

# Isolation by adaptation in *Neochlamisus* leaf beetles: host-related selection promotes neutral genomic divergence

DANIEL J. FUNK,\* SCOTT P. EGAN\*† and PATRIK NOSIL\*‡

\*Department of Biological Sciences, Vanderbilt University, Box 351634 Station B, Nashville, TN 37235, USA, †Department of Biological Sciences, Advanced Diagnostics and Therapeutics, University of Notre Dame, Notre Dame, IN, 46556, USA, ‡EBIO, University of Colorado, Ramaley N122, Campus Box 334, Boulder, CO 80309, USA

## Abstract

This study tests how divergent natural selection promotes genomic differentiation during ecological speciation. Specifically, we use adaptive ecological divergence (here, population divergence in host plant use and preference) as a proxy for selection strength and evaluate the correlation between levels of adaptive and genetic differentiation across pairwise population comparisons. Positive correlations would reveal the pattern predicted by our hypothesis, that of ‘isolation by adaptation’ (IBA). Notably, IBA is predicted not only for selected loci but also for neutral loci. This may reflect the effects of divergent selection on neutral loci that are ‘loosely linked’ to divergently selected loci or on geneflow restriction that facilitates genetic drift at all loci, including neutral loci that are completely unlinked to those evolving under divergent selection. Here, we evaluate IBA in maple- and willow-associated populations of *Neochlamisus bebbianae* leaf beetles. To do so, we collected host preference data to construct adaptive divergence indices and used AFLPs (amplified fragment length polymorphisms) and mitochondrial sequences to quantify genetic differentiation. Partial Mantel tests showed significant IBA in ‘pooled’ analyses of putatively neutral and of putatively selected (‘outlier’) AFLP loci. This pattern was also recovered in 12% of ‘locus-specific’ analyses that separately evaluated genetic differentiation at individual neutral loci. These results provided evidence for widespread effects of selection on neutral genomic divergence. Our collective findings indicate that host-related selection may play important roles in the population genomic differentiation of both neutral and selected gene regions in herbivorous insects.

**Keywords:** adaptive divergence, ecological speciation, genetic population differentiation, *Neochlamisus bebbianae*, neutral evolution, population genomics

Received 6 April 2011; revision received 29 August 2011; accepted 7 September 2011

## Introduction

This study tests the hypothesis that a positive relationship exists between the degree of adaptive divergence and the degree of genetic differentiation exhibited across pairs of populations, independent of geographic distance (e.g. Lu & Bernatchez 1999; Ogden & Thorpe 2002; Crispo *et al.* 2006). This pattern illustrates a phe-

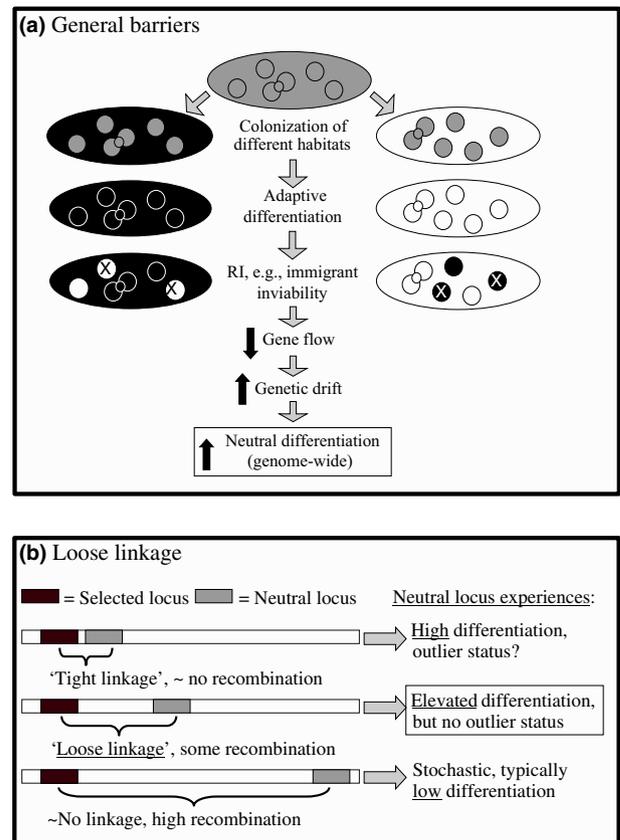
nomenon that has recently been termed ‘isolation by adaptation’ (=IBA; Nosil *et al.* 2008), a word choice that highlights IBA’s analogy to ‘isolation by distance’ (=IBD; Wright 1943; Slatkin 1993; Rousset 1997), the commonly observed positive correlation between geographic distance and neutral genetic divergence. The estimates of ecologically adaptive population divergence used to evaluate IBA represent proxies for the strength of the divergent selection that actually drives genetic differentiation (Schluter 2000), with variation in selection strength across population pairs producing the

Correspondence: Daniel J. Funk, Fax: (615) 343 6707; E-mail: daniel.j.funk@vanderbilt.edu

IBA pattern. It may thus be intuitive that IBA should be observed for divergently selected loci or the loci tightly linked to them. However, furthering its analogy with IBD is the fact that IBA, itself a consequence of divergent natural selection, is predicted to affect even selectively neutral loci that are not tightly linked to selected ones. Thus, IBA could have important consequences for neutral genomic divergence. At least two mechanisms predict this pattern (Nosil *et al.* 2008). Empirically determining the relative roles played by these two mechanisms has scarcely been attempted to date and is beyond the scope of the present study. However, these mechanisms are described here to provide a conceptual framework for thinking about IBA.

First, this pattern could arise via a 'general barriers' to gene flow mechanism (Fig. 1a), whereby greater adaptive divergence results in greater restriction of gene flow between populations, thus facilitating genetic drift at neutral loci (see Barton & Bengtson 1986; Feder & Nosil 2010; Thibert-Plante & Hendry 2010). This mechanism's name invokes both this increased barrier to genetic exchange and the genome-wide effects of increased drift, which affects allele frequencies at all loci. Adaptive divergence might produce such general barriers if it incidentally promotes the evolution of reproductive isolation (Muller 1942; Mayr 1947, 1963), for example, if local adaptation directly reduces the frequency or fitness of migrants in the foreign population, causing habitat isolation or immigrant inviability (Nosil *et al.* 2005). General barriers can yield IBA when levels of adaptive divergence vary among population pairs (see Gavrillets 2004, pp. 147–148 for summary of underlying theory) and gene flow occurs at the spatial scale of population sampling.

Second, the pattern of IBA could reflect 'loose linkage', whereby adaptive divergence affects the frequencies of neutral alleles at loci that are physically linked to those under selection, but somewhat distantly so (Fig. 1b). This can occur because the effects of such genetic hitchhiking may persist at considerable chromosomal distances (Charlesworth *et al.* 1997; Nielsen 2005), particularly if selection acts on numerous loci (Feder & Nosil 2010). Under this scenario, loosely linked loci differentiate more than unlinked neutral alleles. However, although tightly linked neutral loci may sometimes differentiate sufficiently to be identified as putatively divergently selected 'outliers' in genome scans (Beaumont & Nichols 1996; Vitalis *et al.* 2001; Beaumont & Balding 2004; Nielsen 2005; Bonin *et al.* 2007), the same tendency is not expected of loosely linked loci, which should exhibit more modest levels of differentiation. This reflects the greater distance-determined rate of recombination between selected loci and loosely linked as compared to tightly linked neutral



**Fig. 1** Diagrammatic illustration of mechanisms yielding isolation by adaptation (IBA) at neutral loci not tightly linked to loci under selection. IBA may also be observed for selected loci, but we emphasize here the less intuitive association between divergent adaptation and neutral divergence. (a) 'General barriers', under which adaptive differentiation to alternative environments incidentally promotes the evolution of reproductive isolation (RI), causing reduced gene flow and increased genetic drift across the genome. The diagram illustrates divergent adaptation in separate populations, leading to mortality of immigrant individuals (circles with an 'X') prior to mating (joined circles). (b) 'Loose linkage', under which genetic hitchhiking can yield greater differentiation of neutral loci than expected under drift, even for rather distantly linked loci. Bars represent chromosomes.

loci. As is true of the general barriers mechanism, the degree to which loci affected by loose linkage diverge—and thus the likelihood and number of loci exhibiting IBA—should also vary according to variation in the absolute strength of divergent selection exhibited across population pairs (Barton 2000). Notably, however, only linkage can produce IBA at neutral loci among populations experiencing no gene flow (i.e. in allopatry). This is because without gene flow, drift (the basis for divergence under general barriers) should proceed in a manner that is independent of the level of adaptive divergence.

The question of the degree to which natural selection vs. genetic drift drives evolution at the molecular genetic level has dominated population genetic discussion since the advent of allozyme analysis (Kimura 1968; King & Jukes 1969; Gillespie 1991; Ohta 1992; Kreitman & Akashi 1995). The study of IBA can make unique contributions to this discussion—and interestingly elaborate it—by evaluating the proportion of loci in the genome whose phenotypic effects are putatively selectively neutral but whose evolution is nevertheless influenced by selection. By combining such estimates with the estimates of selection's effects from outliers identified in genome scans, a still more comprehensive understanding of selection's role in genomic evolution can be achieved. However, most past studies that have usefully tested for IBA either pooled across loci to obtain population-level estimates of genetic differentiation or treated modest numbers of individual loci (e.g. Lu & Bernatchez 1999; Owuor *et al.* 1999; Cooper 2000; Grahame *et al.* 2006; Pilot *et al.* 2006; Gagnaire *et al.* 2009; Nosil *et al.* 2009 for review; Mendez *et al.* 2010). By contrast, a recent study has individually tested hundreds ( $n = 209$ ) of loci for evidence of IBA (Nosil *et al.* 2008), thus providing enough data for a quantitative estimate of the proportion of the genome influenced by this phenomenon. The present paper describes another such study.

Here, we test for patterns of IBA (and IBD) across pairs of leaf beetle populations. Specifically, we describe the collection of new host preference data from 1001 individuals to calculate indices of adaptive divergence for 15 pairwise population comparisons. To evaluate IBA/IBD, we calculate partial correlations between these indices and estimates of genetic differentiation derived from 75 newly collected ca. 1 kb mtDNA sequences and from AFLP (amplified fragment length polymorphisms) loci that were previously collected for an ecologically comparative genome scan of these populations (Egan *et al.* 2008a,b). The current findings extend our nascent understanding of ecologically divergent selection's varied contributions to population genomic differentiation.

## Methods

### *Study system and sampling localities*

*Neochlamisus bebbiana* (Brown) (Coleoptera: Chrysomelidae) is an eastern North American leaf beetle that uses several tree species as host plants, on which all life activities are conducted (Karren 1972; Funk 1998). Populations associated with different host plants are partially ecologically differentiated in host-use traits and exhibit partial reproductive isolation as a consequence

of divergent host-adaptation (Funk 1998; Funk *et al.* 2002; Egan & Funk 2006, 2009; Funk & Nosil 2008; Egan *et al.* 2011), thus providing informative systems for investigating ecological speciation. This is true, for example, of the focal populations of our study, six incompletely differentiated populations associated with red maple (*Acer rubrum*; Aceraceae) vs. Bebb's willow (*Salix bebbiana*; Salicaceae), between which some gene flow occurs. Gene flow is indicated, for example, by moderate  $F_{ST}$  values across paired population comparisons and by the homogenization of putatively neutral loci in our study's sympatric population comparison (Egan *et al.* 2008a,b), as well as by low levels and polyphyletic patterns of mtDNA variation among and within host-associated populations (Funk 1999; Funk unpublished data).

The present study treats six populations from localities spanning the broad region of northeastern North America in which maple- and willow-associated *N. bebbiana* populations commonly co-occur in the same habitats: Oswego Co., New York; Suffolk Co., Long Island, New York; Wyoming Co., Pennsylvania; near East Angus, Quebec; and Caledonia Co., Vermont, our sympatric site (Fig. 2). This region thus includes many additional, unsampled, sympatric maple- and willow-associated populations, with no obvious geographical barriers to gene flow among them. Study populations were chosen, in part, to exhibit varying degrees of geographic distance between population pairs, these being 162–529 km apart, save for the sympatric pair. These geographic distances were calculated using the ruler tool within Google Earth (version 5.2.1.1329). All data derive from beetle larvae, pupae or adults that were hand-collected across many different

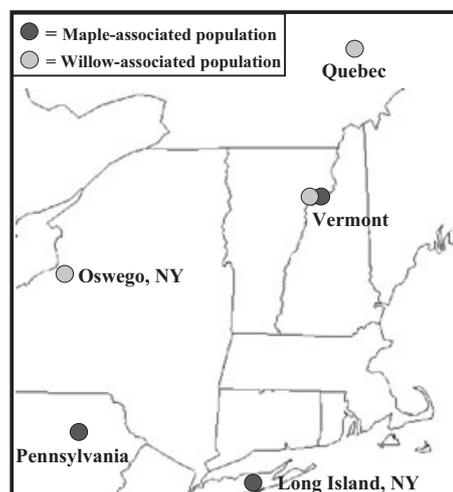


Fig. 2 Map of study populations. Vermont populations are sympatric. Modified from Egan *et al.* (2008a,b).

plant individuals per study population and subsequently maintained and tested on young foliage from northeastern plant genotypes in the laboratory.

Some caveats about our study populations are worth noting. First, not only current gene flow among these populations but also their ancestral colonization history could influence their patterns of genetic differentiation, such as IBD (Le Corre & Kremer 1998). Disentangling these contributions is beyond the scope of the present study but could be investigated in future work, for example, via coalescent analyses of DNA sequence data. For the present, we note that to the degree that ancestral history has influenced genetic differentiation in a manner countering gene flow, we should not be able to detect IBA. Second, the existence of unsampled populations could have implications for population genetic estimates such as migration rates (e.g. Beerli 2004; Slatkin 2005). Beerli (2004), however, concludes that for 'real world' analyses, there is no reason to sample every population in the neighbourhood of a population of interest. Unsampled populations are another issue worthy of future consideration in this system but would be of greater concern if the predicted patterns of IBA were not detected by our study (see Results). Third, although our sampling design was sufficient for us to evaluate IBA/IBD and is similar to the haphazard population sampling used for related studies, a more ideal design might have included many independent pairs of allopatric and sympatric populations.

#### No-choice feeding tests and adaptive divergence indices

Testing for IBA required indices of adaptive population divergence in traits believed to be the subject of the divergent selection promoting this phenomenon. Two host-associated indices were used (Table 1). These reflect divergent adaptation because prior work shows that maple- and willow-associated populations exhibit greatest preference and fitness on their own host plant (Funk 1998; Egan & Funk 2009) and produce hybrids with reduced fitness on both hosts (Egan & Funk 2009; Sæther *et al.* 2007 presents a related example). The first, coarser, index was whether the pair of populations being compared was associated with the same host (=0) or different hosts (=1) in nature. The second index quantified population divergence in host preference based on experimental data from pairs of no-choice trials in which the amount of beetle feeding on maple vs. willow foliage was scored for each test individual in consecutive 24-h trials (Table S1, Supporting information). These trials treated newly emerged adults with no prior feeding experience. For each trial, an individual beetle was placed in a 5-cm petri dish lined with moistened filter paper and included hole-punched leaf discs of the appropriate foliage. Following each trial, feeding on leaf discs was quantified using a gridded ocular under a dissecting microscope. Order of presentation of test plants across trials was systematically varied among individuals to avoid possible bias owing to an

**Table 1** Aspects of divergence for population comparisons in this study

Population comparison	Geographic distance (km)	Host use divergence	Host preference divergence	AFLP $F_{ST}$	% mtDNA divergence
$M_{LI} \times M_{PA}$	162	0	0.034	0.0901	5.77
$M_{LI} \times M_{VT}$	439	0	0.039	0.0757	5.41
$M_{LI} \times W_{OS}$	368	1	0.837	0.1060	5.23
$M_{LI} \times W_{QU}$	529	1	0.837	0.0853	5.29
$M_{LI} \times W_{VT}$	439	1	0.820	0.0937	5.35
$M_{PA} \times M_{VT}$	449	0	0.006	0.0212	1.06
$M_{PA} \times W_{OS}$	238	1	0.803	0.0363	1.57
$M_{PA} \times W_{QU}$	485	1	0.804	0.0339	1.37
$M_{PA} \times W_{VT}$	449	1	0.787	0.0403	1.15
$M_{VT} \times W_{OS}$	376	1	0.797	0.0762	1.56
$M_{VT} \times W_{QU}$	171	1	0.798	0.0455	0.78
$M_{VT} \times W_{VT}$	0	1	0.781	0.0442	0.21
$W_{OS} \times W_{QU}$	325	0	0.001	0.0318	1.31
$W_{OS} \times W_{VT}$	376	0	0.017	0.0457	1.58
$W_{QU} \times W_{VT}$	171	0	0.017	0.0080	0.70

Host use divergence = whether (1) or not (0) populations use the same host plant. Host preference divergence is based on differences in the relative degree to which populations feed on maple. AFLP  $F_{ST}$  values are based on all loci. See text for further details. Populations are identified by a capital letter for the host plant ( $M$  = maple,  $W$  = willow) and a subscripted abbreviation of location (see Fig. 2; LI = Long Island, New York; OS = Oswego, New York, PA = Pennsylvania; QU = Quebec; VT = Vermont).

effect of previous feeding experience in the first trial on expressed host preference in the second trial. Feeding data by each test animal on the two plants were used to calculate an 'individual score', that is, the proportion of total feeding across trials that occurred on maple foliage [=amount fed on maple/(amount fed on maple + amount fed on willow)]. Each 'population-level score' was then calculated as the mean of the individual scores of test animals from a given population. Finally, an index of adaptive divergence (i.e. host preference divergence) for each population pair was obtained as the absolute value of the difference in population-level scores between the two populations. Numbers of individuals assayed were as follows:  $M_{LI} = 97$ ,  $M_{PA} = 101$ ,  $M_{VT} = 41$ ,  $W_{OS} = 180$ ,  $W_{QU} = 75$ ,  $W_{VT} = 507$  (population names abbreviated as in Table 1).

### Mitochondrial sequence divergences

Isolation by adaptation and IBD were evaluated using 903-bp sequences of cytochrome oxidase I that were collected from 75 individuals:  $M_{LI} = 6$ ,  $M_{PA} = 6$ ,  $M_{VT} = 12$ ,  $W_{OS} = 13$ ,  $W_{QU} = 15$ ,  $W_{VT} = 23$ . Although various protocols were used, >2/3 of sequences were obtained as follows: Total genomic DNA was extracted following Bender *et al.* (1983) and purified using Gene-Clean (Qbiogene, Carlsbad, California). PCR included the following steps: (i) 94° for 2 m, (ii) 92° for 1 m, (iii) 50° for 1 m, (iv) 72° for 2 m, (v) go to step 2 and repeat 34 times and (vi) 72° for 5 m. Cocktails and primers followed Funk (1999). PCR products were purified using QIAquick columns (Qiagen, Valencia, CA, USA) and submitted to the Vanderbilt University core facility for sequencing. All individuals were sequenced in both directions and could be unambiguously aligned using Sequencher (GeneCorp, Ann Arbor, MI, USA) because no indels were observed. Sequences were deposited in GenBank (accession nos. JN625283–JN625359). Mean pairwise sequence divergences were calculated for each population comparison using MEGA version 3.1 (Kumar *et al.* 2004) after correcting for multiple hits using the Kimura 2-parameter model (Kimura 1980).

### AFLPs, locus classification and outlier detection

Other estimates of genetic population differentiation used to evaluate IBA and IBD in the present study derived from AFLP-based  $F_{ST}$  calculations from AFLP-SURV (Vekemans *et al.* 2002) (no. beetles:  $M_{LI} = 30$ ,  $M_{PA} = 21$ ,  $M_{VT} = 25$ ,  $W_{OS} = 30$ ,  $W_{QU} = 29$ ,  $W_{VT} = 30$ ) (Egan *et al.* 2008a,b). These analyses included only those loci that were polymorphic in all 15 pairwise population comparisons, to create a completely balanced matrix used to statistically test for IBA. Because some

loci were not polymorphic in particular population comparisons, this reduced the number of loci we analysed from the 447 we assayed to 381.

The present study focuses on identifying patterns of IBA and especially on putatively neutral loci. Nonetheless, because we also draw inferences from outlier patterns, we provide the reader with some background on outlier-related issues. Most relevantly, Egan *et al.* (2008a,b) had assigned each individual AFLP locus considered in this study to one of the following classes, which are invoked later in the present paper: (i) 'Total loci' = all 381 analysed loci; (ii) 'Neutral loci' = the 260 loci that were not outliers in any population comparison; (iii) 'DH-only outliers' = the 21 loci that were outliers in multiple comparisons in which the pair of populations were associated with different host plant species (i.e. maple and willow), and in no population comparisons for which the populations were associated with the same host (i.e. both using maple or both using willow); and (iv) 'Other outliers' = the 42 loci that were outliers in multiple population comparisons, including at least one comparison involving populations using the same host. These seem most likely to be loci evolving under host-independent selective agents not studied here, such as climate. Alternatively, although seemingly less plausibly, they could represent false positives or host-related loci that have gotten shuffled among populations by low levels of gene flow. Classes (iii) and (iv) include only 'repeated' outliers, that is, those loci observed to be outliers in each of multiple population comparisons, while the 58 loci as yet unaccounted for represent 'nonrepeated' outliers, each observed in only one of the 15 population comparisons. Importantly, DH-only outliers are loci that are consistently outliers between hosts and thus represent those loci likely to be involved in host-associated divergent adaptation. In contrast, Other outliers likely represent loci evolving under host-independent sources of selection.

Outlier detection can potentially be influenced by a number of factors, such as type I error, mutation rate variation, demography and linkage across loci, and approaches to outlier detection in this young field continue to be refined (Black *et al.* 2001; Luikart *et al.* 2003; Beaumont & Balding 2004; Storz 2005; Riebler *et al.* 2008; Butlin 2010). However, although our original outlier assessments were undoubtedly imperfect, it is very unlikely that such issues considerably affect the primary interpretations of the present study, for several reasons. First, many of our major inferences, such as those deriving from 'Total loci' and the mtDNA sequences, do not rely on outlier status in any way. Second, a conservative divergence-based method was used to distinguish neutral from outlier loci using the commonly employed simulation program Dfdist (Beaumont

& Nichols 1996; Beaumont & Balding 2004), which is robust to mutation rates, colonization models, sample sizes and IBD (Beaumont & Nichols 1996). Third, by accepting only repeated outliers (see above) for analysis, we greatly diminished the likelihood that their status reflects type I error (e.g. Bonin *et al.* 2006). Our identification of DH-only outliers was yet more conservative as this classification required ecologically consistent patterns expected under host-related selection, but highly unlikely to occur by chance. Fourth, our earlier observation of low linkage disequilibrium among outliers (Egan *et al.* 2008a,b) argues that we have not grossly overestimated the number of independent regions directly influenced by selection. Fifth, factors such as mutation rate variation are not expected to co-vary with adaptive divergence, a necessity if they are to explain the cross-population-comparison associations indicative of IBA.

#### *IBA and IBD analyses*

We used both simple and partial Mantel tests in our analyses (Manly 1997). A simple Mantel test compares the association between two distance matrices. The significance of this relationship is tested by comparing the linear correlation between the two matrices with a null distribution generated by randomizing rows and columns of one matrix while holding the other constant and recalculating the linear association. A partial Mantel test allows one to compare three distance matrices. In this case, the partial Mantel statistic estimates the correlation between two matrices while controlling for the effect of a third and is computed in the same way as a partial correlation coefficient. The test of significance of the partial Mantel compares this partial correlation coefficient with a null distribution generated by randomizing the rows and columns of one of the two focal matrices, while holding the other two constant, and recalculating the partial correlation coefficients (Goslee & Urban 2007). In this way, we evaluated simple, 'uncorrected' associations between adaptive divergence and genetic divergence, and between geographic distance (set to 0.0001 km for the sympatric comparison) and genetic divergence. We used partial Mantel tests to evaluate the first association while controlling for geographic distance (=IBA) and the second association while controlling for adaptive divergence (=IBD). The program IBD version 1.52 (Bohonak 2002), which reports 1-tailed probabilities, was employed for all tests via the implementation of 10 000 randomizations. This program calculates Mantel's  $r$ , the correlation between two matrices. Analyses were carried out using both raw and log-transformed distance matrices, reflecting debate about which is most appropriate for evaluating

IBD (Bohonak 2002). Negative estimates of  $F_{ST}$  were set to 0.0001 before log transformation, and six loci that had negative estimates for all comparisons were excluded from the log-based analyses. Significance testing using simple Mantel tests is unbiased, although the accuracy of partial Mantels has been debated (Raufaste & Rousset 2001; Rousset 2002). Nonetheless, partial Mantel tests seem to perform well under a variety of conditions (Castellano & Balletto 2002), especially when the independent variables are weakly correlated, as in our study (Mantel tests against geographic distance: host use,  $r = 0.06$ ; host preference,  $r = 0.07$ ). Importantly, Mantel tests are insensitive to data distributions and apply equally to categorical and continuous data (Legendre 2000; Goslee & Urban 2007; Legendre & Fortin 2010).

We first conducted 'pooled' analyses of IBA and IBD using  $F_{ST}$  values derived by combining AFLP loci within each class of loci described above. We similarly evaluated mtDNA, which represents a single and often largely neutral locus. Next, we conducted 'locus-specific' analyses using  $F_{ST}$  values for each separate AFLP locus to identify individual loci exhibiting significant IBA or IBD. Having thus evaluated 381 AFLP loci, approximately 19 positive and 19 negative spurious correlations might be expected because of multiple comparisons and associated type I error at a 1-tailed  $\alpha = 0.05$ . Evidence that our results were not simply a reflection of this statistical artefact was provided if: (i) more than 19 significantly positive associations were observed, (ii) a greater number of significantly positive than significantly negative correlations were observed, as evaluated using binomial tests and (iii) the means of Mantel correlation coefficients significantly exceeded zero, as evaluated using separate one-sample  $t$ -tests of all loci within a class and of the significant loci only.

A binomial test compares the observed data for a nominal variable with two possible values with an expected value. In our binomial test, we compared the observed number of significant positive and negative  $r$ -values in our locus-specific analysis to a random 1:1 expected ratio. A one-sample  $t$ -test compares the mean score of a sample to a known value. Our one-sample  $t$ -test compares the mean of 'Mantel's  $r$ ' values to an expected average distribution centred at zero. We made such comparisons for significant 'Mantel's  $r$ ' values and all 'Mantel's  $r$ ' values generated in our locus-specific analyses.

## **Results**

Because our primary interest in this study was IBA, we focus on these patterns, although we also describe our IBD results. Also, although the simple Mantel tests

provided highly similar results, we focus on the partial Mantel analyses, which provide the most definitive evidence for IBA and IBD.

The pooled-locus analyses statistically supported an overall genomic influence of divergent selection on genetic differentiation (Table 2). Comparing results across classes of AFLP loci showed the strength of support for IBA to vary as follows: DH-only outliers ( $r = 0.724$ ,  $P = 0.006$ )  $\gg$  Neutral loci ( $r = 0.350$ ,  $P = 0.014$ ) = Total loci ( $r = 0.313$ ,  $P = 0.019$ )  $>$  Other outliers ( $r = 0.127$ ,  $P = 0.135$ ). mtDNA did not exhibit IBA ( $r = 0.05$ ,  $P = 0.329$ ). Conversely, pooled AFLPs did not exhibit IBD (e.g. Total loci:  $r = 0.217$ ,  $P = 0.157$ ), while mtDNA did ( $r = 0.63$ ,  $P = 0.007$ ).

Our locus-specific analyses revealed significantly positive associations between adaptive divergence and genetic differentiation (=IBA) in 9.4–12.4% of the 381 AFLP loci across the partial Mantel analyses of different adaptive divergence indices and raw vs. log-transformed data (Table 3, Fig. 3). Comparable proportions of neutral and outlier loci showed IBA (across analyses:

neutral = 10.8–11.9%, outliers = 5.8–12.4%). Similar if modestly lower proportions of loci exhibited IBD (overall = 7.6–9.7%, neutral = 7.7–11.5%, outliers = 5.0–8.3%) compared to IBA. (All percentages were calculated from values in Table 3.)

Our inferences of IBA and IBD from the locus-specific analyses are unlikely to be simple artefacts of type I error (Table 3, Fig. 3). First, the proportion of significantly associated loci (approximately 10%) was twice that expected by chance. Thus, minimally 5% of loci demonstrated IBA/IBD. Second, 5% represents an underestimate because the majority (69.0–100%) of significantly associated loci were positively rather than negatively associated and IBA/IBD are illustrated only by positive associations. Indeed, every binomial analysis yielded a highly significant excess of such positive associations. Third, the distribution of correlation coefficients from these loci always exhibited a mean that was significantly greater than zero. This pattern also held for the pooled analyses, supporting a general genomic tendency towards IBA.

**Table 2** Mantel tests of the pooled-locus analyses for each class of loci and mtDNA

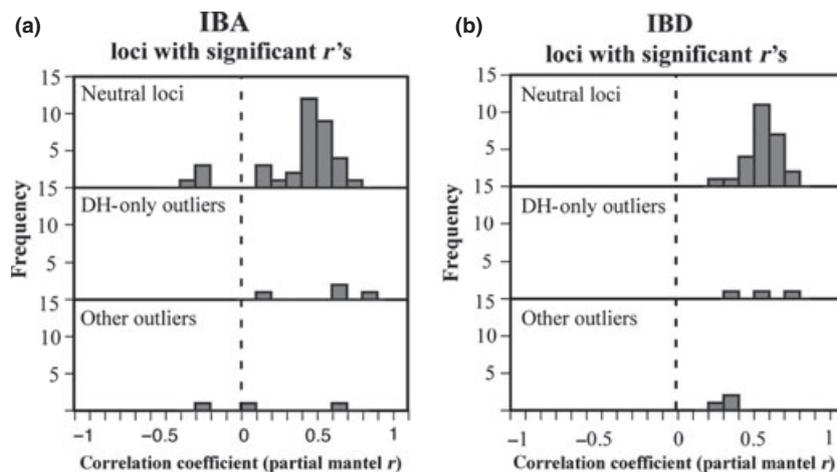
Analysis	Data	Class of AFLP loci									
		Total loci		Neutral loci		DH-only outliers		Other outliers		mtDNA	
		<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>	<i>r</i>	<i>P</i>
AD simple											
Host association	Raw	0.290	0.113	0.329	0.091	0.688	<b>0.001</b>	0.112	0.122	-0.033	0.817
	Log	0.375	0.096	0.417	<b>0.001</b>	0.675	<b>0.001</b>	0.224	<b>0.001</b>	-0.074	0.492
Host preference	Raw	0.322	<b>0.015</b>	0.358	<b>0.012</b>	0.711	<b>0.007</b>	0.139	0.124	-0.038	0.201
	Log	0.400	<b>0.014</b>	0.428	<b>0.014</b>	0.692	<b>0.009</b>	0.245	<b>0.017</b>	-0.071	0.486
AD partial											
Host association	Raw	0.283	0.110	0.323	0.087	0.705	<b>0.027</b>	0.101	0.274	-0.059	0.002
	Log	0.394	<b>0.035</b>	0.426	<b>0.013</b>	0.677	<b>0.030</b>	0.240	<b>0.046</b>	0.054	0.329
Host preference	Raw	0.313	<b>0.019</b>	0.350	<b>0.014</b>	0.724	<b>0.006</b>	0.127	0.135	-0.058	0.003
	Log	0.416	<b>0.011</b>	0.435	<b>0.006</b>	0.692	<b>0.008</b>	0.259	<b>0.037</b>	0.038	0.342
GeoD simple											
Host association	Raw	0.231	0.147	0.180	0.230	0.315	0.105	0.177	0.200	0.346	<b>0.023</b>
	Log	0.066	0.406	0.011	0.277	-0.067	0.452	0.064	0.334	0.635	<b>0.007</b>
GeoD partial											
Host association	Raw	0.222	0.168	0.169	0.242	0.374	0.104	0.171	0.206	0.349	<b>0.016</b>
	Log	0.147	0.240	0.097	0.158	0.077	0.293	0.110	0.296	0.634	<b>0.007</b>
Host preference	Raw	0.217	0.157	0.163	0.221	0.368	0.089	0.168	0.232	0.224	0.166
	Log	0.145	0.219	0.090	0.164	0.064	0.289	0.109	0.315	0.618	<b>0.021</b>

$r$  = association between  $F_{ST}$ , estimated from all AFLP loci within a class, and adaptive divergence or geographic distance, with analogous results for mtDNA. Positive associations are in bold if significant ( $P < 0.05$ ) and in italics if marginally nonsignificant ( $0.10 > P > 0.05$ ). Analyses of adaptive divergence and geographic distance provide evidence on isolation by adaptation (IBA) and isolation by distance (IBD), respectively. AD simple = overall association of  $F_{ST}$  with adaptive divergence (in host association or host preference); AD partial = association of  $F_{ST}$  with adaptive divergence, controlling for geographic distance; GeoD simple = overall association of  $F_{ST}$  with geographic distance. GeoD partial = association of  $F_{ST}$  with geographic distance, controlling for adaptive divergence. Analyses of adaptive divergence and geographic distance provide evidence on IBA and IBD, respectively. Results derive from either raw or log-transformed data.

**Table 3** Distribution of Mantel correlation coefficients from analyses of individual loci

Analysis	Data	No. significant +	No. significant -	Binomial <i>P</i>	Significant loci only			All loci		
					Mean <i>r</i>	<i>t</i> -test	<i>P</i>	Mean <i>r</i>	<i>t</i> -test	<i>P</i>
AD simple										
Host association	Raw	33 (12)	10	<b>0.0006</b>	0.33	6.53	<b>&lt;0.0001</b>	0.08	6.68	<b>&lt;0.0001</b>
	Log	41 (13)	15	<b>0.0007</b>	0.27	5.16	<b>&lt;0.0001</b>	0.08	5.74	<b>&lt;0.0001</b>
Host preference	Raw	45 (13)	4	<b>&lt;0.0001</b>	0.42	11.12	<b>&lt;0.0001</b>	0.09	7.30	<b>&lt;0.0001</b>
	Log	44 (13)	12	<b>&lt;0.0001</b>	0.29	5.64	<b>&lt;0.0001</b>	0.08	6.14	<b>&lt;0.0001</b>
AD partial										
Host association	Raw	36 (8)	8	<b>&lt;0.0001</b>	0.38	7.79	<b>&lt;0.0001</b>	0.08	6.60	<b>&lt;0.0001</b>
	Log	42 (13)	15	<b>0.0005</b>	0.28	4.71	<b>&lt;0.0001</b>	0.08	5.94	<b>&lt;0.0001</b>
Host preference	Raw	38 (7)	5	<b>&lt;0.0001</b>	0.39	8.60	<b>&lt;0.0001</b>	0.09	7.19	<b>&lt;0.0001</b>
	Log	47 (15)	14	<b>&lt;0.0001</b>	0.28	5.27	<b>&lt;0.0001</b>	0.09	6.29	<b>&lt;0.0001</b>
GeoD simple										
Host association	Raw	34 (5)	1	<b>&lt;0.0001</b>	0.48	15.00	<b>&lt;0.0001</b>	0.05	4.36	<b>&lt;0.0001</b>
	Log	30 (7)	6	<b>&lt;0.0001</b>	0.22	4.22	<b>0.0002</b>	0.02	1.71	<i>0.0879</i>
GeoD partial										
Host association	Raw	37 (7)	0	<b>&lt;0.0001</b>	0.52	24.15	<b>&lt;0.0001</b>	0.05	4.16	<b>&lt;0.0001</b>
	Log	29 (9)	13	<b>0.0137</b>	0.15	2.47	<b>0.0179</b>	0.04	2.81	<b>0.0052</b>
Host preference	Raw	35 (6)	0	<b>&lt;0.0001</b>	0.53	24.30	<b>&lt;0.0001</b>	0.05	4.03	<b>&lt;0.0001</b>
	Log	30 (10)	12	<b>0.005</b>	0.16	2.65	<b>0.0113</b>	0.04	2.66	<b>0.0080</b>

Binomial tests evaluated differences in the frequency of loci with significantly positive vs. significantly negative associations. One-sample *t*-tests determined if mean associations across loci were greater than zero. Significant loci only refers to *t*-tests of those loci that exhibited significantly positive or negative *r*-values in the locus-specific analyses. No. significant +, - = number of loci exhibiting positive (with number of repeated outliers in parentheses) vs. negative associations that were significant at  $P < 0.05$ . Results are in bold if significant ( $P < 0.05$ ) and in italics if marginally nonsignificant ( $0.10 > P > 0.05$ ). See notes on Table 2 for further details.



**Fig. 3** The evaluation of locus-specific isolation by adaptation (IBA) and isolation by distance (IBD) using partial Mantel tests. (a) Distributions of correlation coefficients from those individual AFLP loci with significant associations between adaptive divergence (here, in host preference) and  $F_{ST}$ . Loci with positive *r*-values exhibit IBA. (b) Distributions of correlation coefficients from those individual AFLP loci with significant associations between geographic distance and  $F_{ST}$ . Loci with positive *r*-values exhibit IBD. Partial *r*-values for both plots derive from Mantel analyses of the raw data.

## Discussion

This paper provides one of the first investigations of IBA that includes locus-specific analyses of hundreds of loci. The present analyses treat nearly twice as many

AFLP loci as Nosil *et al.* (2008) did in a stick-insect study system and provide yet stronger evidence for widespread IBA in terms of support from the pooled analyses and overall numbers of significant loci. Importantly, they provide a rare many-locus quantitative

estimate of the contributions of divergent selection to genomic differentiation at selectively neutral loci. The interpretation of any such absolute estimates must be tempered by the recognition of earlier-discussed issues that may affect their precision. However, even coarser-scale estimates are critical at this early stage in the development of population genomics. The contributions of this study derive primarily from the demonstration of an under-studied phenomenon (IBA), its exhibition by putatively neutral loci, its nontrivial consequences for genomic divergence, a demonstrated association of IBA with a particular source of divergent selection (host-related selection) and the relative strength of IBA across different classes of loci.

Given the special focus of this study on neutral evolution via IBA, our most important finding may be that adaptive population divergence makes notable contributions to neutral genomic divergence. Egan *et al.* (2008a,b) demonstrated a lack of appreciable within-population linkage disequilibrium within each class of AFLP loci treated here. This apparent physical independence among loci allows our locus-specific-based proportions to be interpreted as approximations of general genomic patterns. These identify 11–12% of neutral gene regions in the genome as exhibiting IBA and thus having evolved under the influence of natural selection. Assuming that a fraction of these loci may reflect type I error, this value may be somewhat inflated. Contrarily, this proportion may represent an underestimate when one considers that increased sampling of population pairs (yielding greater statistical power) and the ability to correct for various biological factors (e.g. recent gene flow that erases the signal of IBA) would surely allow the identification of additional IBA loci. Indeed, both the pooled-locus analyses and the significantly positively skewed distribution of partial correlation coefficients across loci (Table 3) support the hypothesis that IBA makes quite general contributions to genomic differentiation. More specifically, our use of adaptive divergence indices based on host-use ecology argues that observed differentiation can be largely attributed to host-plant-related selection (Funk 1998). The undoubted contribution of additional host-independent sources of selection further argues that we have underestimated the importance of divergent selection *per se* (but see below).

This study also helped corroborate conclusions about host-related selection from the Egan *et al.* (2008a,b) genome scan by assessing them using different analyses and a different index of host-related adaptive divergence (host preference divergence). Notably, we find here that the DH-only outliers provided by far the strongest evidence for IBA in the pooled-locus analyses, while the Other outliers provided the weakest support, even weaker than the Neutral loci. This is as expected if

the earlier ecologically comparative genome scan had accurately identified those loci diverging under host-related selection (i.e. those assigned to the DH-Only outlier class) as well as those loci evolving under host-independent sources of selection, such as climate (i.e. the Other outliers, which were documented in same-host comparisons). That is, while DH-only outliers would be predicted to differentiate and exhibit strong IBA as a function of host preference divergence, Other outliers would not, and this is what we observed. The locus-specific results further complement those of the genome scan. That study found that 5% of loci represented DH-only outliers. Adding that figure to the ca. 12% of total loci exhibiting IBA here increases the estimated genomic divergence resulting from host-related selection to approximately 17%. This illustrates the benefits of combining genome scan and IBA approaches (see also Nosil *et al.* 2009).

Our results are consistent with the hypothesized mechanisms underlying IBA that were described in the Introduction. First, the general barriers assumption that gene flow occurs among study populations is supported by our observation of significant IBD—a pattern frequently (albeit not always) explained by geographically restricted gene flow—in the mitochondrial data. Nearly 10% of individual AFLP loci also exhibited IBD, although no support for this pattern was provided by the pooled analyses. These findings complement the previously mentioned evidence for gene flow in these beetle populations. Second, the considerable evidence for IBA in both the pooled-locus and locus-specific AFLP analyses are consistent with the general barriers mechanism, under which gene flow is restricted by adaptive divergence, yielding general cross-locus effects on genomic differentiation. By contrast, the alternative possibility that gene flow is determining levels of adaptive divergence (Hendry & Taylor 2004; Nosil & Crespi 2004; Räsänen & Hendry 2008) is not supported by our observations of IBA for an adaptive divergence index (based on host association) that is not a phenotypic property of *N. bebbianae*, and thus less likely to be constrained by gene flow.

This study provides no evidence on specific linkage relationships between individual loci, so the contributions of the loose linkage mechanism cannot be directly evaluated here. Nonetheless, comparing our results with the *Timema* stick-insect study (Nosil *et al.* 2008) offers a heuristic, if currently speculative, means of evaluating taxonomic variation in the relative roles of the two proposed mechanisms. Compared to the present study, the *Timema* investigation differed in finding a somewhat lower proportion of loci exhibiting IBA, higher partial *r*'s for those loci that did so and no evidence for IBA in the pooled analyses. These patterns might be observed, for example, if the ratio of contributions from general

barriers vs. loose linkage is greater in *Neochlamisus* than in *Timema*. Clearly, however, teasing apart and comparing the contributions of each mechanism with any degree of rigour will require future studies that, for example, identify linkage relationships (Rogers & Bernatchez 2007; Via & West 2008) and further evaluate whether the spatial scale of gene flow is consistent with a general barriers explanation.

Finally, our investigation provides insights on the outstanding issue of the degree to which divergent adaptation vs. geography restricts gene flow (Crispo *et al.* 2006). This can be evaluated by comparing our IBA and IBD results for our nuclear (AFLP) and mitochondrial data. One might predict a priori that IBD would be observed for both nuclear and mitochondrial loci, but perhaps more strongly in the latter, to the degree that mitochondrial loci are more likely to be neutral and unlinked to loci under selection. IBA should be observed for both genomes if driven by the drift-focused general barriers mechanism but may be less likely observed for mitochondrial loci under loose linkage, again assuming nuclear genes to more likely be targeted by divergent selection. This argument suggests that loose linkage is acting to some extent in our study system, because IBA was detected only for nuclear loci. This possibility warrants further research. Finally, the proportions of individual loci exhibiting IBA and IBD were comparable in our study (Table 3), suggesting that adaptation and geography may make similar contributions to genomic differentiation in *Neochlamisus*. This same finding was observed in the locus-specific analyses in *Timema*. Our pooled analyses also revealed a cross-system pattern, namely, that one source of molecular data exhibited IBA and the other exhibited IBD. Curiously, though, AFLPs exhibited IBA and mtDNA exhibited IBD in *Neochlamisus*, while mtDNA exhibited IBA and AFLPs exhibited IBD in *Timema*. Both these studies treated host-specific herbivorous insect populations that appear to be undergoing speciation (see references above for *Neochlamisus*, Nosil 2007 for *Timema*). And both employed host-plant-associated indices of adaptive divergence in evaluating IBA. Thus, the parallels in their results further corroborate the considerable role that host-related selection plays in promoting neutral and non-neutral differentiation alike during ecological speciation. Additional 'ecologically comparative' studies of this type will help establish how generally important aspects of adaptive divergence are as a major cause of genomic divergence across taxa.

### Acknowledgements

We thank R. Funk and J. Funk for foliage collection, the many students who helped obtain host preference data and P. Prasnik

for obtaining mtDNA sequences. Daniel J. Funk was funded by NSF (DEB #0221262), Scott P. Egan was supported by Vanderbilt University and Patrik Nosil was funded by an NSERC postdoctoral fellowship during this study.

### References

- Barton NH (2000) Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London B*, **355**, 1553–1562.
- Barton N, Bengtson BO (1986) The barrier to genetic exchange between hybridizing populations. *Heredity*, **57**, 357–376.
- Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **263**, 1619–1626.
- Berli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology*, **13**, 827–836.
- Bender W, Spierer P, Hogness DS (1983) Chromosomal walking and jumping to isolate DNA from the ace and rosy loci and the bithorax complex in *Drosophila melanogaster*. *Journal of Molecular Biology*, **168**, 17–33.
- Black WC, Baer CF, Antolin MF, DuTeau NM (2001) Population genomics: genome-wide sampling of insect populations. *Annual Review of Entomology*, **46**, 441–469.
- Bohonak AJ (2002) IBD (isolation by distance): a program for analyses of isolation by distance. *Journal of Heredity*, **93**, 153–154.
- Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution*, **23**, 773–783.
- Bonin A, Ehrlich D, Manel S (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, **16**, 3737–3758.
- Butlin RK (2010) Population genomics and speciation. *Genetica*, **138**, 409–418.
- Castellano S, Balletto E (2002) Is the partial Mantel test inadequate? *Evolution*, **56**, 1871–1873.
- Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genetics Research*, **70**, 155–174.
- Cooper ML (2000) Random amplified polymorphic DNA analysis of southern brown bandicoot (*Isodon obesulus*) populations in Western Australia reveals genetic differentiation related to environmental variables. *Molecular Ecology*, **9**, 469–479.
- Crispo E, Bentzen P, Reznick DN, Kinnison MT, Hendry AP (2006) The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology*, **15**, 49–62.
- Egan SP, Funk DJ (2006) Individual advantages to ecological specialization: insights on cognitive constraints from three conspecific taxa. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **273**, 843–848.

- Egan SP, Funk DJ (2009) Ecologically dependent postmating isolation between sympatric host forms of *Neochlamisus bebbianae* leaf beetles. *Proceedings of the National Academy of Sciences, USA*, **106**, 19426–19431.
- Egan SP, Nosil P, Funk DJ (2008a) Selection and genomic differentiation during ecological speciation: isolating the contributions of host association via a comparative genome scan of *Neochlamisus bebbianae* leaf beetles. *Evolution*, **62**, 1162–1181.
- Egan SP, Nosil P, Funk DJ (2008b) Data from: selection and genomic differentiation during ecological speciation: isolating the contributions of host association via a comparative genome scan of *Neochlamisus bebbianae* leaf beetles. Dryad Digital Repository. doi: 10.5061/dryad.ng09t.
- Egan SP, Janson EM, Brown CG, Funk DJ (2011) Postmating isolation and genetically variable host use in ecologically divergent host forms of *Neochlamisus bebbianae* leaf beetles. *Journal of Evolutionary Biology*, **24**, 2217–2229.
- Feder JL, Nosil P (2010) The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution*, **64**, 1729–1747.
- Funk DJ (1998) Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. *Evolution*, **52**, 1744–1759.
- Funk DJ (1999) Molecular systematics of COI and 16S in *Neochlamisus* leaf beetles and the importance of sampling. *Molecular Biology and Evolution*, **16**, 67–82.
- Funk DJ, Nosil P (2008) Comparative analyses and the study of ecological speciation in herbivorous insects. In: *Specialization, speciation, and radiation: the evolutionary biology of herbivorous insects* (ed. Tilmon K). pp. 117–135, University of California Press, Berkeley.
- Funk DJ, Filchak KE, Feder JL (2002) Herbivorous insects: model systems for the comparative study of speciation ecology. *Genetica*, **116**, 251–267.
- Gagnaire PA, Albert V, Jonsson B, Bernatchez L (2009) Natural selection influences AFLP intraspecific genetic variability and introgression patterns in Atlantic eels. *Molecular Ecology*, **18**, 1678–1691.
- Gavrilets S (2004) *Fitness landscapes and the origin of species*. Princeton University Press, Princeton.
- Gillespie JH (1991) *The causes of molecular evolution*. Oxford University Press, Oxford.
- Goslee SC, Urban DL (2007) The Ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, **22**, 1–19.
- Grahame JW, Wilding CS, Butlin RK (2006) Adaptation to a steep environmental gradient and an associated barrier to gene exchange in *Littorina saxatilis*. *Evolution*, **60**, 268–278.
- Hendry AP, Taylor EB (2004) How much of the variation in adaptive divergence can be explained by gene flow? An evaluation using lake-stream stickleback pairs. *Evolution*, **58**, 2319–2331.
- Karren JB (1972) A revision of the subfamily Chlamisinae north of Mexico (Coleoptera: Chrysomelidae). *University of Kansas Science Bulletin*, **49**, 875–988.
- Kimura M (1968) Evolutionary rate at the molecular level. *Nature*, **217**, 624–626.
- Kimura M (1980) A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.
- King JL, Jukes TH (1969) Non-darwinian evolution. *Science*, **164**, 788–798.
- Kreitman M, Akashi H (1995) Molecular evidence for natural selection. *Annual Review of Ecology and Systematics*, **26**, 403–422.
- Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.
- Le Corre V, Kremer A (1998) Cumulative effects of founding events during colonization on genetic diversity and differentiation in an island and stepping-stone model. *Journal of Evolutionary Biology*, **11**, 495–512.
- Legendre P (2000) Comparison of permutation methods for the partial correlation and partial Mantel tests. *Journal of Statistical Computation and Simulation*, **67**, 37–73.
- 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. *Molecular Ecology Resources*, **10**, 831–844.
- Lu GQ, Bernatchez L (1999) Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): support for the ecological speciation hypothesis. *Evolution*, **53**, 1491–1505.
- Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.
- Manly BFJ (1997) *Randomization and Monte Carlo methods in biology*, 2nd edn. Chapman and Hall, New York.
- Mayr E (1947) Ecological factors in speciation. *Evolution*, **1**, 263–288.
- Mayr E (1963) *Animal species and evolution*. Harvard University Press, Cambridge.
- 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. *Molecular Ecology*, **19**, 2212–2228.
- Muller HJ (1942) Isolating mechanisms, evolution and temperature. *Biological Symposium*, **6**, 71–125.
- Nielsen R (2005) Molecular signatures of natural selection. *Annual Review of Genetics*, **39**, 197–218.
- Nosil P (2007) Divergent host-plant adaptation and reproductive isolation between ecotypes of *Timema cristinae* walking-sticks. *American Naturalist*, **169**, 151–162.
- Nosil P, Crespi BJ (2004) Does gene flow constrain trait divergence or vice-versa? A test using ecomorphology and sexual isolation in *Timema cristinae* walking-sticks. *Evolution*, **58**, 101–112.
- Nosil P, Vines TH, Funk DJ (2005) Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, **59**, 705–719.
- Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes: 'isolation by adaptation' and multiple roles for divergent selection. *Evolution*, **63**, 316–336.
- Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375–402.

- Ogden R, Thorpe RS (2002) Molecular evidence for ecological speciation in tropical habitats. *Proceedings of the National Academy of Sciences, USA*, **99**, 13612–13615.
- Ohta T (1992) The nearly neutral theory of molecular evolution. *Annual Review of Ecology and Systematics*, **23**, 263–286.
- Owuor ED, Fahima T, Beharav A, Korol A, Nevo E (1999) RAPD divergence caused by microsite edaphic selection in wild barley. *Genetica*, **105**, 177–192.
- Pilot M, Jedrzejewski W, Branicki W *et al.* (2006) Ecological factors influence population genetic structure of European grey wolves. *Molecular Ecology*, **15**, 4533–4553.
- Räsänen K, Hendry AP (2008) Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecology Letters*, **11**, 624–636.
- Raufaste N, Rousset F (2001) Are partial Mantel tests adequate? *Evolution*, **55**, 1703–1705.
- Riebler A, Held L, Stephan W (2008) Bayesian variable selection for detecting adaptive genomic differences among populations. *Genetics*, **178**, 1817–1829.
- Rogers SM, Bernatchez L (2007) The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp.: Salmonidae) species pairs. *Molecular Biology and Evolution*, **24**, 1423–1438.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.
- Rousset F (2002) Partial Mantel tests: reply to Castellano and Balletto. *Evolution*, **56**, 1874–1875.
- Sæther SA, Sætre GP, Borge T *et al.* (2007) Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science*, **318**, 95–97.
- Schluter D (2000) *The ecology of adaptive radiation*. Oxford University Press, Oxford.
- Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.
- Slatkin M (2005) Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. *Molecular Ecology*, **14**, 67–73.
- Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.
- Thibert-Plante X, Hendry AP (2010) When can ecological speciation be detected with neutral loci? *Molecular Ecology*, **19**, 2301–2314.
- Vekemans X, Beauwens T, Lemaire M, Roldan-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, **11**, 139–151.
- Via S, West J (2008) The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Molecular Ecology*, **17**, 4334–4345.
- Vitalis R, Dawson K, Boursot P (2001) Interpretation of variation across marker loci as evidence of selection. *Genetics*, **158**, 1811–1823.
- Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

---

All authors adopt a diversity of approaches in studying ecological speciation, especially in herbivorous insects.

---

### Data accessibility

The host preference data set for this study is available as an online supporting information document. All mtDNA sequences have been submitted to GenBank under accession nos JN625283–JN625359. AFLP data have been deposited in the Dryad database at doi: 10.5061/dryad.ng09t.

### Supporting information

Additional supporting information may be found in the online version of this article.

**Table S1** Data for the calculation of population-level host preference.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.