posterior probability that each locus is subject to selection is estimated. A locus is considered to be influenced by selection if its F_{ST} is significantly higher or lower than the expectation provided by the coalescent simulations. For all subsequent analyses, outlier loci were defined as loci detected by both methods at the 95% confidence level. Neutral loci were defined as loci detected by neither Dfdist nor BayeScan at the 95% thresholds. Loci detected as outliers by only one of the two methods were not considered in further analysis of either outlier or neutral loci. We further tested for pairwise linkage disequilibrium (LD) of the outlier loci detected by both methods, using 1000 steps in the Markov chain and a dememorization of 1000 steps in ARLEQUIN 3.5 [30].

In outlier detection, the possibility of detecting false positives (i.e. committing type-I errors), among the large number of screened loci is a general concern. For example, the number of loci expected to be outliers by random chance = $0.05 \times$ number of loci (when a 95% criterion is employed). Correlation of allele frequencies across populations also potentially leads to a high rate of false positives because it invalidates the estimation of the variance of F_{ST} [31,32]. Although the stochastic simulation approach [28] implemented in Dfsist and BayeScan improved the performance of the test, the correlations might be the case in our sampling design where high gene flow is expected among the geographically close sites. Especially, Dfdist is likely to generate false positives when gene flow is asymmetric across populations, and/or when some populations experience bottlenecks [33–35]. Therefore, a false discovery rate (FDR) of 10% was adopted for Dfdist analysis by using the program MCHEZA ([36]), which implements the method of [37]. BayeScan attempts to address the issue of detecting false positives by calculating a locus-specific effect in the overall pattern of genetic differentiation [19]. Therefore, the outliers conservatively defined in our study as loci detected by both methods including the Bayesian method are less likely to represent false positives.

Analysis of genetic diversity and divergent selection

Three mean values of genetic differentiation (global F_{ST}) among sampling sites were calculated for each species: for all loci, for only neutral loci, and for only non-neutral loci. Global heterozygosity (H_T) and mean heterozygosity within sampling sites (H_W) was estimated separately for neutral and non-neutral loci using a Bayesian approach with a uniform prior distribution of allele frequencies [29] implemented in AFLP-SURV v 1.0 [38].

We used simple and partial Mantel tests [39] to test for correlation between pairwise genetic distances at non-neutral loci and environmental distance between sites. Some authors report that partial Mantel tests can result in incorrect p-values [40,41]. Although several alternative methods based on the likelihood analysis [42] and mixed modelling approach [43,44] have been proposed, the partial test remains a standard approach. Genetic distance between each pair of sites was quantified using the mean pairwise F_{ST} for all non-neutral loci. Each pairwise F_{ST} was calculated using the Bayesian-estimated allele frequencies generated by AFLP-SURV. The environmental difference between each pair of sites was quantified as a one-dimensional Euclidian

distance for each variable (Fig. 2). Simple Mantel tests were used to examine patterns of Isolation By Environmental Distance (IBED; [45]). Because genetic isolation by geographical distance (IBD) was significant in three of the four species examined (Table S3), partial Mantel tests were carried out to control for geographic distance. All tests were implemented using 3000 randomizations and were conducted using a program written in C (available from the authors) that reports one-tailed probabilities.

Results

Genetic structure and outlier loci

Global F_{ST} among sampling sites of four study species (Fig. 1) ranged from 0.04–0.09 (Table 2) and standard errors overlapped 0.00 in each case (data not shown). Measured using only neutral loci (see below), F_{ST} was even lower ($F_{ST} = 0.01$ – 0.04 ; Table 2) throughout all species (Permutation test, $p < 0.0125$). These results at neutral loci indicate high levels of gene flow in all of the species across the spatial scale of our study. Total genetic variance (H_T) was similar in all species. Values were lower at neutral loci than at non-neutral loci (Table 2) and the same was true for mean genetic variance within sampling sites (H_W).

Between 6 and 21 outlier loci were detected in the four species examined (Table 1). The proportion of outlier loci (using our criterion of significance with both methods) was not correlated to number of individuals used for AFLP analysis (n) or global F_{ST} (data not shown). Loci detected using BayeScan (Table 1) were under directional selection ($\alpha_i > 0$). Between 11 (Ha) and 35 (Sm) loci were detected as outliers by only one of the two methods. These were removed from subsequent analysis because they could not be confidently considered outliers nor neutral. LD analysis found that 8–20% of possible pair-wise combinations of outlier loci (i.e. one of the outlier pairs in one of the sites) were statistically linked (Randomization test, $p < 0.01$). The proportions of significant pairs were higher than expected by chance; however, there was no locus pair that was consistently in disequilibrium in multiple populations.

Habitat specialization and adaptive divergence

Based on the highest significant values of the specialization index (S), species were most constrained by altitude (for Ho and Sm), algal concentration in the water (measured as chlorophyll- a) (for Ha), and ammonium-nitrogen (NH_4) concentration in the water (for Ej) (Fig. 2) (Table S3). Other variables with large values of S included ammonium-nitrogen (Ho), water velocity (Ho), algal concentration (Ha), and substrate size (Ej).

At non-neutral loci, pairwise F_{st} among study sites were significantly correlated to between 1 and 3 environmental variables, depending on the species examined. This was based on partial Mantel tests of non-neutral F_{ST} and environmental (Euclidean) distance (Fig. 2, Table S3 in File S1). In all four species, the strongest correlation of pairwise genetic and environmental distances (i.e., highest partial correlation coefficient (r_p)) was with the variable most limiting species distributions (i.e., highest S). Specifically, pairwise F_{st} at non-neutral loci were most closely related to changes in altitude (Ho, Sm), transported chl- a (Ha), and ammonium-nitrogen (Ej) (Fig. 2). With 15 environmental variables, the probability that the variable with highest S also has the highest r_p for the 4 species was significantly higher than would be expected by chance ($p < 0.01$, Fisher's combined probability test). Partial mantel correlations for the other significant variables were also significant in each case, but lower (Fig. 2).

Discussion

Adaptive genetic divergence along narrow environmental gradients

There is a broad range of evidence for adaptive divergence of populations [7–9], and techniques are now available for the study of the genetic signatures of adaptive divergence in non-model organisms (e.g., [14,15]). An important next step in our understanding of adaptive divergence is to identify general patterns by examining multiple species and environmental characteristics simultaneously in nature. Our anonymous marker approach applied to four co-occurring species and a range of environmental characteristics is complementary to studies that test whether one or a few *a priori* loci or phenotypic traits are under selection in single species (e.g., [14,16,17]). We used genome scans (AFLP) to detect outlier loci presumably experiencing adaptive divergence. We then used Mantel test to identify several environmental factors contributing to the divergence at these loci. We used a “reverse ecology” approach (*sensu* [46]), whereby a suite of habitat and resource components of the niche were measured and correlation was applied to identify those driving adaptive divergence. We uncovered strong evidence for adaptive divergence in these riverine insects, and by simultaneously studying four species, we identified a striking trend: the environmental variables that were most significantly associated with genetic distance at non-neutral loci were the same variables that limit the species geographical distribution, as measured using the specialization index (S) of Hirzel et al. [47].

Adaptive divergence thus appears to be driven by the narrowest ecological gradient in each species, i.e. the most restricted spectrum of possible environments (*sensu* [48]). That the steepest ecological gradient is also the narrowest is not necessarily obvious, and many studies of adaptive divergence focus on the extremes of a particular environmental factor that occur over broad scales such as temperature extremes across broad latitudinal [16] and elevational [17] gradients. Nonetheless, if a niche is narrow, then genetic change (i.e., adaptive divergence) may be required to accommodate even a small environmental distance regardless of how steep a particular gradient may appear in our observations. Maladaptive gene flow from central to peripheral populations of a species' range (“migration load”; [49–51]) may account for the observed patterns. The immigrant genes may be adapted to the range center and thus inhibit adaptation at the periphery [52], making a species' range a function of gene flow and the steepness of a given selection gradient. What is interesting in our findings is that the species studied occur throughout Japan, thus while a strict range-margin explanation seems unlikely, it may be that similar processes operate between locally adapted populations on relatively small spatial scales. Another explanation is an intrinsic limit of species' evolutionary capacity [53]. The low frequency of mutations may not allow peripheral populations to adapt to greater environmental extremes. Finally, an intriguing alternative explanation is interspecific competition for space or resources. In theory [54], populations tend to diverge along the narrowest niche dimension because the dimension still has a space to drive the populations to use alternative resources, which provide a greater chance to escape from the high competition however this process is determined by the balance between the inherent costs and benefits of widening the niche [55]. We do not have any measures of competition to test this hypothesis, although local population densities can be very high in these species.

Before drawing firm conclusions about any general role of the particular environmental gradients identified here, it is important to consider both the characteristics themselves and the temporal

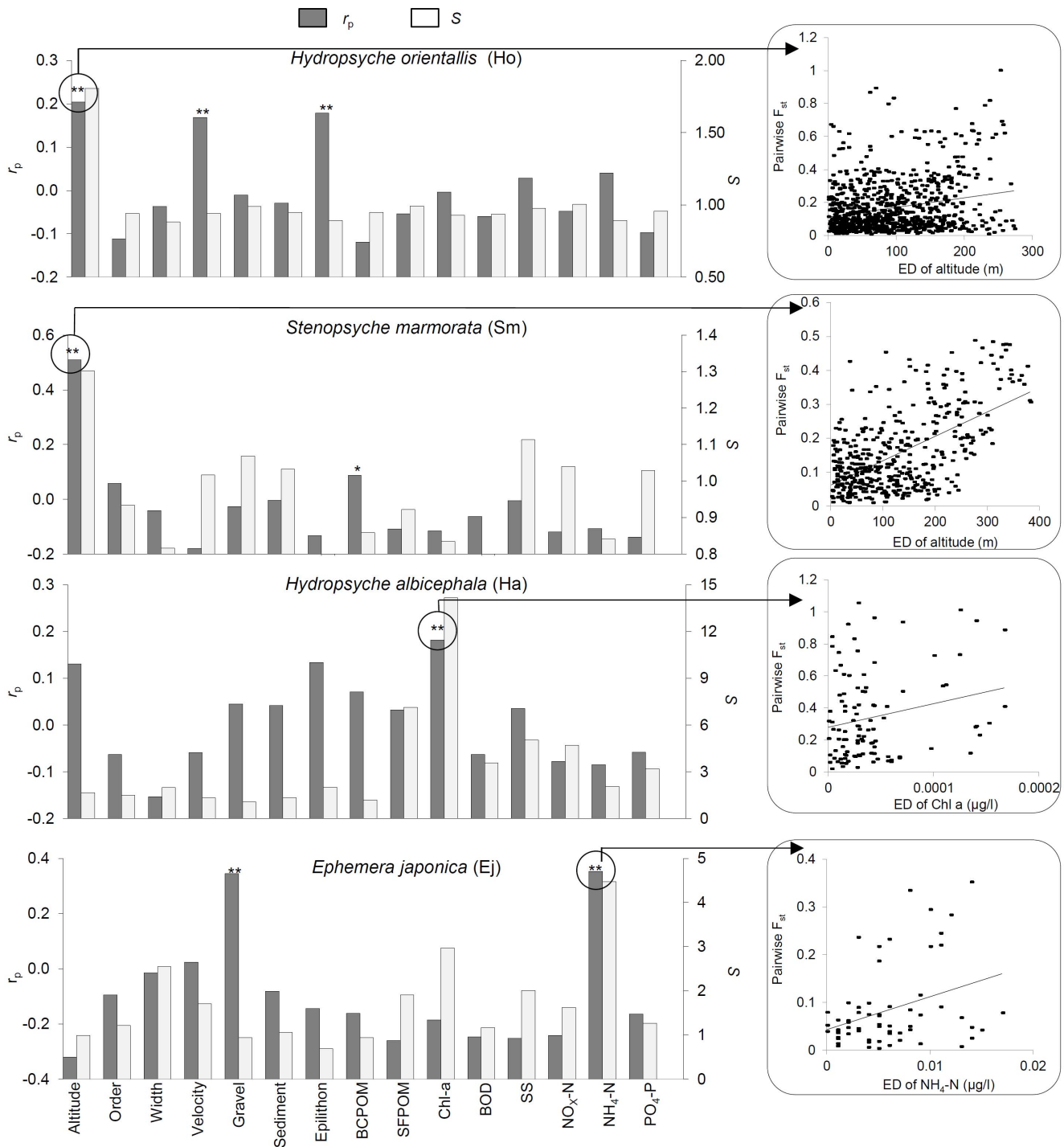


Figure 2. Specialization index (open bars) and genetic-environment correlation coefficients at outlier loci (filled bars) for each species and each environmental variable. Partial Mantel tests were used to calculate r_p in order to control for any effect of geographic distance. ** = significant partial Mantel correlation ($p < 0.001$, randomizations test). In each species, highest r_p and highest S occurred with the same environmental variable. These were altitude (Ho, Sm), transported chlorophyll-*a* (Ha), and ammonium-nitrogen (Ej). Correlations with highest r_p are shown in the right-hand panels. doi:10.1371/journal.pone.0093055.g002

scale of measurement. Altitude was most strongly correlated with non-neutral pairwise F_{st} in two species (Ho and Sm). Altitude itself is related to a number of resource and habitat variables in stream ecosystems, some of which were measured here (e.g., stream size, water velocity), and some of which were not measured (solar

insolation, primary production). Temperature is often an *a priori* factor for which adaptation is tested (e.g., [17]) because of its overwhelming influence on organismal physiology. In stream environments like those studied here, altitude is often strongly associated with mean annual water temperature, and the stronger

Table 2. Summary of genetic structure and diversity measured using all loci, neutral loci, and outlier loci.

Statistic	Species			
	<i>Hydropsyche orientalis</i> (Ho)	<i>Stenopsyche marmorata</i> (Sm)	<i>Hydropsyche albicephala</i> (Ha)	<i>Ephemera japonica</i> (Ej)
F_{ST}				
All loci	0.04 (.18)	0.07 (.14)	0.09 (.18)	0.04 (.28)
Neutral loci	0.00	0.03	0.03	0.02
Outlier loci	0.23	0.22	0.34	0.22
H_T				
Neutral loci	0.11	0.11	0.12	0.07
Outlier loci	0.25	0.22	0.43	0.18
H_w				
Neutral loci	0.11	0.12	0.12	0.07
Outlier loci	0.20	0.17	0.28	0.14

F_{ST} = Wright's fixation index among sampling sites (with standard error in parentheses); H_T = total expected heterozygosity within species; H_w = mean expected heterozygosity within sampling sites.

doi:10.1371/journal.pone.0093055.t002

genetic selection for altitude compared to temperature may well reflect the fact that altitude is a better proxy for complex differences in thermal regime than our two point measurements in the course of one year. Species' distributions are often constrained by extreme thermal conditions (i.e. maxima or minima) rather than mean conditions, therefore, in future study, further analysis using more detailed data such as time-series thermal measures may be feasible. For example, there is experimental evidence for thermal limits in the two species that showed strongest adaptation along the altitudinal gradient. Ho cannot oviposit in air temperatures lower than 11°C (unpublished data) and Sm cannot pupate or emerge in water temperatures lower than 11°C and 13°C, respectively [56].

Ammonium-nitrogen and chlorophyll-*a* that reflect benthic and planktonic algae also exhibited clear relationships with pairwise F_{st} . Like altitude, nitrogen may act as a general indicator of environmental pollution and habitat degradation. The sensitivity of mayflies to such factors is well documented in biomonitoring studies [57]. The link with chlorophyll-*a* in the water column is less clear, but this can be influenced by benthic algae that has been dislodged from the stream bottom or planktonic algae that has been transported downstream from standing water zones. Both are important components of the diet of filter-feeding Trichoptera. Other unmeasured factors such as disease [58] or predation [59] also may play a role in adaptive divergence.

The observed genetic divergences at loci under selection were maintained despite relatively high levels of gene flow observed at neutral loci. This is in contrast to previous studies that were conducted at larger spatial scales and reported some level of neutral population differentiation (e.g., [14,16]) and indicates that selection pressures are strong enough to counter homogenization by migration [60]. The fact that species showed little or no population differentiation at neutral loci is similar to what has been observed in other riverine caddisflies at the spatial scale examined here (up to ca. 50 km; see [61,62]) although it is risky to attempt a generalization across species [63]. Individuals are therefore able to disperse widely, with local conditions determining the genetic makeup of those individuals who persist.

Genetic structure using neutral and non-neutral loci

By comparing neutral and non-neutral genetic differentiation, we can draw some inferences about the process of genetic

divergence in natural populations in a way analogous to $F_{ST} - Q_{ST}$ comparisons, where F_{ST} represents neutral variation and Q_{ST} represents variation in quantitative trait such as morphology [64]. Similar selection pressures (i.e., stabilizing selection) acting throughout the range of sites should lead to lower differentiation at non-neutral loci compared to neutral loci which are primarily governed by genetic drift. We did not observe this pattern, but found the opposite pattern of diversifying selection showing higher non-neutral differentiation than neutral differentiation. Several empirical studies for animals such as snails [65], spiders [66], and damselflies [67–69] compared levels of divergence between morphological traits (analogous to non-neutral markers) and neutral molecular markers. Most of these studies support our results, showing greater morphological divergence than differentiation as measured through F_{ST} based on neutral markers. However, one damselfly study observed the opposite pattern [70], and another damselfly study reported a temporal shift of the pattern from stabilizing selection (lower morphological divergence) to diversifying selection over the course of two generations [71].

Genetic diversity was higher when measured with non-neutral loci compared to neutral loci, suggesting that adaptive divergence is an important driver of total genetic diversity at this spatial scale [72]. The greater diversity within subpopulations at non-neutral loci supports a theory of migration-selection balance which predicts that, as long as adaptive divergence is strong on the spatial scale at which gene flow occurs [73–75], diversity is increased through mixing of diverse alleles from different subpopulations. One shortcoming of our study design is that selection that occurs within individual sampling sites (i.e., along microhabitat gradients) would be obscured because our analysis of alleles and environmental variables was at the site-scale. Thus while our results provide clear evidence of adaptive divergence among different stream reaches (ca. 300 m long), it may also be that within-site, micro-habitat scale divergence also contributes to total genetic diversity. From a conservation standpoint, our results suggest that ecologically marginal habitats, for example those that may occur at the edge of a species' range of niche, harbor genetic diversity that is potentially important for adaptive diversification. When marginal habitats are lost, rare genotypes adapted to these margins will be lost.

Supporting Information

File S1 Table S1, the 31 environmental variables measured at each site and Pearson's correlation coefficients between them. Table S2, oligonucleotide primer pairs used for selective amplification (AFLP PCR). Table S3, characteristics of the 15 non-redundant environmental variables and their relationship to allele frequency at outlier loci. (DOCX)

Acknowledgments

We thank Hikaru Kusano and Yuji Kikuchi for assistance with fieldwork and laboratory analysis and Thomas Mehner for helpful discussion.

References

- Hendry AP (2009) Ecological speciation! Or the lack thereof? *Can J Fish Aquat Sci* 66: 1383–1398.
- Schluter D (2009) Evidence for ecological speciation and its alternative. *Science* 323(5915):737–741.
- Nosil P, Vines TH, Funk DJ (2005) Perspective: Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* 59: 705–719.
- Saint-Laurent R, Legault M, Bernatchez L (2003) Divergent selection maintains adaptive differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus mordax* Mitchell). *Mol Ecol* 12: 315–330.
- Van Valen L (1965) Morphological variation and width of ecological niche. *Am Nat* 99: 377–389.
- Hawthorne DJ, Via S (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* 412: 904–907.
- Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Lett* 7: 1225–1241.
- Leimu R, Fischer M (2008) A meta-analysis of local adaptation in plants. *PLoS ONE* 3: e4010.
- Hereford J (2009) A quantitative survey of local adaptation and fitness trade-offs. *Am Nat* 173: 579–588.
- Mitchell-Olds T, Willis JH, Goldstein DB (2007) Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nat Rev Genet* 8: 845–856.
- Akey JM, Ruhe AL, Akey DT (2010) Tracking footprints of artificial selection in the dog genome. *Proc Natl Acad Sci USA* 107: 1160–1165.
- Counterman BA, Arajuo-Perez F, Hines HM, Baxter SW, Morrison CM, et al. (2010) Genomic Hotspots for adaptation: population genetics of Mullerian mimicry in *Heliconius erato*. *PLoS Genet* 6: e1000796.
- Carneiro M, Afonso S, Geraldes A, Garreau H, Bolet G, et al. (2011) The genetic structure of domestic rabbits. *Mol Biol Evol* 28: 1801–1816.
- Bonin A, Miaud C, Taberlet P, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the commonfrog (*Rana temporaria*). *Mol Biol Evol* 23: 773–783.
- Nosil P, Funk DJ, Ortiz-Barriontos D (2009) Divergent selection and heterogeneous genomic divergence. *Mol Ecol* 18: 375–402.
- Cox K, Vanden Broek A, Van Calster H, Mergeay J (2011) Temperature-related natural selection in a wind-pollinated tree across regional and continental scales. *Mol Ecol* 20: 2724–2738.
- Narum SR, Campbell NR, Kozfay CC, Meyer KA (2010) Adaptation of redband trout in desert and montane environments. *Mol Ecol* 19: 4622–4637.
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proc R Soc Lond B Biol Sci* 263: 1619–1626.
- Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol* 13: 969–980.
- Fischer MC, Foll M, Excoffier L, Heckel G (2011) Enhanced AFLP genome scans detect local adaptation in high-altitude populations of a small rodent (*Microtus arvalis*). *Mol Ecol* 20: 1450–1462.
- Foll M, Gaggiotti OE (2006) Identifying the environmental factors that determine the genetic structure of populations. *Genetics* 174: 875–891.
- Foll M, Gaggiotti OE (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 180: 977–993.
- Ward JV.1992. *Aquatic Insect Ecology 1*, New York: Wiley.
- Hirzel AH, Hauser J, Chessel D, Perrin N (2002) Ecological-niche factor analysis: How to compute habitat- suitability maps without absence data? *Ecology* 83: 2027–2036.
- Hauer FR, Lamberti GA (1996) *Methods in stream ecology*. San Diego (CA): Academic Press.
- Watanabe K, Monaghan MT, Takemon Y, Omura T (2008) Biodilution of heavy metals in a stream macroinvertebrate food web: evidence from stable isotope analysis. *Sci Total Environ* 394: 57–67.
- Vos PR, Hogers M, Bleeker M, Reijfers M, van de Lee T, et al. 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Res* 23: 4407–4414.
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proc R Soc Lond B Biol Sci* 263: 1619–1626.
- Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Mol Ecol* 8: 907–913.
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10: 564–567.
- Nei M, Maruyama T (1975) Lewontin-Krakauer test for neutral genes—comment. *Genetics* 80: 395.
- Robertson A (1975) Lewontin-Krakauer test for neutral genes—comment. *Genetics* 80: 396.
- Bonin A, Ehrlich D, Manel S (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Mol Ecol* 16: 3737–3758.
- Caballero A, Quesada H, Rolan-Alvarez E (2008) Impact of amplified fragment length polymorphism size homoplasy on the estimation of population genetic diversity and the detection of selective loci. *Genetics* 179: 539–554.
- Pe' rez-Figueroa A, Garcia-Pereira M, Saura M, Rola'n-Alvarez E, Caballero A (2010) Comparing three different methods to detect selective loci using dominant markers. *J Evol Biol* 23: 2267–2276.
- Antao T, Beaumont MA (2011) Mchexa: A workbench to detect selection using dominant markers. *Bioinformatics* 27: 1717–1718.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B* 57: 289–300.
- Vekemans X (2002) AFLP-SURV version 1.0. Distributed by the author. Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.
- Manly BJF (1997) *Randomization, bootstrap and Monte Carlo methods in biology*. London: Chapman and Hall.
- Raufaste N, Rousset F (2001) Are partial Mantel tests adequate? *Evolution* 55: 1703–1705.
- Legendre P, Fortin M-J (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Mol Ecol Resour* 10: 831–844.
- Yang RC (2004) A likelihood-based approach to estimating and testing for isolation by distance. *Evolution* 58: 1839–1845.
- Frichot E, Schoville SD, Bouchard G, François O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol Biol Evol* 30: 1687–1699.
- Van Strien MJ, Keller D, Holderegger R (2012) A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Mol Ecol* 21: 4010–4023.
- Mendez M, Rosenbaum HC, Subramaniam A, Yackulic C, Bordino P (2010) Isolation by environmental distance in mobile marine species: molecular ecology of franciscana dolphins at their southern range. *Mol Ecol* 19: 2212–2228.
- Li YF, Costello JC, Holloway AK, Hahn MW (2008) "Reverse ecology" and the power of population genomics. *Evolution* 62: 2984–2994.
- Hirzel AH, Hauser J, Chessel D, Perrin N (2002) Ecological-niche factor analysis: How to compute habitat- suitability maps without absence data? *Ecology* 83: 2027–2036.
- Poisot T, Bever JD, Nemri A, Thrall PH, Hochberg ME (2011) A conceptual framework for the evolution of ecological specialisation. *Ecology Lett* 14: 841–851.
- Haldane J (1956) The relation between density regulation and natural selection. *Proc R Soc Lond B Biol Sci* 145: 306–308.
- Kirkpatrick M, Barton N (1997) Evolution of a species' range. *Am Nat* 150: 1–23.
- Kawecki TJ (2008) Adaptation to marginal habitats. *Ann Rev Ecol Evol Syst* 39: 321–342.
- Mayr E (1963) *Animal species and evolution*. Cambridge: Harvard University Press.
- Antonovics J (1976) The nature of limits to natural selection. *Ann Mo Bot Gard* 63: 224–247.

Yasuhiro Takemon kindly helped taxonomic identification. Yoshihisa Suyama graciously allowed us to use their laboratory facilities in the Graduate School of Agriculture, Tohoku University, and Motoshi Tomita provided assistance genotyping.

Author Contributions

Conceived and designed the experiments: KW TO. Performed the experiments: KW SK. Analyzed the data: KW MTM. Contributed reagents/materials/analysis tools: KW MTM. Wrote the paper: KW MTM.

54. Van Valen L (1965) Morphological variation and width of ecological niche. *Am Nat* 99: 377–389.
55. Ackermann M, Doebeli M (2004) Evolution of niche width and adaptive diversification. *Evolution* 58: 2599–2612.
56. Aoya K, Yokoyama N (1987) Life cycle of two species of *Stenopsyche* (Trichoptera: Stenopsychidae) in Tohoku district. *Jpn J Limnol* 48: 41–53. (In Japanese.)
57. Brittain JE, Sartori M (2003) Ephemeroptera (Mayflies). In: Resh VH, Carde RT, editors. *Encyclopedia of Insects*. Amsterdam: Academic Press. pp. 373–380.
58. Tollenare C, Duplantier JM, Rahalison L, Ranjalaly M, Brouat C (2011) AFLP genome scan in the black rat (*Rattus rattus*) from Madagascar: detecting genetic markers undergoing plague-mediated selection. *Mol Ecol* 20: 1026–1038.
59. Cousyn C, De Meester L, Colbourne JK, Brendonck L, Verschuren D, et al. (2001) Rapid, local adaptation of zooplankton behavior to changes in predation pressure in the absence of neutral genetic changes. *Proc Natl Acad Sci USA* 98: 6256–6260.
60. Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nat Rev Gene* 11: 697–709.
61. Hughes JM (2007) Constraints on recovery: using molecular methods to study connectivity of aquatic biota in rivers and streams. *Freshwater Biol* 52: 616–631.
62. Watanabe K, Monaghan MT, Takemon Y, Omura T (2010) Dispersal ability determines the genetic effects of habitat fragmentation caused by reservoirs in three species of aquatic insect. *Aquat Conservat Mar Freshwat Ecosyst* 20: 574–579.
63. Wilcock HR, Bruford MW, Nichols RA, Hildrew AG (2007) Landscape, habitat characteristics and the genetic population structure of two caddisflies. *Freshwater Biol* 52: 1907–1929.
64. McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends Ecol Evol* 17: 285–291.
65. Cook LM (1992) The neutral assumption and maintenance of color morph frequency in mangrove snails. *Heredity* 69: 184–189.
66. Gillespie RG, Oxford GS (1998) Selection on the color polymorphism in Hawaiian happy-face spiders: evidence from genetic structure and temporal fluctuations. *Evolution* 52: 775–783.
67. Andrés JA, Sánchez-Guillén RA, Cordero Rivera A (2000) Molecular evidence for selection on female colour polymorphism in the damselfly *Ischnura graellsii*. *Evolution* 54: 2156–2161.
68. Wong A, Smith ML, Forbes MR (2003) Differentiation between subpopulations of a polychromatic damselfly with respect to morph frequencies, but not neutral genetic markers. *Mol Ecol* 12: 3505–3513.
69. Abbott JK, Bensch S, Gosden TP, Svensson EI (2008) Patterns of differentiation in a colour polymorphism and in neutral markers reveal rapid genetic changes in natural damselfly populations. *Mol Ecol* 17: 1597–1604.
70. Andrés JA, Sánchez-Guillén RA, Cordero Rivera A (2000) Molecular evidence for selection on female colour polymorphism in the damselfly *Ischnura graellsii*. *Evolution* 54: 2156–2161.
71. Abbott JK, Bensch S, Gosden TP, Svensson EI (2008) Patterns of differentiation in a colour polymorphism and in neutral markers reveal rapid genetic changes in natural damselfly populations. *Mol Ecol* 17: 1597–1604.
72. Lynch M, Walsh JB (1998) *Genetics and Analysis of Quantitative Traits*, Sunderland (MA): Sinauer Associates, Inc.
73. Barton NH (1999) Clines in polygenic traits. *Genet Res* 74: 223–236.
74. Spichtig M, Kawecki TJ (2004) The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments. *Am Nat* 164: 70–84.
75. Yeaman S, Jarvis A (2006) Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proc R Soc Lond B Biol Sci* 273: 1587–1593.