

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

Divergent Selection and Heterogeneous Genomic Divergence

PATRIK NOSIL¹, DANIEL J. FUNK², and DANIEL ORTIZ-BARRIENTOS³

¹ *Zoology Department and Biodiversity Research Centre, University of British Columbia
Vancouver BC, V6T 1Z4, Canada*
pnosil@zoology.ubc.ca

² *Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee, 37235, USA*
daniel.j.funk@Vanderbilt.edu

³ *School of Integrative Biology, University of Queensland, St. Lucia QLD, 4702, Australia*
d.ortizbarrientos@uq.edu.au

1 **Abstract**

2
3 Levels of genetic differentiation between populations can be highly variable across the
4 genome, with divergent selection contributing to such heterogeneous genomic divergence. For
5 example, loci under divergent selection and those tightly physically-linked to them may exhibit
6 stronger differentiation than neutral regions with weak or no linkage to such loci. Divergent
7 selection can also increase genome-wide neutral differentiation by reducing gene flow (e.g., by
8 causing ecological speciation), thus promoting divergence via the stochastic effects of genetic drift.
9 These consequences of divergent selection are being reported in recently accumulating studies that
10 identify: (1) ‘outlier loci’ with higher levels of divergence than expected under neutrality, and (2) a
11 positive association between the degree of adaptive phenotypic divergence and levels of molecular
12 genetic differentiation across population pairs (‘Isolation-By-Adaptation’ = IBA). The latter pattern
13 arises because as adaptive divergence increases, gene flow is reduced (thereby promoting drift) and
14 genetic hitchhiking increased. Here, we review and integrate these previously disconnected
15 concepts and literatures. We find that studies generally report 5-10% of loci to be outliers. These
16 selected regions were often dispersed across the genome, commonly exhibited replicated divergence
17 across different population pairs, and could sometimes be associated with specific ecological
18 variables. IBA was not infrequently observed, even at neutral loci putatively unlinked to those
19 under divergent selection. Overall, we conclude that divergent selection makes diverse
20 contributions to heterogeneous genomic divergence. Nonetheless, the number, size, and distribution
21 of genomic regions affected by selection varied substantially among studies, leading us to discuss
22 the potential role of divergent selection in the growth of regions of differentiation (i.e., genomic
23 islands of divergence), a topic in need of future investigation.

24
25 *keywords:* chromosomal inversions, ecological genetics, F_{ST} , introgression, neutral gene flow,
26 outlier loci, population genomics, QTL mapping, speciation

1 Introduction

2
3 Levels of genetic differentiation can be highly variable across the genome, a pattern we refer
4 to here as ‘heterogeneous genomic divergence’ (Harrison 1991; Avise 2000; Via 2001; Rieseberg
5 2001; Wu 2001; Ortiz-Barrientos et al. 2002; Wu and Ting 2004; Gavrillets and Vose 2005; Mallet
6 2005; Turner et al. 2005; Harr 2006; Begun et al. 2007; Noor and Feder 2006; Mallet et al. 2007;
7 Via and West 2008; see glossary for terminology). Genomic divergence may be particularly
8 heterogeneous during the process of population divergence and speciation, during which genetic
9 differentiation accumulates in some regions, while the homogenizing effects of gene flow or
10 inadequate time for random differentiation by genetic drift precludes divergence in other regions
11 (Wu 2001; Gavrillets and Vose 2005). Many factors potentially contribute to heterogeneous
12 genomic divergence, including selection arising from ecological causes (Schluter 2000; Wu 2001)
13 or genetic conflict (Rice 1998; Presgraves et al. 2003; Haig 2004; Arnqvist and Rowe 2005; Crespi
14 2007; Presgraves 2007), the stochastic effects of genetic drift (Kimura 1968; King and Jukes 1969;
15 Kimura 1986; Ohta 1992, 2002), variable mutation rates (Balloux and Lugon-Moulin 2002; Hedrick
16 2005; Noor and Feder 2006), the genomic distribution and effect size of genes under selection (Orr
17 2005), and chromosomal structure (Noor et al. 2001; Rieseberg 2001; Ortiz-Barrientos et al. 2002).

18 We focus here on the contributions of divergent selection, defined as selection that acts in
19 contrasting directions in two populations (c.f., Schluter 2000; Rundle and Nosil 2005). Divergent
20 selection itself can promote molecular genetic differentiation via two main mechanisms: (1) by
21 acting on specific loci and those physically-linked to them (Fisher 1930; Haldane 1930, 1932;
22 Endler 1973, Lewontin and Krakauer 1973; Barton 2000), and (2) by promoting reproductive
23 isolation that causes barriers to gene flow (i.e., ‘ecological speciation’, Mayr 1963; Funk 1998;
24 Schluter 2000; Rundle and Nosil 2005), thereby facilitating even genome-wide neutral divergence
25 via genetic drift. The first mechanism involves a relatively direct role for selection in genetic
26 differentiation and promotes divergence both in the presence and absence of gene flow (Fig. 1). The
27 second mechanism facilitates differentiation by a different process (drift). This second mechanism
28 applies only to divergence with gene flow because in allopatric scenarios divergent selection is not
29 required to counter gene flow in order for neutral population differentiation to proceed. While these
30 basic mechanisms account for heterogeneity in the origin and frequency of highly differentiated
31 genomic regions, and in their degree of differentiation, selection may also affect the actual size of
32 such ‘islands of genomic divergence’ on a chromosome (Turner et al. 2005; Harr 2006; Begun et al.
33 2007; Turner and Hahn 2007; see glossary).

34 Here, we review these roles for divergent selection in generating heterogeneous genomic
35 divergence, and further consider the nature and growth of islands of genomic divergence. We
36 consider mainly conceptual issues and empirical patterns, because methodology has been well
37 covered elsewhere (e.g., Beaumont and Nichols 1996; Andolfatto 2001; Black et al. 2001;
38 Schlötterer 2002; Luikart et al. 2003; Beaumont and Balding 2004; Beaumont 2005; Nielsen 2005;
39 Noor and Feder 2006; Storz 2005; Vasemagi and Primmer 2005; Hahn 2006; Hedrick 2006; Bonin
40 et al. 2007; Foll and Gaggioti 2008; Riebler et al. 2008; Stinchcombe and Hoekstra 2008). We focus
41 on divergence during the process of population differentiation and speciation, and note that during
42 this process loci under divergent selection and loci causing reproductive isolation behave similarly,
43 differentiating more strongly (even during allopatric divergence), and introgressing less freely than
44 other loci (Barton 1979, 1983; Barton and Hewitt 1989; Mallet 1995, 2005, 2006; Wu 2001; Wu
45 and Ting 2004; Nosil et al. 2005). While acknowledging this similarity (see online supplementary
46 materials for further discussion), we focus on divergent selection per se. Our frequent use of

1 “selection” is shorthand for “divergent selection”, while “genetic differentiation” refers to
2 “molecular genetic differentiation”.

3 In the order presented, the specific aims of this paper are to: (1) discuss theory and make
4 explicit predictions about divergent selection and heterogeneous genomic divergence, aided by the
5 metaphor of genomic islands of divergence, (2) review empirical studies testing for loci whose
6 genetic divergence exceed neutral expectations, that is, ‘outlier loci’, which putatively represent the
7 genetic signature of divergent selection, (3) review empirical studies testing whether adaptive
8 phenotypic divergence facilitates molecular genetic differentiation, (4) describe how selection may
9 promote the growth of genomic islands of divergence, and (5) integrate our findings and offer
10 suggestions for future research.

12 **How selection generates heterogeneous genomic divergence: processes and patterns**

14 We classify loci into three categories according to how they are affected by selection: loci
15 under divergent selection and those tightly linked to them, loci loosely linked to those under
16 selection, and neutral loci that are unlinked to selected loci. This categorization scheme is somewhat
17 arbitrary because the effects of selection vary continuously, but is nonetheless useful because it
18 relates to what can be empirically determined, for example whether an affected locus is an outlier
19 (see Fig. 1 and below).

21 *1) Direct selection and its effect on tightly linked neutral loci (outlier loci)*

23 Divergent selection on a locus will cause its adaptive divergence to a degree that often
24 reflects a balance between the strength of selection and rates of gene flow (Fisher 1930; Wright
25 1931, 1940; Haldane 1930, 1932; Bulmer 1972; Endler 1973; Felsenstein 1976, 1981; Barton 1983;
26 Slatkin 1985; Hendry et al. 2001; Butlin 2005). However, selection on one locus can also strongly
27 affect the frequency of alleles at physically close or ‘tightly linked’ loci, even when the latter are
28 selectively neutral. This process, whereby the frequency of alleles changes in concert with linked
29 selected ones, has been termed genetic hitchhiking (Maynard Smith and Haigh 1974; Kaplan et al.
30 1989; for reviews see Barton 2000; Andolfatto 2001). In essence, under hitchhiking, the effect of
31 selection on loci affecting fitness spills over to neutral loci. The effect of this process depends on a
32 number of factors, but strongly on the ratio of recombination frequency and selection strength ($r /$
33 s), with stronger selection and lower recombination yielding greater effects on linked loci (Bodmer
34 and Parsons 1962; Felsenstein 1981; Charlesworth et al. 1997; Barton 2000; Andolfatto 2001;
35 Ortiz-Barrientos et al. 2002; Butlin 2005). This concept is illustrated in Figure 2A and also implicit
36 in Figure 2B, which further illustrates how genetic divergence at neutral loci decreases with
37 distance from a selected locus. Hitchhiking can also have stochastic effects; by reducing N_e at loci
38 affected by selection, it can amplify divergence via genetic drift (Robertson 1961; Hill and
39 Robertson 1966; Santiago and Caballero 1998).

41 Thus, loci under selection and those tightly physically-linked to them should exhibit greater
42 differentiation than distantly-linked or unlinked neutral regions (Fig. 2; Lewontin and Krakauer
43 1973; Bowcock et al. 1991; Beaumont and Nichols 1996; Black et al. 2001; Schlötterer 2002;
44 Luikart et al. 2003; Beaumont and Balding 2004; Beaumont 2005; Nielsen 2005; Storz 2005;
45 Vasemagi and Primmer 2005; Nachman 2006; Foll and Gaggioti 2008; Stinchcombe and Hoekstra
46 2008; Riebler et al. 2008). Divergent selection thus results in outlier loci whose genetic divergence
47 exceeds neutral expectations. When testing for outliers, simulations are generally used to determine
the upper level of genetic divergence expected under neutrality, and loci whose genetic

1 differentiation exceeds this neutrality threshold are deemed outliers (Fig. 3A). Genetic
 2 differentiation itself is often quantified using F_{ST} , a fixation index that measures the degree of
 3 population differentiation (Wright 1940; Hudson et al. 1992). A final point is that new mutations
 4 are more likely to diverge between populations if they arise in genomic regions already under
 5 divergent selection (i.e., that already exhibit reduced introgression). Consequently, regions of strong
 6 differentiation are predicted to accumulate in clusters within the genome, rather than being
 7 randomly distributed throughout it (see discussion of genomic island growth) (Rieseberg 2001;
 8 Navarro and Barton 2003; Gavrillets 2004, chapter 6; Kirkpatrick and Barton 2006).

9 Here, we review findings from studies testing for outlier loci in natural populations. Other
 10 methods of evaluating the molecular signatures of selection exist, such as McDonald-Kreitman tests
 11 and candidate gene approaches (Nielsen 2005, Nachman 2006 for reviews O'Malley et al. 2007).
 12 However, because our paper concerns population-level patterns of genomic heterogeneity, we focus
 13 on methods that can evaluate divergent selection in large numbers of gene regions across the
 14 genome. In particular, we discuss 'genome scans', in which many individuals are screened for
 15 variation in dozens or hundreds of molecular markers (Lewontin and Krakauer 1973; Bowcock et
 16 al. 1991; Beaumont and Nichols 1996; Andolfatto 2001; Black et al. 2001; Vitalis et al. 2001, 2003;
 17 Schlötterer 2002; Luikart et al. 2003; Beaumont and Balding 2004; Beaumont 2005; Nielsen 2005;
 18 Noor and Feder 2006; Storz 2005; Vasemagi and Primmer 2005; Hahn 2006; Hedrick 2006; Bonin
 19 et al. 2007; Foll and Gaggioti 2008; Riebler et al. 2008; Stinchcombe and Hoekstra 2008). By
 20 distinguishing outliers from putatively neutrally evolving (non-outlier) loci, genome scans enable
 21 the quantification of regions under selection, the evaluation of their distribution across the genome
 22 (e.g., by evaluating linkage disequilibrium), tests for replicated differentiation across population
 23 comparisons, the comparison of evolutionary patterns at outlier versus neutral loci, etc. Genome
 24 scans therefore allow the investigation of how and why divergence varies across the genome,
 25 outstanding questions in evolutionary genetics (Orr 2005). Genome scans typically employ
 26 anonymous molecular markers whose chromosomal positions are unknown, but may also be
 27 combined with classical genetic approaches such as quantitative trait locus (QTL) mapping (Rogers
 28 et al. 2005; see Fig. 3B).

30 *II) Loosely linked neutral loci: additional effects of hitchhiking*

31
 32 Loci that are too far removed along a chromosome from a selected site for hitchhiking to
 33 elevate them to outlier status may nonetheless be somewhat influenced by selection, because
 34 hitchhiking effects can extend a considerable distance from the selected locus (Charlesworth et al.
 35 1997; Nielsen 2005; Fig. 3B). Such non-outlier loci are referred to as loosely linked (= weakly
 36 linked) and are expected to exhibit moderately elevated differentiation compared to completely
 37 unlinked neutral loci. Selection acting through loosely linked loci thus further contributes to
 38 heterogeneous genomic differentiation (Fig. 1). These effects of loose linkage were characterized by
 39 Charlesworth et al. (1997) using a combination of simulations and analytical theory. These authors
 40 found that strong divergent selection ($s = 0.5$) resulted in elevated differentiation of neutral loci
 41 relative to the case where divergent selection was absent, with such effects extending far along the
 42 chromosome and persisting despite high recombination rates. Under moderate selection ($s = 0.1$),
 43 differentiation was less elevated from neutral expectations and approached neutral expectations
 44 when the distance from the selected site exceeded two Morgans (Fig. 2B). Thus, the effects of
 45 hitchhiking on genetic differentiation are positively related to the strength of selection.

46 These observations yield the prediction that at neutral loci loosely linked to those under
 47 selection, levels of genetic differentiation among population pairs will be positively correlated with

1 degree of adaptive phenotypic divergence (a proxy for the strength of divergent selection) across
 2 population pairs (Fig. 3C). We hereafter refer to such associations as ‘Isolation-by-Adaptation’
 3 (IBA, following Nosil et al. 2008; see also Foll and Gaggiotti 2006; Faubet and Gaggiotti 2008;
 4 Funk et al. in review). IBA is most convincing when associations are demonstrated while
 5 controlling for the effects of geographic distance. IBA is analogous (and complementary) to the
 6 well-known pattern of Isolation-by-Distance (IBD; Wright 1943; Slatkin 1993; Rousset 1997), in
 7 which genetic differentiation increases with geographic distance, rather than adaptive divergence
 8 (see also Hendry and Day 2005 for temporal effects on genetic differentiation).

9 10 *III) Unlinked neutral loci: general barriers to neutral gene flow*

11
12 Divergent selection can also have more indirect yet widespread effects on heterogeneous
 13 genomic divergence, by reducing gene flow between populations (Barton and Bengtsson 1986;
 14 Pialek and Barton 1997; Gavrilets and Cruzan 1998; Gavrilets 2004) and thereby facilitating neutral
 15 divergence across the genome via genetic drift (Fig. 2C). The basic scenario is one in which a
 16 population is subject to continuous immigration. Due to divergent local adaptation, immigrants have
 17 lower fitness than residents, yielding selection against immigrants (Mallet and Barton 1989; Funk
 18 1998; Via et al. 2000; Hendry 2004; Nosil et al. 2005 for review). In this fashion, selection against
 19 incoming, locally-maladapted alleles will act as a ‘general barrier’ to the spread of neutral alleles
 20 between populations (Bengtsson 1985; Barton and Bengtsson 1986; Pialek and Barton 1997;
 21 Gavrilets and Cruzan 1998; Navarro and Barton 2003; Gavrilets and Vose 2005).

22 The effective immigration rate of neutral alleles is slowed even further under assortative
 23 mating (Gavrilets 2004, p. 148; for a theoretical summary see the online supplementary materials).
 24 For example, when populations evolve preferences for their native habitat, the resulting decrease in
 25 between-habitat dispersal may reduce opportunities for between-population mating, yielding habitat
 26 isolation and reducing gene flow between populations. In fact, any reproductive barrier, by
 27 definition, reduces gene flow, thereby potentially facilitating neutral genetic divergence via drift.
 28 Notably, adaptive divergence can promote the evolution of all such barriers, including ‘non-
 29 ecological’ ones such as intrinsic hybrid inviability (Bateson, 1909; Muller 1940, 1942;
 30 Dobzhansky 1936, 1937, 1951; Orr 1995; Orr and Turelli 2001; Gavrilets 2004; Dettman et al.
 31 2007), via by-product models of ecological speciation (Mayr 1947, 1963; Funk 1998; Lu and
 32 Bernatchez 1998; Schluter 2000; Ogden and Thorpe 2002; Rundle and Nosil 2005; Funk et al.
 33 2006; Vines and Schluter 2006). We hereafter refer to this scenario, whereby selection facilitates
 34 drift by reducing gene flow, as the ‘general barriers’ mechanism.

35 The ‘general barriers’ mechanism can generate heterogeneous genomic divergence due the
 36 stochastic nature of drift, which causes different neutral loci within a population to differentiate to
 37 varying degrees. The process also predicts that more adaptively divergent populations will
 38 experience greater gene flow reduction and associated neutral differentiation, yielding the pattern of
 39 IBA. Under the general barriers scenario, IBA is particularly analogous to IBD, as both can affect
 40 neutral loci unlinked to those under selection. An open empirical question is how commonly
 41 adaptive divergence restricts gene flow strongly enough to generate IBA at such loci, given that
 42 even small amounts of gene flow overwhelm the ability of drift to cause neutral differentiation (Fig.
 43 2C; Wright 1931, 1940; Barton and Bengtsson 1986).

44 Another issue is that if time since population divergence is the main predictor of adaptive
 45 and genetic divergence, then the pattern of IBA may arise without ‘general barriers’, simply
 46 because both types of divergence increase with time (Roger 1986; Lande 1992; Whitlock 1999;
 47 Merila and Crnokrak 2001; McKay and Latta 2002). Nonetheless, theory indicates that IBA may

1 commonly be generated by general barriers because gene flow represents the predominant force
2 affecting levels of genetic differentiation, even under low migration rates (Wright 1931, 1943;
3 Beaumont and Nichols 1996; Slatkin 1993; Hartl and Clark 1997; Rousset 1997; Balloux and
4 Lugon-Moulin 2002; Hedrick 2005). From an empirical perspective, recently-developed coalescent-
5 based methods can be used to estimate migration rate (m) separately from divergence time (Nielsen
6 and Wakeley 2001; Hey and Nielsen 2004; Hey 2006; Nosil 2008), thereby allowing IBA to be
7 tested while controlling for time since divergence. Another point is that if IBA arises via selective
8 processes, then absolute values of adaptive divergence (e.g., quantitative trait divergence measured
9 using Q_{ST}) are expected to exceed those of F_{ST} , despite the correlation of these two measures (Roger
10 1986; Lande 1992; Whitlock 1999). Thus, time versus adaptation can potentially be distinguished as
11 alternative causes of IBA.

12 In our review, we focus on the exhibition of IBA by neutral loci that are not tightly-linked to
13 those under selection, because such IBA is an expression of the somewhat counterintuitive effects
14 of selection on non-selected genomic regions (via either loose linkage or general barriers). We
15 discuss factors affecting the prevalence and extent of IBA. For example, under the general barriers
16 scenario, the proportion of loci exhibiting IBA should vary according to factors that affect genetic
17 drift, such as levels of gene flow and effective population size. Another prediction is that if IBA is
18 detected in the absence of gene flow (e.g., at spatial scales beyond that at which gene flow occurs),
19 then it is more likely to have arisen via loose linkage than via general barriers.

20 21 *Genomic islands of divergence: an integrated view of genomic heterogeneity*

22
23 To help understand how the above processes combine to generate heterogeneous genomic
24 divergence, we use the concept of ‘genomic islands of divergence’ (Wu 2001; Turner et al. 2005;
25 Harr 2006; Turner and Hahn 2007). Here, we define a ‘genomic island’ as any gene region, be it a
26 single nucleotide or an entire chromosome, that exhibits significantly greater differentiation than
27 expected under neutrality. We conceptually extend the metaphor of genomic islands of divergence
28 by comparing the heterogeneous genetic differentiation observed along a chromosome to the
29 topography of oceanic islands and the contiguous sea floor to which they are connected. The
30 concept is visually depicted in Figure 4.

31 Following this metaphor, sea level represents the threshold above which observed
32 differentiation is significantly greater than expected by neutral evolution alone. Thus, an island is
33 composed of loci – both selected (dark grey) and tightly linked neutral (white) loci – that should be
34 identifiable as outliers in a genome scan. Island elevation (differentiation) is a function of selection
35 strength. Selection also contributes to island size, representing the length of contiguous highly
36 differentiated chromosome. Loosely linked (light grey) loci are depicted as regions far enough from
37 selected loci to fall below sea level as non-outliers, but still close to the surface, being more highly
38 differentiated than most unlinked neutral loci. The differentiation of such loci will also be a function
39 of selection strength, yielding IBA. Farther still from the selected locus, the sea floor drops more
40 steeply, with variation in its topography representing the variable levels of baseline neutral
41 differentiation occurring by genetic drift alone. Thus, the sea floor may be rather homogeneous
42 during divergence with gene flow, where gene exchange between populations homogenizes most
43 neutral differentiation. However, if selection promotes ‘general barriers’ to gene flow during such
44 non-allopatric divergence, genome-wide differentiation under drift is facilitated. This is illustrated
45 by a shallow sea with undersea mounts whose proximity to sea level will vary as a function of
46 selection strength (light grey caps depict loci whose differentiation was facilitated by general
47 barriers such that they contribute to IBA). In contrast, the sea floor is expected to be more

1 heterogeneous during allopatric divergence, where the absence of gene flow allows the stochasticity
 2 of divergence due to drift and variable mutation rates to be more fully expressed (Balloux and
 3 Lugon-Moulin 2002; Hedrick 2005). Notably, the islands themselves may grow through time,
 4 because new mutations are more likely to differentiate if they arise in or adjacent to regions already
 5 subject to divergent selection, thus contributing to island growth (Rieseberg 2001; Navarro and
 6 Barton 2003; Gavrilets 2004, chapter 6; Kirkpatrick and Barton 2006).

7 8 **Literature reviews**

9
10 We conducted two literature reviews to evaluate critical issues relating to genomic
 11 heterogeneity. First, we review genome scan studies seeking to identify outlier loci exhibiting
 12 elevated divergence (Table 2). Second, we review studies that allow the evaluation of IBA (Table
 13 3).

14 15 **Literature review 1: genome scans and outlier loci**

16
17 Numerous methodological issues arise when interpreting genome scans, including aspects of
 18 experimental design and data collection. Major issues include the relationship between type I error
 19 and outlier detection (reflecting the multiple comparisons involved in screening large numbers of
 20 loci), disentangling the effects of mutation rate variation, demography, and selection on levels of
 21 genetic divergence, and determining the ecological causes of the divergent selection. We report on
 22 the methodological robustness of each study described here in the online supplementary materials,
 23 and refer readers to the previous reviews cited in the Introduction for more detailed treatment of
 24 methodological issues.

25 26 *Identifying relevant genome scan studies*

27
28 Relevant papers were identified from a Web of Science (WOS) search on “genome scan and
 29 selection”, and from the studies cited by these papers. In order to focus on genome-wide patterns in
 30 natural populations, our review excluded studies that were genomically restricted (e.g., those
 31 evaluating a single chromosome or assaying markers known *a priori* to be linked to QTL), treated
 32 humans or domesticated species, or did not use divergence-based methods (e.g., those evaluating
 33 selective sweeps within populations), although we cite such studies where appropriate. This
 34 approach located 20 relevant studies (Table 2; a few QTL studies data are treated separately), a
 35 modest number that demonstrates the nascent state of this field. Nonetheless, these studies
 36 illustrated a number of emerging patterns, which are discussed below.

37 38 *Proportion of genome exhibiting outlier behavior*

39
40 We found substantial, but not extreme, variation in the proportion of loci identified as
 41 outliers across studies. This observation holds for studies considering all reported outlier loci
 42 (range: 0.4%-24.5%, mean: 8.5%, $n = 18$), those that only tallied outliers observed in multiple
 43 population pairs (i.e., studies examining ‘replicated divergence’: range: 1.4%-12.0%, mean: 5.8%, n
 44 = 7), and studies examining the distribution of outliers within and between ecological types (range:
 45 0.4%-9.0%, mean: 4.5%, $n = 5$). Note that these percentages quantify the proportion of loci that
 46 were outliers in at least one (or, in the case of replicated outliers, at least two) of a study’s pairwise
 47 population comparisons, rather than mean proportions across individual comparisons. These

1 estimates should be interpreted with great caution, as different studies varied in the number of
 2 populations and individuals examined, molecular markers employed, methods for estimating
 3 baseline neutral differentiation and criteria determining outlier status (Table 2 and online
 4 supplementary materials for details). Nonetheless, the results imply that roughly 5–10% of the
 5 genome is strongly affected by divergent selection. More standardization in the reporting of future
 6 results should help to refine the estimates reported here. What this survey most clearly demonstrates
 7 is that non-trivial proportions of the genomes of disparate taxa show evidence of divergent
 8 selection. These findings are highly consistent with studies examining a few loci or particular parts
 9 of the genome (e.g. sex chromosomes), and those comparing divergence at coding versus non-
 10 coding regions, which also regularly report heterogeneous genomic divergence (Wang et al. 1997;
 11 Machado et al. 2002; Broughton and Harrison 2003; Hoekstra et al. 2004; Payseur et al. 2004;
 12 Llopart et al. 2005; Basset et al. 2006; Geraldès et al. 2006; Zayed and Whitfield 2008).

13 14 *Genomic distribution of outliers*

15
16 To what extent are outlier loci clustered within the genome, as predicted by some theory
 17 (Gavrilets 2004, chapter 6; Kirkpatrick and Barton 2006)? The data in Table 2 suggest a range of
 18 answers, with genomic clustering of outliers ranging from low to reasonably high. Three studies
 19 located outliers on a linkage map. Two of these mapped them to many different linkage groups
 20 (13,14 in Table 2), while the third observed a clustering of loci associated with between-race
 21 divergence in particular regions on just a few of more than twenty chromosomes examined (6 in
 22 Table 2). Two other studies also identified the specific location of genomic differentiation. Turner
 23 et al. (2005) found that differentiation between forms of *Anopheles gambiae* mosquitoes that
 24 involved 1.2% of the genome was clustered into three genomic regions comprising less than 2Mb.
 25 Clustering was somewhat weaker between house mouse (*Mus musculus*) subspecies, where
 26 differentiated regions (7.5% of the autosomal genome) were distributed among eight genomic
 27 regions from the 14-16 chromosomes examined (Harr 2006).

28 Four additional studies indirectly examined outlier distribution by quantifying within-
 29 population linkage disequilibrium (LD) among them (1,8,9,17 in Table 2). If outlier loci are
 30 physically linked, their LD should be elevated relative to neutral loci (Kim and Nielsen 2004),
 31 assuming the latter are widely distributed throughout the genome. Importantly, physical linkage of
 32 outliers is expected to be associated with similar levels of LD within allopatric and parapatric
 33 populations (Arnold 1992). In contrast, if LD forms between physically unlinked loci, due to
 34 migration between differentiated populations (Nei and Li 1973; Kirkpatrick et al. 2002), then LD
 35 will be greater where migration is higher, such as within parapatric populations (Nosil et al. 2006)
 36 or at the center of a cline (Grahame et al. 2006). For all four relevant studies, levels of LD for
 37 outlier loci were very low and, except for one study, similar to those of non-outlier (neutral) loci.
 38 One study did find slightly elevated LD for outlier loci relative to non-outliers (Nosil et al. 2008).
 39 Because levels of within-population LD among outliers were independent from levels of between-
 40 population gene flow (migration), this study provides indirect evidence for weak physical linkage
 41 among outlier loci.

42 In summary, the reviewed studies provided evidence for both genomic dispersion and
 43 genomic clustering of outliers. The former observation suggests that islands of genomic divergence
 44 may be greater in number, and perhaps smaller in size, than currently thought (see also Nakazato et
 45 al. 2007; Turner et al. 2008; Wood et al. 2008) and is contrary to both the theoretical predictions
 46 outlined above and some empirical observations, such as quantitative genetic evidence on the
 47 tendency of genes involved in host adaptation to be sex-linked in herbivorous insects (Prowell

1 1998). More studies are required before explanations for these variable empirical outcomes can be
2 evaluated, although the extent of genomic coverage (e.g., marker density) in a genome scan could
3 contribute to this variability. We suggest that future studies report patterns of LD across different
4 classes of loci and geographic contexts.

5 6 *Comparative insights: 'ecological' genome scans*

7
8 The evaluation of replicated comparisons of particular types of population pairs represents a
9 powerful application of the genome scan approach. This is because replicated divergence across
10 multiple population pairs of a given type is unlikely to arise via non-selective factors such as type I
11 error, genetic drift, or mutation rate variation (Luikart et al. 2003; Campbell and Bernatchez 2004;
12 Bonin et al. 2006, 2007). Such studies might further allow the specific ecological causes of outlier
13 behavior to be identified, for example by contrasting genetic differentiation for population pairs that
14 are ecologically similar versus those that are ecologically (and presumably adaptively) divergent in
15 a specific ecological variable (Nagel and Schluter 1995; Funk 1998; Funk et al. 2002; Nosil 2007).
16 The few studies making comparisons between population pairs with different versus similar
17 ecologies (studies 6-12 in Table 2) suggest that a relatively large proportion (25-100%) of outliers
18 are associated with divergence in a specific ecological variable. A related approach examines
19 correlations between outliers and environmental factors (e.g., studies 5 and 18 in Table 2).

20 Consider the specific example of adaptation to different plant species by herbivorous insect
21 populations. Recent studies have identified loci that are outliers in comparisons of multiple
22 'different-host' population pairs, but that are never outliers in comparisons of different populations
23 that use the same host plant (Egan et al. 2008, Nosil et al. 2008). Such patterns suggest that these
24 loci have likely evolved under the influence of host-plant-related sources of selection. By contrast,
25 loci that are outliers only in 'same-host' population pairs are best interpreted as being affected by
26 host-independent sources of selection (e.g., climate). For example, in a study of nine different-host
27 and six same-host population comparisons, Egan et al. (2008) identified 23 outlier loci (representing
28 5% of all loci examined) associated exclusively with different-host population pairs of
29 *Neochlamisus bebbianae* leaf beetles, and only five outliers associated exclusively with same-host
30 population pairs. From this they concluded that host-related-selection plays a major role in the
31 adaptive genomic differentiation of these populations. This same study identified three loci that
32 were especially highly differentiated outliers in all nine different-host population comparisons and
33 none of the same-host comparisons, thus illustrating how comparative genome scans can identify
34 robust candidate loci (genomic regions) for further molecular characterization and evolutionary
35 study (e.g., Wood et al. 2008).

36 37 *Replicated divergence: adaptation via the same or different mutations?*

38
39 A major question in evolutionary genetics is the extent to which divergent adaptation in
40 different geographic localities or taxa involves the same versus different genes, alleles, and
41 substitutions (Orr 2005). The frequency with which outliers are replicated across different
42 population pairs in a given study provides some insight into this question. In the relevant studies, a
43 large proportion of outlier loci (roughly 50%) were replicated in this fashion, exhibiting outlier
44 status across multiple population pairs (Table 2). This tendency implies that divergent adaptation
45 may often proceed in two ways. First, it may occur by the repeated and parallel fixation of the same
46 allele in different populations (Table 2; see also Storz and Nachman 2003; Panova et al. 2006;
47 Turner et al. 2008). Second, an adaptive allele may arise and be locally fixed only once, followed

1 by the subsequent spread of the ecological type carrying the new allele to multiple geographic
2 localities.

3 Replicated outlier behavior is of particular interest when divergent adaptation in different
4 localities can be demonstrated to have evolved independently (e.g., multiple origins of ecological
5 types), thus providing evidence for truly ‘parallel’ divergence (e.g., as in *Coregonus* whitefish,
6 Campbell and Bernatchez 2004, and *Timema* walking-sticks, Nosil et al. 2002, Nosil et al. 2008;
7 Table 2). A definitive example of this process comes from threespine sticklebacks (*Gasterosteus*
8 *aculeatus*), in which independent instances of adaptation to freshwater environments have
9 repeatedly involved the loss of lateral plates (an anti-predator trait) via the fixation of the same
10 allele at the ectodysplasin gene (Colosimo et al. 2005). Alternatively, divergent adaptation may
11 proceed via different mutations/loci in different localities such that particular outliers are not highly
12 consistently observed across population comparisons (e.g., Panova et al. 2006; Acheré et al. 2005).
13 These different mutations could either be recently derived or reflect the differential sorting of
14 standing genetic variation (Barrett and Schluter 2008).

15 16 *Genealogical discordance: population trees from neutral versus selected loci*

17
18 When genetic exchange between populations varies among loci, different loci can yield
19 different gene trees (i.e., genealogical discordance) (Maddison 1997; Hey 2006; Shaw 2002; Funk
20 and Omland 2003). Loci involved in divergent adaptation and reproductive isolation are expected to
21 reflect boundaries between biological species or ecotypes more strongly than neutral loci, because
22 the former flow less readily between populations (Wu 2001; Dopman et al. 2005; Hey 2006; Xie et
23 al. 2007). This process generates the prediction that selected loci are likely to group populations by
24 adaptively relevant ecological variables whereas neutral loci are likely to phylogenetically group
25 populations by geographic proximity (reflecting spatial patterns of gene flow). A number of studies
26 that have evaluated genetic structure at each of a few loci report such patterns (Beltrán et al. 2002;
27 Hoekstra et al. 2004; Dopman et al. 2005; Llopart et al. 2005; Bull et al. 2006; Cano et al. 2006;
28 Geraldès et al. 2006; Putnam et al. 2007; Roe and Sperling 2007). Five genome scan studies
29 (1,2,4,6,9 in Table 2) have analogously compared population trees for datasets that either include or
30 exclude outliers. All these studies observed that outlier loci group populations more as a function of
31 ecology whereas putatively neutral (non-outlier) loci group populations in a manner more consistent
32 with geography. The extent to which outlier loci group populations according to ecology should
33 increase as the frequency of outlier loci and population pairs exhibiting replicated divergence
34 increases, and these studies are consistent with this prediction. Another factor is how tightly linked
35 the outliers are to the actual targets of selection.

36 37 *Genome scans incorporating QTL*

38
39 We conclude this section by considering studies that complement genome scans with QTL
40 approaches. A major goal of such studies is to determine the phenotypic traits and associated
41 genomic regions that contribute to adaptive population divergence. This can be achieved by
42 determining which (if any) QTL from mapping experiments are also outlier loci in genome scans.
43 Such work is relevant to heterogeneous genomic divergence because it compares levels of genetic
44 divergence between QTL and other markers. Finding QTL that are also outlier loci puts the study of
45 heterogeneous genomic divergence on steadier ‘ecological footing’, as it allows stronger inferences
46 about underlying ecological traits and associated divergent selection than genome scans of
47 anonymous markers alone (Stinchcombe and Hoekstra 2008). For treatment of allele frequency

1 expectations at neutral loci versus QTL, we refer readers to past reviews (Latta 1998, 2003; LeLorre
2 and Kremer 2003).

3 To date, few studies have combined QTL and genome scan approaches. Thus, we focus on a
4 few key examples. The first involves whitefish (*Coregonus clupeaformis*) ecotypes studied by
5 Rogers et al. (2005) and Rogers and Bernatchez (2007). Genetic mapping identified nine QTL for
6 phenotypic traits related to adaptation of the two ecotypes to different lake environments, while a
7 genome scan of four sympatric pairs of ecotypes identified 24 outlier loci (among 440 loci
8 examined). These outliers were associated with QTL, as opposed to other genomic regions, more
9 often than expected by chance. These findings increased confidence that regions under selection had
10 indeed been identified, and that outlier differentiation reflected divergent adaptation to lake
11 environments (see also Ólafsdóttir et al. 2006). Although the authors point out that demographic,
12 spatial, and local selective effects can influence QTL-outlier correlations (Beaumont and Balding
13 2004; Hahn 2006), this work clearly illustrates the inferential advantages of a more integrated
14 approach.

15 The second example treats *Helianthus annuus* and *H. petiolaris* sunflowers (Yatabe et al.
16 2007). These species hybridize extensively, F1 hybrids exhibit extremely low pollen fertility, and
17 these fertility barriers map to chromosomal rearrangements (Rieseberg et al. 1999). This study,
18 however, detected no association of outliers with QTL for morphological differences or hybrid
19 sterility, and only weak associations of outliers with chromosomal inversions. The authors conclude
20 that regions of differentiation between these species are very small.

21 A third study considers alfalfa and clover host races of *Acyrtosiphon pisum* pea aphids (Via
22 and West 2008), and reports that outlier loci between the races are significantly clustered around
23 QTL for traits that cause ecologically based reproductive isolation, while also demonstrating that
24 genetic divergence decreases with increasing map distance from QTL (Fig. 3B). A final example
25 (Mäkinen et al. 2008a,b) examined 103 microsatellites (many linked to known QTL) in four
26 freshwater and three marine populations of threespine stickleback fishes. This study reported 2.8%
27 of loci to be outliers, with the clearest signature of selection exhibited by a marker associated with
28 the ectodysplasin gene (which codes for the ecologically important trait of lateral plate number).
29 However, other markers associated with QTL showed no signature of selection, and two strong
30 outliers were not associated with known QTL. These findings highlight certain limitations of this
31 otherwise informative approach. Because QTL studies examine only a subset of phenotypic traits
32 potentially evolving under divergent selection, outlier loci truly subject to such selection may
33 nonetheless map to non-QTL regions (Martin et al. 2005). Moreover, QTL regions are often very
34 large (tens of cM) so that in the absence of extensive LD along the chromosome, genome scans
35 might not identify loci within these regions as outliers (Cano et al. 2006).

36 37 **Literature review 2: isolation-by-adaptation**

38 39 *Identifying relevant IBA studies*

40
41 We identified 22 studies relevant to the evaluation of IBA (controlling for geographic
42 distance) from a WOS search on “ecology and genetic and divergence”, and from the studies cited
43 by the recovered papers. In these studies, adaptive divergence was inferred using either the degree
44 of divergence in habitat (n = 15) or phenotype (n = 7). We excluded phylogeographic studies that
45 examine genetic divergence in relation to habitat, because they generally do not analyze genetic
46 divergence among population pairs in relation to adaptive divergence. Similarly, studies of mosaic
47 hybrid zones relate genetic differentiation to habitat type, but generally do not explicitly evaluate

1 genetic divergence as a function of ecology versus geographic distance, and have been reviewed
 2 elsewhere (Nosil et al. 2005). One study of each of these latter two types was nonetheless included
 3 for illustrative purposes (but excluded when calculating summary statistics). Due to the broad range
 4 of studies that could conceivably be suitable for examining IBA, we acknowledge that additional
 5 relevant studies surely exist. This contrasts with the genome scan review above, where only a more
 6 specific type of study was relevant, such that the majority of relevant studies were likely identified.
 7 Nonetheless, we have hopefully identified a representative sample of studies, and investigations
 8 using distance matrices are likely well covered. We classified the focal studies into those that
 9 examined matrices of pairwise differences among population pairs, for example, using Mantel tests
 10 (Manly 1997), versus those that quantified the proportion of genetic variation explained by habitat
 11 within an AMOVA framework (Excoffier et al. 1992). The studies treated here are assumed to
 12 evaluate loci that are neutral and not tightly-linked to those under selection. Such loci represent
 13 perhaps the most intriguing and widely-evaluated aspect of IBA.

14 15 *Prevalence of IBA*

16
17 Our standard for the presence of IBA was either a significant positive correlation between
 18 adaptive and neutral genetic divergence across population pairs or significant genetic structuring
 19 between ecological types or habitats by AMOVA. With this in mind, we found evidence for IBA in
 20 15/22 studies across a variety of taxa (Table 3). Of the 16 tests employing distance matrices, 75%
 21 support IBA, as do half of the AMOVA studies. As our survey is not a formal meta-analysis, the
 22 results should not be over-interpreted. The clearest result is that IBA is not uncommon in nature.

23 An unresolved question is what proportion of the genome exhibits IBA. Most studies testing
 24 for IBA pool across loci to obtain population-level estimates of genetic differentiation, precluding
 25 an estimate of this value. However, four studies did report locus-specific results. In whitefish
 26 ecotypes (*Coregonus clupeaformis*), five of six microsatellite loci exhibited fairly strong IBA,
 27 showing correlations between ecomorphological and neutral genetic divergence that ranged from r
 28 = 0.72 - 0.84 (Lu and Bernatchez 1999). In host plant ecotypes of *Timema cristinae* walking-stick
 29 insects, 10% of non-outlier AFLP loci, as well as mitochondrial DNA, exhibited significant IBA
 30 (Nosil et al. 2008). In host forms of *Neochlamisus bebbianae* leaf beetles, 11% of non-outlier AFLP
 31 loci exhibit IBA, but mtDNA did not (Funk et al., in review). Finally, in wild barley (*Hordeum*
 32 *spontaneum*), 44% of RAPD loci exhibited a strong correlation with soil type (Owuor et al. 1999).
 33 Thus, although a genome-wide signature of IBA is not rare, the proportion of the genome exhibiting
 34 a particularly strong pattern of IBA may vary considerably. The combined trends suggest a
 35 potentially important role for natural selection in neutral genomic differentiation between
 36 populations.

37 38 *The role of ecological versus geographic factors in reducing gene flow*

39
40 What are the relative roles of ecology versus geography in facilitating neutral genetic
 41 divergence? IBD was detected in 50% of studies where it was evaluated ($n = 16$), somewhat less
 42 than for IBA. Perhaps more interesting is that the presence versus absence of IBA did not appear
 43 strongly associated with the presence versus absence of IBD. For example, of the sixteen studies in
 44 Table 3 that tested for both IBA and IBD, six detected both, two detected neither, six detected only
 45 IBA, and two detected only IBD ($p > 0.25$, Fisher's exact test).

46 Consider some examples of these different types of results. Ogden and Thorpe (2002) report
 47 a primary role for ecological divergence, rather than geographic distance, in the neutral genetic

1 divergence of a Caribbean lizard (*Anolis roquet*). Using seven microsatellite loci, genetic distance
2 was compared among pairs of adjacent localities from three different transects, one of which cut
3 through an ecological gradient (the “habitat transect”), and two of which did not. Strong genetic
4 differentiation was observed only in the habitat transect, particularly at habitat boundaries, and
5 genetic structuring by habitat was further supported by AMOVA. The results are thus consistent
6 with IBA. In contrast to these findings, a study of guppies (*Poecilia reticulata*) reported no
7 evidence for IBA and a primary role for geographic distance and physical barriers to dispersal in
8 reducing gene flow (Crispo et al. 2006). Finally, a study of grey wolves (*Canis lupus*) related
9 genetic distance to both habitat type (tundra, taiga, or boreal coniferous forest) and geographic
10 distance, reporting both IBA and IBD (Musiani et al. 2007). Below we propose some hypotheses for
11 this variability.

12

13 *Hypotheses for variability among studies*

14

15 IBA appears relatively common, yet was not uniformly observed. What factors explain
16 variability in IBA? An obvious possibility is that selection is sometimes too weak to cause IBA,
17 whether via linkage or general barriers to gene flow (Figs. 1, 2). Another possibility is that the
18 particular phenotypic/ecological traits evaluated are not good proxies for the major sources of
19 divergent selection acting on study populations. Reciprocal transplant experiments indicate that
20 divergent selection between alternative environments is very common (Schluter 2000; Nosil et al.
21 2005 for reviews), suggesting that habitat-based indices of adaptive divergence should provide
22 informative proxies for selection. Notably, however, in the limited sample of studies available to
23 date, the detection of IBA appears independent from whether adaptive divergence was inferred
24 using habitat or phenotypic data ($p > 0.25$, Fisher’s exact test).

25 Finally, levels of gene flow can affect IBA. For example, high gene flow can overwhelm
26 adaptive divergence (Saint-Laurent et al. 2003; Hendry and Taylor 2004; Smith et al. 2005; Crispo
27 et al. 2006; Yatabe et al. 2007), precluding the generation of IBA. This raises the issue of reversed
28 causality, where even if IBA is detected, levels of gene flow may be affecting the degree of
29 adaptive divergence, rather than vice versa (Hendry et al. 2001; Hendry and Taylor 2004; Nosil and
30 Crespi 2004). In addition to levels of biological gene flow, the spatial scale of sampling is also
31 important. At a spatial scale greater than that at which gene flow occurs, neutral divergence can
32 occur without general barriers to gene exchange, potentially precluding the detection of IBA at this
33 scale (even if IBA would be detectable by sampling at a smaller spatial scale).

34

35 **Studies combining tests for outlier loci with the examination of IBA**

36

37 As described above, divergent selection can cause heterogeneous genomic divergence via
38 different processes. Questions remain concerning how these different effects of selection interact,
39 because outlier loci and IBA have rarely been evaluated within the same system. Here we
40 summarize findings from the few existing examples.

41 One of the first applications of a genome scan to natural populations of a non-model
42 organism was that of Wilding et al. (2001), who identified AFLP loci that were outliers between
43 upper and lower shore ecotypes of *Littorina saxatilis* snails. A subsequent study further examined
44 these outliers, as well as the effects of adaptive divergence on neutral differentiation (Grahame et al.
45 2006). These studies showed outlier loci to form sharp clines in allele frequencies in transects
46 between the upper and lower shore (Fig. 5A). LD among outlier loci was low except in the middle
47 of the cline, as expected when migration between differentiated populations generates LD between

1 unlinked loci (see also Wood et al. 2008). Non-outlier loci exhibited higher F_{ST} between ecotypes
 2 relative to within-ecotype comparisons (independent of geographic distance; Fig. 5B). These results
 3 are consistent with selection acting on unlinked outlier loci and creating a general barrier to gene
 4 flow across a contact zone, resulting in IBA at neutral loci.

5 Another combined analysis was conducted by Ólafsdóttir et al. (2006), who examined both
 6 QTL putatively under selection and neutral markers. These authors reported elevated divergence
 7 between stickleback (*Gasterosteus aculeatus*) ecotypes at microsatellites associated with QTL
 8 relative to putatively neutral non-QTL markers. They also observed IBA at neutral loci as a function
 9 of divergent predation regimes.

10 Finally, studies on two herbivorous insect systems each used AFLPs to identify outliers and
 11 evaluate IBA among populations associated with either of two host plants. In an investigation of
 12 *Timema cristinae* walking-stick insects (Nosil et al. 2008), 8% of AFLP loci proved to be outliers
 13 and 1-2% of loci were identified as likely subjects of host-plant-related selection, using the
 14 comparative approach described above. This study found little evidence for IBA when loci were
 15 pooled (Fig 5C). However, as described earlier, locus-specific analyses revealed that 10% of non-
 16 outlier loci exhibited IBA, as did mitochondrial DNA (Fig 5D). Results from Egan et al. (2008) on
 17 outlier proportions and host-related selection in *Neochlamisus bebbianae* leaf beetles were reviewed
 18 above. Funk et al. (in review) further report that about 11% of non-outlier loci in *N. bebbianae*
 19 exhibited IBA in individual-locus analyses (Fig. 5E), similar to the proportion in *Timema*. However,
 20 in contrast to the *Timema* study, IBA was detected when loci were pooled, with stronger IBA for
 21 host-specific outliers than for non-outlier (neutral) loci (Fig. 5F). In summary, patterns for outlier
 22 loci and IBA vary among systems, with ecologically related systems exhibiting both similarities and
 23 differences.

24 **The growth of genomic islands of divergence: alternative models**

25
 26
 27 Our literature reviews detected some general trends, but also variability in patterns of outlier
 28 behavior and of IBA. This variability may reflect factors determining the size and number of
 29 differentiated regions (i.e., genomic islands of divergence) across the genome. As described, the
 30 effects of divergent selection can extend across a chromosome via different effects on tightly
 31 linked, loosely linked, and unlinked loci. The further the effects of divergent selection extend along
 32 chromosomes, the larger the associated regions of elevated differentiation. The stronger the
 33 selective barriers to gene flow, the greater the number of loci expected to exhibit elevated
 34 differentiation, via processes that contribute to IBA. But how do genomic islands of divergence
 35 (hereafter “islands” for simplicity) grow? Here we present models for island growth (see also Fig. 4,
 36 Table 1, and the online supplementary materials).

37 *Model I) Allopatric model*

38
 39
 40 Genetic divergence and reproductive isolation during allopatric differentiation is unimpeded
 41 by gene flow and increases with time (reviewed in Coyne and Orr 2004). Thus, the allopatric model
 42 predicts that the number of islands should be positively correlated with time since population
 43 divergence. As compared to models of divergence with gene flow (see below), regions of
 44 differentiation are not predicted to be as highly clustered within the genome during allopatric
 45 divergence, because divergence at all regions (i.e., not only those already exhibiting reduced
 46 introgression) can proceed unimpeded by gene flow. An empirical study comparing an allopatric
 47 species pairs of *Drosophila* to a sympatric pair supports this prediction (Brown et al. 2004). This

1 raises the issue of the degree to which low clustering of differentiated regions within the genome
 2 represents a genetic signature of allopatric divergence. Another pattern that may be associated with
 3 this model is high baseline levels of neutral differentiation, because physical barriers to gene flow
 4 make divergence via genetic drift likely (i.e., even without the evolution of ‘general barriers’ to
 5 gene flow). Also, islands might be relatively small owing to the absence of certain conditions
 6 characterizing island growth in the other models. Nonetheless, even in the allopatric case, regions
 7 under stronger selection are expected to produce islands of greater size and height through their
 8 stronger effects on selected and linked neutral regions within each population. And islands may
 9 reach nontrivial height since adaptive divergence is unconstrained by gene flow. Thus, the allopatric
 10 model predicts many small to modest-sized islands, whose number and height are a positive
 11 function of time and selection strength (Yatabe et al. 2007), and which are distributed throughout
 12 the genome, as recently observed between allopatric populations of ferns (Nakazato et al. 2007).

13

14 *Model II) Ecological (divergent selection) model with gene flow*

15

16 The growth of islands in the face of gene flow is likely to differ from their growth in
 17 allopatric conditions, with the effects of time being less clear. Genes under divergent selection, and
 18 those tightly linked to them, will experience reduced introgression relative to neutral, unlinked loci.
 19 The chance that a new mutation (whether adaptive or neutral) will persist and increase in frequency
 20 is highest in regions of reduced introgression (Gavrilets 2004, chapter 6). Therefore, differentiated
 21 loci are expected to accumulate in genomic regions that already harbor genes under divergent
 22 selection, leading to increases in the number of genes within an island and thus in island size, and
 23 potentially generating a positive feedback loop. This model thus predicts that genes affecting local
 24 adaptation will form clusters within the genome rather than being more evenly distributed across it.
 25 QTL studies demonstrating that different adaptive traits map to similar genomic regions, but not
 26 within known inversions, are consistent with this prediction (e.g., *Acyrtosiphon* pea aphids,
 27 Hawthorne and Via 2001; *Heliconius* mimetic butterflies, Kronforst et al. 2006; *Coregonus*
 28 whitefish ecotypes, Rogers and Bernatchez 2007; *Gasterosteus* sticklebacks, Albert et al. 2008),
 29 although pleiotropy could also contribute to these results.

30 A number of issues will affect the generality of the process outlined above. A major one is
 31 whether divergent selection typically persists long enough for new mutations to arise and be
 32 captured by regions of reduced introgression. Empirical evidence suggests that anciently diverged,
 33 sympatric species that have presumably been subject to divergent selection for extended periods of
 34 time sometimes exhibit very small regions of differentiation (Mallet et al. 2007; Yatabe et al. 2007),
 35 but other times larger ones (Llopart et al. 2005). Why do islands sometimes appear to grow, and
 36 other times not? Some of this variability may be due to how many different islands (i.e., gene
 37 regions) affect a given trait under divergent selection, with more islands perhaps associated with
 38 more opportunity to capture new mutations within a least one island, thereby resulting in island
 39 growth. A counterargument is that more genes (i.e., islands) affecting a trait can result in weaker
 40 per-locus selection coefficients (Gavrilets and Vose 2005), thereby constraining divergence with
 41 gene flow.

42 Another issue is the capacity of natural selection to favor the evolution of tighter linkage
 43 among loci, for example to keep beneficial genotypic combinations together (Kimura 1956; Bodmer
 44 and Parsons 1962; Kojima and Schaffer 1964; Butlin 2005). This process could proceed via the
 45 evolution of modifier loci that suppress recombination (Kouyus et al. 2006) and facilitate the
 46 growth of genomic islands. Support for the evolution of tighter linkage is provided by the evolution
 47 of ‘supergenes’, that is, groups of neighboring genes on a chromosome that are inherited together.

1 Different genes within a supergene tend to affect different, if sometimes related, traits (Nabour et al.
 2 1933; Sheppard 1953; Clarke and Sheppard 1960; Turner 1967a; Sinervo and Svensson 2002).
 3 Classic examples are genes affecting different color-pattern traits in insects, such as instances where
 4 the fitness of an allele at a locus affecting one color-pattern element depends on which alleles are
 5 present at a different locus affecting another color-pattern element (Nabour et al. 1933; Sheppard
 6 1953). Such a scenario generates selection favoring the retention of particular genotypic
 7 combinations (i.e., tighter linkage; Nabour et al. 1933; Sheppard 1953; Clarke and Sheppard 1960;
 8 Turner 1967a; Sinervo and Svensson 2002), but can increase linkage only within chromosomes
 9 (Turner 1967b; Charlesworth and Charlesworth 1975; Joron et al. 2006). In sum, compared to the
 10 allopatric model, the ecological model predicts greater genomic clustering of genomic regions
 11 under selection and fewer islands. The ecological model further predicts that islands can be small,
 12 but will sometimes be large and include multiple and sometimes interacting selected genes.

13
 14 *Model III) Structural model with gene flow*

15
 16 The origin and growth of sizeable islands might be further facilitated by the structural
 17 organization of the genome, notably by chromosomal inversions (Noor et al. 2001; Rieseberg 2001;
 18 Ortiz-Barrientos et al. 2002; Butlin 2005; Machado et al. 2007; Noor et al. 2007). Inversions
 19 themselves might cause postmating isolation, but could further promote differentiation by
 20 facilitating adaptive divergence via various processes. The ‘protection from introgression’
 21 hypothesis proposes that inversions are initially established by some unknown mechanism, perhaps
 22 in allopatry. When the inversion-bearing populations come into contact, the inversions promote
 23 adaptive divergence (and associated genomic divergence) by reducing introgression at large regions
 24 of the genome and protecting favorable genotypic combinations that arise within these regions from
 25 being broken up by recombination (Rieseberg 2001; Noor et al. 2001; Brown et al. 2004). The
 26 related ‘inversions first’ scenario posits that once inversions are established, genetic differences
 27 between taxa can easily build up within them (Navarro and Barton 2003). Finally, the recent
 28 ‘selective spread’ hypothesis posits that a newly formed inversion captures locally adapted alleles at
 29 two or more loci in hybridizing populations (Kirkpatrick and Barton 2006; Manoukis et al. 2008).
 30 These co-occurring loci confer a fitness advantage to the inversion by keeping well-adapted
 31 genotypes intact, thereby allowing the inversion to spread via selection. This scenario differs from
 32 the others in positing that locally adapted alleles within an inversion are the cause of the inversion’s
 33 spread, rather than a consequence of its existence.

34 The extent to which inversions promote island growth under these scenarios depends on two
 35 main factors. The first is how frequently genes under selection or promoting reproductive isolation
 36 occur within inversions. The second factor is the extent to which inversions reduce introgression at
 37 regions outside of them. Recent work reported that levels of genetic differentiation between
 38 *Drosophila* species were indeed somewhat elevated just outside of inversions, but dropped off
 39 markedly even just a few megabases outside the inversion (Machado et al. 2007; Noor et al. 2007).

40 Despite its simplicity and potentially great explanatory capacity, empirical evidence on the
 41 importance of the structural model appears mixed. On one hand, examples exist where inversions
 42 appear to have promoted genomic divergence and speciation (Rieseberg et al. 1999; Rieseberg
 43 2001; Noor et al. 2001; Feder et al. 2003a; Basset et al. 2006; Butlin 2005; Machado et al. 2007;
 44 Noor et al. 2007; Yatabe et al. 2007; Manoukis et al. 2008). On the other hand, our genome scan
 45 review found that islands were often genomically scattered rather than clustered, contrary to
 46 expectation if such regions tend to reside within inversions (Table 1). Moreover, of the two studies
 47 explicitly focusing on ‘islands of divergence’ (Turner et al. 2005; Harr 2006), only the first found

1 regions of high differentiation to be associated with regions of reduced recombination (specifically,
2 centromeres, Fig. 5B). Thus, it appears that although inversions can promote island formation and
3 growth, they are not required.

4 The structural and ecological models are similar in some aspects, but differ most clearly in
5 two respects. First, a smaller number and greater size of genomic islands is predicted for the
6 structural model because inversions are often not plentiful within genomes, but commonly consume
7 extensive regions of chromosomes. Second, the structural model may allow longer evolutionary
8 persistence of islands, due to the strong inhibition of recombination caused by inversions.

9 10 *Integration of the Different Models*

11
12 The models above are not mutually exclusive, and might interact. For example, the
13 ecological and structural models could operate simultaneously, resulting in clustering of genes
14 affecting local adaptation within (and around) chromosomal inversions, as predicted by recent
15 theory (Kirkpatrick and Barton 2006). Another possible interaction is between the allopatric model
16 and models involving gene flow, given that the often extended process of speciation can involve
17 multiple geographic modes. For example, some divergence might occur in allopatry and some in
18 sympatry (Feder et al. 2003a; Rundle and Schluter 2004; Rundle and Nosil 2005; Xie et al. 2007).
19 Such a geographically pluralistic view of speciation has implications for the heterogeneity of
20 genomic divergence observed. For example, the amount of adaptive divergence and reproductive
21 isolation that evolves during an initial allopatric period will influence patterns of genomic
22 differentiation following secondary contact. If very little reproductive isolation evolved, then
23 widespread gene flow will ultimately erode accumulated differentiation at islands not possessing the
24 introgression-resisting characteristics described above. If strong reproductive isolation evolved,
25 then accumulated differentiation might largely be maintained, and further divergence can occur.
26 Thus, the number of genomic islands should be positively associated with the degree of divergence
27 during the allopatric period, itself a function, in part, of the duration of this period. A related point is
28 that the age of a contact (e.g., hybrid) zone will affect patterns of heterogeneous genomic
29 divergence, because it takes time for gene flow to erode divergence upon secondary contact
30 (Strasburg and Rieseberg 2008). Thus, the size and number of islands might decrease through time
31 since secondary contact.

32 33 *Factors affecting all models*

34
35 A number of additional factors are relevant to island growth under all models. One is the
36 genomic distribution of genes subject to divergent selection, as island formation and growth will be
37 facilitated by any tendency for selected genes to occur in physical proximity. Another is the
38 distribution of linkage disequilibrium (LD) and its rate of decay with increasing distance from
39 selected regions (Charlesworth et al. 1997; Nielsen 2005; Cano et al. 2006). The more rapidly LD
40 decays, the smaller the resulting islands. The relevance of this issue is indicated by evidence that
41 levels of LD can vary among genomic regions and taxa, for example due to variation in
42 recombination rates (Reich et al. 2001; Scotti-Saintagne et al. 2004; Liu and Burke 2006;
43 Arunyawat et al. 2007). In wild sunflower (*Helianthus annuus*), for instance, LD among loci falls to
44 negligible levels within 200bp, whereas in cultivated varieties of this species it extends up to
45 1100bp (Liu and Burke 2006). Another factor is whether forms of reproductive isolation (e.g.,
46 assortative mating) evolve, with genomic differentiation via selection being facilitated when such
47 barriers to gene flow evolve. Reproductive isolation itself can involve epistatic interactions between

1 different loci, perhaps even loci on different chromosomes. The effects of such epistasis on the
2 nature of islands are poorly understood, illustrating a need for work on ‘speciation’ islands.

3 A factor of general importance, as noted in our initial description of islands, is the strength
4 of selection, with stronger selection leading to stronger hitchhiking effects (Fig. 1A), and thus
5 presumably to larger and higher islands. Thus, islands may be expected to grow through time as
6 populations in alternative environments gradually become more divergently adapted, resulting in
7 stronger selection against immigrants. Relatedly, patterns of island growth might vary across
8 different stages of the speciation process (for consideration of the speciation continuum see Wu
9 2001; Berlocher and Feder 2002; Dres and Mallet 2002; Hey et al. 2003; Coyne and Orr 2004; De
10 Queiroz 2005; Funk et al. 2006; Rueffler et al. 2006; Mallet et al. 2007; Svensson et al. 2007; Nosil
11 and Sandoval 2008). Many of the examples for our genome scan review (Table 2) compare
12 conspecific populations, rather than distinct species. Thus, it is possible that our findings reflect a
13 bias towards the earlier stages of speciation, when islands may still be small, reflecting the few
14 regions under strong selection and low levels of associated reproductive isolation. The later stages
15 of speciation might be characterized by different types of divergence, for example larger islands
16 than contain inversions that facilitate the long-term persistence of differentiation. Studies of
17 population/species pairs spanning the range of divergences that cumulatively represent the
18 speciation continuum might contribute to a more comprehensive understanding of how divergent
19 selection affects heterogeneous genomic divergence.

20 21 **Conclusions and future directions**

22
23 Divergent selection is predicted to yield heterogeneous divergence across the genome. And
24 indeed, recent work is beginning to document variation in the prevalence, distribution, and size of
25 differentiated genomic regions. A variety of outstanding questions exist about the causes and
26 consequences of these patterns. Here, we have attempted to integrate emerging ideas and findings
27 from different research traditions in order to facilitate future exploration of the role of divergent
28 selection in the evolution of heterogeneous genomic divergence, including the testing of associated
29 predictions (Table 1). A major conclusion is that divergent selection plays multiple roles. These
30 include its effects on fitness-associated loci, its effects on both tightly and loosely linked neutral
31 loci via hitchhiking, its facilitation of genetic drift by countering gene flow, and its influence on the
32 growth of genomic islands. These sometimes non-intuitive contributions illustrate the capacity of
33 selection to affect many aspects of genomic differentiation. More specifically, the modest data
34 collected so far illustrate considerable variability in the number, size, and genomic distribution of
35 strongly differentiated regions. Avenues for future research should include, first and foremost, the
36 collection of more data from more taxa. This is clearly required for generalities to more fully
37 emerge and to distinguish between alternative hypotheses. Also important will be evaluating how
38 particular ecological, genetic, and geographic factors help explain observed variability. Finally,
39 increased integration of different methods and the incorporation of additional methods (e.g., gene
40 expression analysis, Roberge et al. 2005; Derome et al. 2006) will allow new questions to be
41 addressed. We hope that the ideas and data reviewed here help promote the advancement of the
42 emerging field of population genomics.

43 44 **Acknowledgements**

45 We thank B. Crespi, A. Hendry, J. Galindo, J. Lee-Yaw, H. Collin, S. Rogers, and D.
46 Schluter for discussions about genomic divergence and speciation. S. Rogers, S.P. Egan, J. Galindo,
47 L. Bernatchez, and three anonymous reviewers provided many useful comments on earlier versions

1 of this manuscript. During the preparation of this manuscript, P.N. was funded by a post-doctoral
 2 fellowship from the Natural Sciences and Engineering Research Council (NSERC) Canada, D.J.F.
 3 was funded by Vanderbilt University and by NSF grants IOB0447659 and DEB0640741, and
 4 D.O.B. was funded by The University of Queensland.

6 Literature Cited

- 7 Achere V, Favre JM, Besnard G, Jeandroz, S (2005) Genomic organization of molecular
 8 differentiation in Norway spruce (*Picea abies*). *Molecular Ecology*, **14**, 3191-3201.
- 9 Albert AYK, Sawaya S, Vines TH, Knecht AK, Miller CT, Summers BR, Balabhadra S,
 10 Kingsley DM, Schluter D (2007). The genetics of adaptive shape shift in stickleback:
 11 pleiotropy and effect size. *Evolution*, **62**, 76-85.
- 12 Andolfatto P (2001) Adaptive hitchhiking effects on genome variability. *Current Opinion in*
 13 *Genetics and Development*, **11**, 635-641.
- 14 Arnqvist G, Rowe, L (2005) *Sexual conflict*. Princeton University Press, Princeton.
- 15 Arunyawat U, Stephan W, Stadler T (2007) Using multilocus sequence data to assess
 16 population structure, natural selection, and linkage disequilibrium in wild tomatoes.
 17 *Molecular Biology and Evolution*, **24**, 2310 - 2322.
- 18 Avise, JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University
 19 Press, Cambridge, MA.
- 20 Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with
 21 microsatellite markers. *Molecular Ecology*, **11**, 155-165.
- 22 Basset P, Yannic G, Brunner H, Hausser J (2006) Restricted gene flow at specific parts of the
 23 shrew genome in chromosomal hybrid zones. *Evolution*, **60**, 1718-1730.
- 24 Barrett R, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology*
 25 *and Evolution*, **23**, 38-44.
- 26 Barton NH (1979) The dynamics of hybrid zones. *Heredity* **43**, 341-359.
- 27 Barton, NH (1983) Multilocus clines. *Evolution*, **37**, 454-471.
- 28 Barton, NH (2000) Genetic hitchhiking. *Philosophical transactions of the Royal Society of*
 29 *London B*, **355**, 1553-1562.
- 30 Barton N, Bengtsson BO (1986) The barrier to genetic exchange between hybridizing
 31 populations. *Heredity*, **57**, 357-376.
- 32 Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature* **341**, 497-503.
- 33 Bateson W (1909) Heredity and variation in modern lights. In *Darwin and modern science*, ed.
 34 A.C. Seward, pp. 85-101. Cambridge University Press, Cambridge.
- 35 Beaumont MA (2005) Adaptation and speciation: what can F_{st} tell us? *Trends in Ecology*
 36 *and Evolution*, **20**, 435-440.
- 37 Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of
 38 population structure. *Proceedings of the Royal Society of London B*, **263**, 1619-1626.
- 39 Beaumont MA, Balding, DJ (2004) Identifying adaptive genetic divergence among
 40 populations from genome scans. *Molecular Ecology*, **13**, 969-980.
- 41 Begun DJ, Holloway AK, Stevens K, Hillier LW, Poh YP, Hahn MW, Nista PM, Jones CD,
 42 Kern AD, Dewey CN, Pachter L, Myers E, Langley CH (2007) Population genomics: whole
 43 genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biology*,
 44 **5**, e310.
- 45 Beltrán M, Jiggins CD, Bull V, Linares M, Mallet J, McMillan WO, Bermingham E (2002)
 46 Phylogenetic discordance at the species boundary: comparative gene genealogies among
 47 rapidly radiating *Heliconius* butterflies. *Molecular Biology and Evolution*, **19**, 2176-2190.

- 1 Bengtsson BO (1985) The flow of genes through a genetic barrier. In: Evolution essays in
2 honor of John Maynard Smith. pp. 31-42 (J.J. Greenwood, P.H. Harvey, and M. Slatkin,
3 eds.), Cambridge University Press, Cambridge.
- 4 Berlocher SH, Feder JL (2002) Sympatric speciation in phytophagous insects: moving
5 beyond controversy? *Annual Review of Entomology*, **47**, 773-815.
- 6 Black, WC, Baer CF, Antolin MF, DuTeau, NM (2001) Population genomics: genome-wide
7 sampling of insect populations. *Annual Review of Entomology*, **46**, 441-469.
- 8 Blank J, Stauss MJ, Tomiuk J, Fietz J, Segelbacher G (2007) Habitat type does not affect
9 population genetic structure in sympatric great tits (*Parus major*) and blue tits (*P.*
10 *caeruleus*). *Journal of Negative Results*, **4**, 1-14.
- 11 Bodmer WF, Parsons PA (1962) Linkage and recombination in evolution. *Advances in*
12 *Genetics*, **11**, 1-100.
- 13 Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect
14 candidate loci for adaptation along a gradient of altitude in the common frog (*Rana*
15 *temporaria*). *Molecular Biology and Evolution*, **23**, 773-783.
- 16 Bonin A, Ehrich D, Manel S (2007) Statistical analysis of amplified fragment length
17 polymorphism data: a toolbox for molecular ecologists and evolutionist. *Molecular Ecology*,
18 **16**, 3737-3758.
- 19 Bowcock AM, Kidd JR, Mountain JL, Hebert JM, Carotenuto L, Kidd KK, Cavalli-Sforza, LL
20 (1991) Drift, admixture and selection in human evolution – a study with DNA
21 polymorphisms. *Proceedings of the National Academy of Sciences USA*, **88**, 839-843.
- 22 Bradshaw HD, Schemske DW (2003). Allele substitution at a flower colour locus produces a
23 pollinator shift in monkeyflowers *Nature*, **426**, 176-178.
- 24 Broughton RE, Harrison RG (2003) Nuclear gene genealogies reveal historical, demographic,
25 and selective factors associated with speciation in field crickets. *Genetics*, **163**, 1389-1401.
- 26 Brown KM, Burk LM, Henagan LM, Noor, MA (2004) A test of the chromosomal
27 rearrangement model of speciation in *Drosophila pseudoobscura*. *Evolution*, **58**, 1856-60.
- 28 Bull V, Beltran M, Jiggins CD, McMillan WO, Bermingham E, Mallet J (2006) Polyphyly and
29 gene flow between non-sibling *Heliconius* species. *BMC Biology*, **21**, 11.
- 30 Bulmer MG (1972) Multiple niche polymorphism. *American Naturalist*, **106**, 254-257.
- 31 Butlin RK (2005) Recombination and speciation. *Molecular Ecology*, **14**, 2621-2635.
- 32 Campbell D, Bernatchez L (2004) Generic scan using AFLP markers as a means to assess
33 the role of directional selection in the divergence of sympatric whitefish ecotypes.
34 *Molecular Biology and Evolution*, **21**, 945-956.
- 35 Cano JM, Matsuba C, Mäkinen H, Merilä J (2006) The utility of QTL-linked markers to detect
36 selective sweeps in natural populations - a case study of the *Eda* gene and a linked marker
37 in threespine stickleback. *Molecular Ecology*, **15**, 4613-4621.
- 38 Carmichael LE, Krizan J, Nagy JA, Fuglei E, Dumond M, Johnson D, Veitch A, Berteaux D,
39 Strobeck C. 2007. Historical and ecological determinants of genetic structure in arctic
40 candid. *Molecular Ecology*, **16**, 3466-3483.
- 41 Charlesworth D, Charlesworth B (1975) Theoretical genetics of Batesian mimicry II. Evolution
42 of supergenes. *Journal of Theoretical Biology*, **55**, 305-324.
- 43 Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection, balanced
44 polymorphism and background selection on equilibrium patterns of genetic diversity in
45 subdivided populations. *Genetical Research*, **70**, 155-174.
- 46 Clarke CA, Sheppard PM (1960) Supergenes and mimicry. *Heredity*, **14**, 175-185.
- 47 Colosimo PF, Hosemann KE, Balabhadra S, Villarrea G, Dickson M, Grimwood J, Schmutz J,

- 1 Myers RM, Schluter D, Kingsley DM (2005) Widespread parallel evolution in sticklebacks
2 by repeated fixation of ectodysplasin alleles. *Science*, **5717**, 1928-1933.
- 3 Cooper ML (2000) Random amplified polymorphic DNA analysis of southern brown bandicoot
4 (*Isoodon obesulus*) populations in Western Australia reveals genetic differentiation related
5 to environmental variables. *Molecular Ecology*, **9**, 469-479.
- 6 Coyne JA, Orr HA. 2004. *Speciation*. Sinauer Associates, Inc. Sunderland, MA.
- 7 Crespi BJ (2007) Sly FOXP2: Genomic conflict in the evolution of language. *Trends in*
8 *Ecology and Evolution*, **22**, 174-175.
- 9 Crispo E, Bentzen, P, Reznick, DN, Kinnison MT, Hendry AP (2006) The relative influence of
10 natural selection and geography on gene flow in guppies. *Molecular Ecology*, **15**, 49-62.
- 11 Curry RA, Currie SL, Bernatchez L, Saint-Laurent R (2004). The rainbow smelt, *Osmerus*
12 *mordax*, complex of Lake Utopia: threatened or misunderstood? *Environmental Biology of*
13 *Fishes*, **69**, 153–166.
- 14 De Queiroz K (2005). Ernst Mayr and the modern concept of species. *Proceedings of National*
15 *Academy of Sciences USA*, **102**, 6600-6607.
- 16 Derome N, Duchesne P, Bernatchez L (2006) Parallelism in gene transcription among sympatric
17 lake whitefish (*Coregonus clupeaformis* Mitchell) ecotypes. *Molecular Ecology*, **15**, , 1239–
18 1249.
- 19 Dettman JR, Sirjusingh C, Kohn LM, Anderson JB (2007) Incipient speciation by
20 divergent adaptation and antagonistic epistasis in yeast. *Nature*, **447**, 585-588.
- 21 Dobzhansky T (1936) Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila*
22 *pseudoobscura* hybrids. *Genetics*, **121**, 113-125.
- 23 Dobzhansky T (1937) *Genetics and the Origin of Species*. Columbia University Press, New
24 York.
- 25 Dobzhansky T (1951) *Genetics and the origin of species*. 3rd. Columbia University Press, New
26 York.
- 27 Dopman EB, Perez L, Bogdanowicz SM, Harrison RG (2005) Consequences of
28 reproductive barriers for genealogical discordance in the European corn borer. *Proceedings*
29 *of the National Academy of Sciences, USA*, **102**, 14706-14711.
- 30 Drès M, Mallet J (2002). Host races in plant-feeding insects and their importance in
31 sympatric speciation. *Philosophical Transactions of the Royal Society of London B*, **357**,
32 471-492.
- 33 Egan SP, Nosil P, Funk DJ (2008) Selection and genomic differentiation during ecological
34 speciation: isolating the contributions of host-association via a comparative genome scan of
35 *Neochlamisus bebbianae* leaf beetles. *Evolution*, **62**, 1162-1181.
- 36 Emelianov I, Marec F, Mallet J (2004) Genomic evidence for divergence with gene flow
37 in host races of the larch budmoth. *Proceedings Royal Society of London B*, **271**, 97-105.
- 38 Endler JA (1973) Gene flow and population differentiation. *Science*, **179**, 243-250.
- 39 Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric
40 distances among DNA haplotypes : application to human mitochondrial DNA restriction
41 data. *Genetics*, **131**, 479-491.
- 42 Faubet P, Gaggiotti OE (2008) A new Bayesian method to identify the environmental factors that
43 influence recent migration. *Genetics*, **178**, 1491-1504.
- 44 Feder JL, Berlocher SH, Roethele JB, Dambroski H, Smith JJ, Perry WL, Gavrilovic V,
45 Filchak KE, Rull J, Aluja M (2003a) Allopatric genetic origins for sympatric host-plant
46 shifts and race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences*
47 *USA*, **100**, 10314-10319.

- 1 Feder JL, Roethele FB, Filchak K, Niedbalski J, Romero-Severson J (2003b). Evidence for
2 inversion polymorphism related to sympatric host race formation in the apple maggot fly,
3 *Rhagoletis pomonella*. *Genetics*, **163**, 939-953.
- 4 Felsenstein J (1976) The theoretical population genetics of variable selection and migration.
5 *Annual Review of Genetics*, **10**, 253-280.
- 6 Felsenstein J (1981) Skepticism towards Santa Rosalia, or why are there so few kinds of
7 animals? *Evolution*, **35**, 124-38.
- 8 Fisher RA (1930) *The Genetical Theory of Natural Selection*. Carendon Press, Oxford.
- 9 Foll M, Gaggiotti O (2006) Identifying the environmental factors that determine the genetic
10 structure of populations. *Genetics*, **174**, 875-891.
- 11 Foll M, Gaggiotti O (2008) A genome scan method to identify selected loci appropriate
12 for both dominant and codominant markers: A Bayesian perspective. *Genetics*, in press.
- 13 Friar EA, Cruse-Sanders JM, McGlaughlin EM (2007) Gene flow in *Dubautia arborea* and *D.*
14 *ciliolata*: the roles of ecology and isolation by distance in maintaining species boundaries
15 despite ongoing hybridization. *Molecular Ecology*, **16**, 4028-4038.
- 16 Funk DJ (1998) Isolating a role for natural selection in speciation: host adaptation and sexual
17 isolation in *Neochlamisus bebbianae* leaf beetles. *Evolution*, **52**, 1744-1759.
- 18 Funk DJ, Omland KE (2003) The frequency, causes and consequences of species level paraphyly
19 and polyphyly: insights from animal mitochondrial DNA. *Annual Review of Ecology,*
20 *Evolution and Systematics*, **34**, 397-423.
- 21 Funk DJ, Filchak KE, Feder JL (2002) Herbivorous insects: model systems for the
22 comparative study of speciation ecology. *Genetica*, **116**, 251-267.
- 23 Funk DJ, Nosil P, Etges B (2006) Ecological divergence exhibits consistently positive
24 associations with reproductive isolation across disparate taxa. *Proceedings of the National*
25 *Academy of Sciences USA*, **103**, 3209-3213.
- 26 Funk DJ, Egan SP, Nosil P. "Isolation by Adaptation" in *Neochlamisus* leaf beetles:
27 Host-related selection promotes neutral genomic divergence. *Proceedings of the Royal*
28 *Society of London B*, in review.
- 29 Gavrilets S (2004) *Fitness landscapes and the origin of species*. Princeton Univ. Press,
30 Princeton, NJ.
- 31 Gavrilets S, Cruzan MB (1998) Neutral gene flow across single locus clines. *Evolution*, **52**,
32 1277-1284.
- 33 Gavrilets S, Vose A (2005) Dynamic patterns of adaptive radiation. *Proceedings of the National*
34 *Academy of Sciences USA*, **102**, 18040-18045.
- 35 Geraldes A, Ferrand N, Nachman NW (2006) Contrasting patterns of introgression at X-linked
36 loci across the hybrid zone between subspecies of the European rabbit (*Oryctolagus*
37 *cuniculus*). *Genetics*, **173**, 919-933.
- 38 Grahame JW, Wilding CS, Butlin RK (2006) Adaptation to a steep environmental gradient and
39 an associated barrier to gene exchange in *Littorina saxatilis*. *Evolution*, **60**, 268-278.
- 40 Haig D (2004) Genomic imprinting and kinship: how good is the evidence? *Annual Review of*
41 *Genetics*, **38**, 553-585.
- 42 Haldane JBS (1930) A mathematical theory of natural and artificial selection. Part IV. Isolation.
43 *Proceedings of the Cambridge Philosophical Society*, **26**, 220-230.
- 44 Haldane JBS (1932) *The Causes of Evolution*. Longmans, New York.
- 45 Hahn MW (2006) Accurate inference and estimation in population genomics. *Molecular*
46 *Biology and Evolution*, **23**, 911-918.
- 47 Harr B (2006) Genomic islands of differentiation between house mouse subspecies. *Genome*

- 1 Res, **16**, 730-737.
- 2 Harrison RG (2001) Molecular changes at speciation. *Annual Review of Ecology and*
3 *Systematics*, **22**, 281-308.
- 4 Hartl DL, Clark AG (1997) Principles of population genetics, 3rd edn. Sinauer Associates, Inc.
5 Sunderland, MA.
- 6 Hawthorne DJ, Via S (2001) Genetic linkage of ecological specialization and reproductive
7 isolation in pea aphids. *Nature*, **412**, 904-907.
- 8 Hedrick (PW) 2005 A standardized genetic differentiation measure. *Evolution*, **59**, 1633-1638.
- 9 Hedrick PW (2006) Genetic polymorphism in heterogeneous environments: The age of
10 genomics. *Annual Review of Ecology and Systematics*, **37**, 67-93.
- 11 Hendry AP (2004) Selection against migrants contributes to the rapid evolution of
12 ecologically-dependent reproductive isolation. *Evolutionary Ecology Research*, **6**, 1219-
13 1236.
- 14 Hendry AP, Day T, Taylor EB (2001) Population mixing and the adaptive divergence of
15 quantitative traits in discrete populations: A theoretical framework for empirical tests.
16 *Evolution*, **55**, 459-466.
- 17 Hendry AP, Taylor EB (2004) How much of the variation in adaptive divergence can be
18 explained by gene flow? - An evaluation using lake-stream stickleback pairs. *Evolution*, **58**,
19 2319-2331.
- 20 Hendry AP, Day T (2005) Population structure attributable to reproductive time:
21 isolation by time and adaptation by time. *Molecular Ecology*, **14**, 901-906.
- 22 Hey J (2006) Recent advances in assessing gene flow between diverging populations and
23 species. *Current Opinion in Genetics and Development*, **16**, 592-596.
- 24 Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration
25 rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura*
26 and *D-persimilis*. *Genetics*, **167**, 747-760.
- 27 Hey J, Waples RS, Arnold ML, Butlin RK, Harrison RG (2003) Understanding and confronting
28 species uncertainty in biology and conservation. *Trends in Ecology and Evolution* **18**, 597-
29 603.
- 30 Hill WG, Robertson A (1966) The effect of linkage on limits to artificial selection. *Genetical*
31 *Research*, **8**, 269-294.
- 32 Hoekstra HE, Drumm KE, Nachman MW (2004) Ecological genetics of adaptive color
33 polymorphism in pocket mice: geographic variation in neutral and selected genes.
34 *Evolution*, **58**, 1329-1341.
- 35 Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA
36 sequence data. *Genetics*, **132**, 583-589.
- 37 Joost S, Bonin A, Bruford MW, Després L, Conord C, Erhardt G, Taberlet P (2007) A
38 spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape
39 genomics approach of adaptation. *Molecular Ecology*, **16**, 3955-3969.
- 40 Joron M, Papa R, Beltrán M, Chamberlain N, Mavárez J, Baxter S, Bermingham E, Humphray
41 S, Rogers J, Beasley H, Barlow K, ffrench-Constant R, Mallet J, McMillan WO, Jiggins CD
42 (2006) A Conserved Supergene Locus Controls Colour Pattern Diversity in *Heliconius*
43 Butterflies *Plos Biology*, **4**, e303.
- 44 Jump AS, Hunt JM, Martinez-Izquierdo JA, Penuelas J (2006) Natural selection and climate
45 change: temperature-linked spatial and temporal trends in gene frequency in *Fagus*
46 *sylvatica*. *Molecular Ecology*, **15**, 3469-3480.
- 47 Kaplan NL, Hudson RR, Langley CH (1989) The "hitchhiking effect" revisited *Genetics*, **123**,

- 1 887-899.
- 2 Kim Y, Nielsen R (2004) Linkage disequilibrium as a signature of selective sweeps.
3 *Genetics*, **167**, 1513-1524.
- 4 Kimura M (1956) A model of genetic system which leads to closer linkage by natural
5 selection. *Evolution*, **10**, 278-287.
- 6 Kimura M (1968) Evolutionary rate at the molecular level. *Nature*, **217**, 624-626.
- 7 Kimura M (1986) DNA and the neutral theory. *Philosophical Transactions of the Royal
8 Society of London B*, **312**, 343-354.
- 9 King JL, Jukes TH (1969) Non-darwinian evolution. *Science*, **164**, 788-798.
- 10 Kirkpatrick M, Barton NH (2006) Chromosome inversions, local adaptation and speciation.
11 *Genetics*, **173**, 419-434.
- 12 Kirkpatrick M, Johnson T, Barton N (2002) General models of multilocus evolution. *Genetics*,
13 **161**, 1727-1750.
- 14 Kojima K, Schaffer HE (1964) Accumulation of epistatic gene complexes. *Evolution*, **18**,
15 127-129.
- 16 Kouyos RD, Otto SP, Bonhoeffer S (2006) Effect of varying epistasis on the evolution of
17 recombination. *Genetics*, **173**, 589-597.
- 18 Kronforst MR, Young LG, Kapan DD, McNeely C, O'Neill RJ, Gilbert LE (2006) Linkage of
19 butterfly mate preference and wing color preference cue at the genomic location of wingless.
20 *Proceedings of the National Academy of Sciences USA*, **103**, 6575-6580.
- 21 Lande R (1992) Neutral theory of quantitative genetic variance in an island model with local
22 extinction and colonization. *Evolution*, **46**, 381-389.
- 23 Latta RG (1998) Differentiation of allelic frequencies at quantitative trait loci affecting locally
24 adaptive traits. *American Naturalist*, **151**, 283-292.
- 25 Latta RG (2003) Gene flow, adaptive population divergence and comparative population
26 structure across loci. *New Phytologist*, **161**, 51-58.
- 27 Le Corre V, Kremer A (2003) Genetic variability at neutral markers, quantitative trait loci and
28 trait in a subdivided population under selection. *Genetics*, **164**, 1205-1219.
- 29 Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of
30 selective neutrality of polymorphisms. *Genetics*, **74**, 175-195.
- 31 Liu A, Burke JM (2006) Patterns of nucleotide diversity in wild and cultivated sunflower.
32 *Genetics*, **173**, 321-330.
- 33 Llopart A, Lachaise D, Coyne JA (2005) Multilocus analysis of introgression between two
34 sympatric sister species of *Drosophila*: *Drosophila yakuba* and *D. santomea*. *Genetics*, **171**,
35 197 - 210.
- 36 Lu GQ, Bernatchez L (1999) Correlated trophic specialization and genetic divergence in
37 Sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): Support for the ecological
38 speciation hypothesis. *Evolution*, **53**, 1491-1505.
- 39 Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of
40 population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981-
41 994.
- 42 MacCallum CJ, Nürnberger B, Barton NH, Szymura JM (1998) Habitat preference in *Bombina*
43 hybrid zone in Croatia. *Evolution*, **52**, 227-239.
- 44 Machado C, Kliman, RM, Markert JM, Hey J (2002) Inferring the history of speciation from
45 multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and its close
46 relatives. *Molecular Biology and Evolution*, **19**, 472-488.

- 1 Machado CA, Haselkorn TS, Noor MAF (2007) Evaluation of the genomic extent of effects of
2 fixed inversion differences on intraspecific variation and interspecific gene flow in
3 *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **175**, 1289-1306.
- 4 Maddison WP (1997) Gene trees in species trees. *Systematic Biology*, **46**, 523-536.
- 5 Mäkinen HS, Cano JM and Merilä J. 2008a. Identifying footprints of directional and balancing
6 selection in marine and freshwater threespine stickleback (*Gasterosteus aculeatus*)
7 populations. *Molecular Ecology* doi:10.1111/j.1365-294X.2008.03714.x
- 8 Mäkinen HS, Shikano T, Cano JM and Merilä J 2008b. Hitchhiking mapping reveals a candidate
9 genomic region for natural selection in three-spined stickleback chromosome VIII. *Genetics*
10 **178**, 435-465
- 11 Mallet J, Barton NH (1989) Strong natural selection in a warning-color hybrid zone. *Evolution*,
12 **43**, 421-431.
- 13 Mallet J (1995) A species definition for the modern synthesis. *Trends in Ecology and Evolution*,
14 **10**, 294-299.
- 15 Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*,
16 **20**, 229-237.
- 17 Mallet J (2006) What has *Drosophila* genetics revealed about speciation? *Trends in Ecology and*
18 *Evolution*, **21**, 186-193.
- 19 Mallet J, Beltran M, Neukirchen W, Linares M (2007) Natural hybridization in heliconiine
20 butterflies: the species boundary as a continuum. *BMC Evolutionary Biology*, **7**, 28.
- 21 Manly BFJ (1997) Randomization and Monte Carlo methods in biology. 2nd edition. Chapman
22 and Hall, New York.
- 23 Martin NH, Bouck AC, Arnold ML (2005) Loci affecting long-term hybrid survivorship in
24 Louisiana irises: implications for reproductive isolation and introgression. *Evolution*, **59**,
25 2116-2124.
- 26 Maynard Smith J, Haigh J (1974) The hitch-hiking effect of a favorable gene. *Genetical*
27 *Research Cambridge*, **23**, 23-35.
- 28 Mayr E (1947) Ecological factors in speciation. *Evolution*, **1**, 263-288.
- 29 Mayr E (1963) *Animal Species and Evolution*. Harvard University Press, Cambridge.
- 30 McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends*
31 *in Ecology and Evolution*, **17**, 285-291.
- 32 Merila J, Crnokrak P (2001) Comparison of genetic differentiation at marker loci and
33 quantitative traits. *Journal of Evolutionary Biology*, **14**, 892-903.
- 34 Miller NJ, Ciosi M, Sappington TW, Ratcliffe ST, Spencer JT, Guillemaud T (2007) Genome
35 scan of *Diabrotica virgifera virgifera* for genetic variation associated with crop rotation
36 tolerance. *Journal of Applied Entomology*, **131**, 378-385.
- 37 Manoukis NC, Powell JR, Toure MB, Sacko A, Edillo FE, Coulibaly MB, Traore SF, Taylor CE,
38 Besansky NJ (2008) A test of the chromosomal theory of ecotypic speciation in *Anopheles*
39 *gambiae*. *Proceedings of the Royal Society of London B* **105**, 2940-2945.
- 40 Morjan CL, Rieseberg LH (2004) How species evolve collectively: implications of gene
41 flow and selection for the spread of advantageous alleles. *Molecular Ecology*, **13**, 1341-
42 1356.
- 43 Musiani M, Leonard JA, Cluff HD, Gates CC, Mariani S, Paquet PC, Vila C, Wayne RK
44 (2007) Differentiation of tundra/taiga and boreal coniferous forest wolves: genetics, coat
45 colour and association with migratory caribou. *Molecular Ecology*, **16**, 4149-4170.
- 46 Muller HJ (1940) Bearings of the *Drosophila* work on systematics. In: *The new systematics*
47 (ed. J.S. Huxley), pp. 185-268. Clarendon Press, Oxford.

- 1 Muller HJ (1942) Isolating mechanisms, evolution and temperature. *Biological Symposia*, **6**,
2 71-125.
- 3 Murray MC, Hare MP (2006) A genomic scan for divergent selection in a secondary
4 contact zone between Atlantic and Gulf of Mexico oysters, *Crassostrea virginica*. *Molecular*
5 *Ecology*, **15**, 4229-4242.
- 6 Nabours RK, Larson I, Hartwig N (1933) Inheritance of color patterns in the grouse locust
7 *Acrydium arenosum* Burmeister. (Tettigidae). *Genetics*, **18**, 159-171.
- 8 Nachman MW (2006) Detecting selection at the molecular level. In: *Evolutionary Genetics,*
9 *Concepts and Case Studies*, pp. 103-118 (edited by C.W. Fox and J.B. Wolf). Oxford
10 University Press, Oxford.
- 11 Nakazato T, Jung M-K, Housworth EA, Rieseberg KH, Gastony GJ (2007) A genome wide study
12 of reproductive barriers between allopatric populations of a homosporous fern, *Ceratopteris*
13 *richardii*. *Genetics*, **177**, 1141-1150
- 14 Navarro A, Barton NH (2003) Accumulating postzygotic isolation genes in parapatry: A
15 new twist on chromosomal speciation. *Evolution*, **57**, 447-459.
- 16 Nei N, Li W-H (1973) Linkage disequilibrium in subdivided populations. *Genetics*, **75**, 213-
17 219.
- 18 Nielsen R (2005) Molecular signatures of natural selection. *Annual Review Genetics*, **39**, 197-
19 218.
- 20 Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: A Markov chain
21 Monte Carlo approach. *Genetics*, **158**, 885-896.
- 22 Noor MAF, Feder JL (2006) Speciation genetics: evolving approaches. *Nature Review*
23 *Genetics*, **7**, 851-861.
- 24 Noor MAF, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the
25 reproductive isolation of species. *Proceedings of the National Academy of Sciences USA*,
26 **98**, 12084-12088.
- 27 Noor MAF, Garfield DA, Schaeffer SW, Machado CA (2007) Divergent between the *Drosophila*
28 *pseudoobscura* and *D. persimilis* genome sequences in relation to chromosomal inversions.
29 *Genetics*, **177**, 1417-1428.
- 30 Nosil P (2007) Divergent host-plant adaptation and reproductive isolation between ecotypes of
31 *Timema cristinae* walking-sticks. *American Naturalist*, **169**, 151-162.
- 32 Nosil P (2008) Speciation with gene flow may be common. *Molecular Ecology*, **17**, 2103-2106
- 33 Nosil P, Crespi BJ (2004) Does gene flow constrain trait divergence or vice-versa? A test
34 using ecomorphology and sexual isolation in *Timema cristinae* walking-sticks. *Evolution*,
35 **58**, 101-112.
- 36 Nosil P, Sandoval CP (2008) Ecological niche dimensionality and the evolutionary diversification
37 of stick insects. *PLoS ONE*, **3**, e1907.
- 38 Nosil, P, Crespi BJ, Sandoval CP (2002) Host-plant adaptation drives the parallel evolution of
39 reproductive isolation. *Nature*, **417**, 440-443.
- 40 Nosil P, Vines TH, Funk DJ (2005) Perspective: Reproductive isolation caused by natural
41 selection against immigrants from divergent habitats. *Evolution*, **59**, 705-719.
- 42 Nosil P, Crespi BJ, Sandoval CP, Kirkpatrick M (2006) Migration and the genetic covariance
43 between habitat preference and performance. *American Naturalist*, **167**, E66-E78.
- 44 Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-
45 stick ecotypes: 'isolation-by-adaptation' and multiple roles for divergent selection.
46 *Evolution*, **62**, 316-336.
- 47 Oetjen K, Reusch TBH (2007) Genome scans detect consistent divergent selection among

- 1 subtidal vs. intertidal populations of the marine angiosperm *Zostera marina*. *Molecular*
2 *Ecology*, **16**, 5156-5167.
- 3 Ogden R, Thorpe RS (2002) Molecular evidence for ecological speciation in tropical
4 habitats. *Proceedings of the National Academy of Sciences USA*, **99**, 13612-13615.
- 5 Ohta T (1992) The nearly neutral theory of molecular evolution. *Annual Review of Ecology*
6 *and Systematics*, **23**, 263-286.
- 7 Ohta T (2002) Near-neutrality in evolution of genes and gene regulation. *Proceedings of the*
8 *National Academy of Sciences USA*, **99**, 16134-16137.
- 9 Ólafsdóttir, GÁ, Snorrason SS, Ritchie MG (2006) Morphological and genetic divergence of
10 intralacustrine stickleback morphs in Iceland: a case for selective differentiation? *Journal*
11 *Evolution Biology*, **20**, 603-616.
- 12 O'Malley KG, Camara MD, Banks MA (2007) Candidate loci reveal genetic differentiation
13 between temporally divergent migratory runs of Chinook salmon (*Oncorhynchus*
14 *tshawytscha*). *Molecular Ecology*, in press.
- 15 Orr HA (1995) The population genetics of speciation: the evolution of hybrid
16 incompatibilities. *Genetics*, **139**, 1805-1813.
- 17 Orr HA (2005) The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, **6**,
18 119-127.
- 19 Orr HA, Turelli M (2001) The evolution of postzygotic isolation : accumulating Dobzhansky-
20 Muller incompatibilities. *Evolution*, **55**, 1085-1094.
- 21 Orr HA, Masly JP, Presgraves DC (2004) Speciation genes. *Current Opinion in Genetics and*
22 *Development*, **14**, 675-679.
- 23 Ortiz-Barrientos D, Reiland J, Hey J, Noor MAF (2002) Recombination and the divergence of
24 hybridizing species. *Genetica*, **116**, 167-178.
- 25 Owuor ED, Fahima T, Beharav A, Korol A, Nevo E (1999) RAPD divergence caused
26 by microsite edaphic selection in wild barley. *Genetica*, **105**, 177-192.
- 27 Panova M, Hollander J, Johannesson K (2006) Site-specific genetic divergence in parallel
28 hybrid zones suggests nonallopatric evolution of reproductive barriers. *Molecular Ecology*,
29 **15**, 4021-4031.
- 30 Parchman TL, Benkman CW, Britch SC (2006) Patterns of genetic variation in the adaptive
31 radiation of New World crossbills (Aves : *Loxia*). *Molecular Ecology*, **15**, 1873-1887.
- 32 Parisod C, Christin P-A (2008) Genome-wide association to fine-scale ecological
33 heterogeneity within a continuous population of *Biscutella laevigata* (Brassicaceae). *New*
34 *Phytologist*, in press.
- 35 Payseur BA, Krenz JG, Nachman MW (2004) Differential patterns of introgression across the
36 X chromosome in a hybrid zone between two species of house mice. *Evolution*, **58**, 2064-
37 78.
- 38 Pialek J, Barton NH (1997) The spread of an advantageous allele across a barrier: The
39 effects of random drift and selection against heterozygotes. *Genetics*, **145**, 493-504.
- 40 Pilot M, Jedrzejewski W, Branicki W, Sidorovich VE, Jedrzejewska B, Stachura K, Funk SM
41 (2006) Ecological factors influence population genetic structure of European grey wolves.
42 *Molecular Ecology*, **15**, 4533-4553.
- 43 Presgraves DC, Balagopalan L, Abmayr SA, Orr HA (2003) Adaptive evolution drives
44 divergence of a hybrid inviability gene between two species of *Drosophila*. *Nature*, **243**,
45 715-719.
- 46 Presgraves DC (2007) Does genetic conflict drive molecular evolution of nuclear transport genes
47 in *Drosophila*? *BioEssays*, **29**, 386-391.

- 1 Prowell, D.P. 1998. Sex linkage and speciation in Lepidoptera. pp. 309-319 in S. Berlocher and
2 D. Howard (eds.), *Endless Forms: Species and Speciation*. Oxford Press, NY.
- 3 Putnam AS, Scriber JM, Andolfatto P (2007) Discordant divergence times among Z
4 chromosome regions between two ecologically distinct swallowtail butterfly species.
5 *Evolution*, **61**, 912-927.
- 6 Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R,
7 Farhadian SF, Ward R, Lander ES (2001) Linkage disequilibrium in the human genome.
8 *Nature*, **411**, 199-204.
- 9 Rice WR (1988) Intergenomic conflict, interlocus antagonistic coevolution, and the evolution
10 of reproductive isolation. Pp. 261-270 in D.J. Howard and S.H. Berlocher (eds.) *Endless
11 forms : Species and Speciation*. Oxford University Press, New York.
- 12 Riebler A, Held L, Stephan W (2008) Bayesian variable selection for detecting adaptive genomic
13 differences among populations. *Genetics*, **178**, 1817-1829.
- 14 Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology and
15 Evolution*, **16**, 351-358.
- 16 Rieseberg LH, Whitton J, Gardner K. (1999) Hybrid zones and the genetic architecture of a
17 barrier to gene flow between two wild sunflower species. *Genetics*, **152**, 713-727.
- 18 Roberge C, Guderley H, Bernatchez L (2007) Genome-wide Identification of Genes under
19 Selection: Gene Transcription Qst Scan in Diverging Atlantic Salmon Subpopulations.
20 *Genetics*, **177**, 1011-1022.
- 21 Robertson A (1961) Inbreeding in artificial selection programmes. *Genetical Research*, **2**, 189-
22 194.
- 23 Rocha LA, Robertson DR, Roman J, Bowen BW (2005) Ecological speciation in tropical reef
24 fishes. *Proceedings of the Royal Society London B*, **272**, 573-579.
- 25 Roe AD, Sperling FH (2007) Population structure and species boundary delimitation of cryptic
26 *Dioryctria* moths: an integrative approach. *Molecular Ecology*, **16**, 3617-3633.
- 27 Rogers RA (1986) Population differences in quantitative characters as opposed to gene
28 frequencies. *American Naturalist*, **127**, 729-730.
- 29 Rogers SM, Bernatchez L (2005) Integrating QTL mapping and genomic scans towards the
30 characterization of candidate loci under parallel directional selection in these lake whitefish
31 (*Coregonus chupeaformis*). *Molecular Ecology*, **14**, 351-361.
- 32 Rogers SM, Bernatchez L (2007) The genetic architecture of ecological speciation and
33 the association with signatures of selection in natural lake whitefish (*Coregonus* sp.
34 Salmonidae). *Molecular Biology and Evolution*, **24**, 1423-1438.
- 35 Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under
36 isolation by distance. *Genetics*, **145**, 1219-1228.
- 37 Rueffler C, Van Dooren TJM, Leimar O, Abrams PA (2006). Disruptive selection and
38 then what? *Trend in Ecology and Evolution*, **21**, 238-245.
- 39 Rundle HD, Schluter D (2004) Natural Selection and Ecological Speciation in
40 Sticklebacks. In: *Adaptive Speciation* (eds. Dieckmann, U., Doebeli, M., Metz, J.A.J. &
41 Tautz, D.). Cambridge University Press, Cambridge, pp. 192-209.
- 42 Rundle H, Nosil P (2005) Ecological speciation. *Ecology Letters*, **8**, 336-352.
- 43 Saint-Laurent R, Legault M, Bernatchez L (2003) Divergent selection maintains adaptive
44 differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus
45 mordax* Mitchell). *Molecular Ecology*, **12**, 315-330.
- 46 Santiago E, Caballero A. (1998) Effective size and polymorphism of linked neutral loci in
47 populations under directional selection. *Genetics*, **149**, 2105-2117.

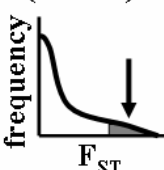
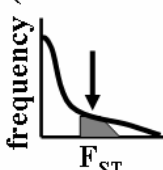
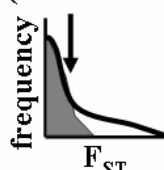




- 1 Savolainen V, Anstett MC, Lexer C, Hutton I, Clarkson JJ, Norup MV, Powell MP, Springate
2 D, Salamin N, Baker WJ (2006) Sympatric speciation in palms on an oceanic island. *Nature*,
3 **441**, 210-213.
- 4 Scheffer SJ, Hawthorne DJ (2007) Molecular evidence of host-associated genetic
5 divergence in the holly leafminer *Phytomyza glabricola* (Diptera: Agromyzidae): apparent
6 discordance among marker systems. *Molecular Ecology*, **16**, 2627-2637.
- 7 Schlötterer C (2002) A microsatellite-based multilocus screen for the identification of local
8 selective sweeps. *Genetics*, **160**, 753-763.
- 9 Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.
- 10 Scotti-Saintagne C, Mariette S, Porth I, Goicoechea PG, Barreneche T, Bodenes K, Burg K,
11 Kremer A (2004) Genome scanning for interspecific differentiation between two closely
12 related oak species [*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.]. *Genetics*, **168**, 1615-
13 1626.
- 14 Shaw, K. L. 2002. Conflict between mitochondrial and nuclear DNA phylogenies of a recent
15 species radiation: what mitochondrial reveals and conceals about modes of speciation in
16 Hawaiian crickets. *Proceedings of the National Academy of Sciences USA*, **99**, 16122-
17 16127.
- 18 Sheppard PM (1953) Polymorphism, Linkage and the Blood Groups. *American Naturalist*,
19 **87**, 283-294.
- 20 Sinervo B, Svensson E (2002) Correlational selection and the evolution of genomic
21 architecture. *Heredity*, **89**, 329-338.
- 22 Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*,
23 **16**, 393-430.
- 24 Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations.
25 *Evolution*, **47**, 264-279.
- 26 Smith TB, Calsbeek R, Wayne RK, Holder KH, Pires D, Bardeleben C (2005) Testing
27 alternative mechanisms of evolutionary divergence in an African rain forest passerine bird.
28 *Journal of Evolutionary Biology*, **18**, 257-268.
- 29 Steinfartz S, Weitere M, Tautz D (2007) Tracing the first step to speciation: ecological and
30 genetic differentiation of a salamander population in a small forest. *Molecular Ecology*, **16**,
31 4550-4561.
- 32 Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative
33 genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158-
34 170.
- 35 Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population
36 divergence. *Molecular Ecology*, **14**, 671-688.
- 37 Storz JF, Nachman MW (2003) Natural selection on protein polymorphism in the rodent
38 genus *Peromyscus*: evidence from interlocus contrasts. *Evolution*, **57**, 2628-2635.
- 39 Strasburg JL, Rieseberg LH (2008) Molecular demographic history of the annual sunflowers
40 *Helianthus annuus* and *H. petiolaris* – large effective population sizes and rates of long-term
41 gene flow. *Evolution*, in press.
- 42 Svensson EI, Abbot JK, Gosden TP, Coreau A (2007) Female polymorphisms, sexual
43 conflict and limits to speciation processes in animals. *Evolutionary Ecology*, in press.
- 44 Sword GA, Joern A, Senior LB (2005) Host plant-associated genetic differentiation in the
45 snakeweed grasshopper, *Hesperotettix viridis* (Orthoptera: Acrididae). *Molecular Ecology*,
46 **7**, 2197-2205.
- 47 Tsumura Y, Kado T, Takahashi T, Tani N, Ujino-Ihara T, Iwata H (2007) Genome scan to

- 1 detect genetic structure and adaptive genes of natural populations of *Cryptomeria japonica*.
2 Genetics, **176**, 2393–2403.
- 3 Turner JRG (1967a) On supergenes. I. The evolution of supergenes. American Naturalist,
4 **101**, 195–221.
- 5 Turner JRG (1967b) Why does the genotype not congeal? Evolution, **21**, 645–656.
- 6 Turner TL, Hahn MW (2007) Locus- and population specific selection and differentiation
7 between incipient species of *Anopheles gambiae*. *Molecular Biology and Evolution*, **24**,
8 2132–2138.
- 9 Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles*
10 *gambiae*. PLOS Biology, **3**, 1572–1578.
- 11 Turner TL, Levine MT, Eckert ML, Begun DJ (2008) Genomic analysis of adaptive
12 differentiation in *Drosophila melanogaster*. Genetics, **179**, 455–475.
- 13 Vandepitte K, Jacquemyn H, Roldan-Ruiz I, Honnay O (2007). Landscape genetics of the self-
14 compatible forest herb *Geum urbanum*: effects of habitat age, fragmentation and local
15 environment. *Molecular Ecology*, **16**, 4171–4179.
- 16 Vasemagi A, Primmer CR (2005) Challenges for identifying functionally important genetic
17 variation: the promise of combining complementary research strategies. *Molecular Ecology*,
18 **14**, 3623–3642.
- 19 Vasemagi A, Nilsson J, Primmer CR (2005) Expressed sequence tag-linked microsatellites as a
20 source of gene-associated polymorphisms for detecting signatures of divergent selection in
21 Atlantic salmon (*Salmo salar* L.). *Molecular Biology and Evolution*, **22**, 1067–1076.
- 22 Via S, Bouck AC, Skillman S (2000) Reproductive isolation between divergent races of pea
23 aphids on two hosts. II. Selection against migrants and hybrids in the parental environments.
24 *Evolution*, **54**, 1626–1637.
- 25 Via S, West J (2008) The genetic mosaic suggests a new role for hitchhiking in ecological
26 speciation. *Molecular Ecology*, in press.
- 27 Via S (2001) Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology*
28 *and Evolution*, **16**, 381–390.
- 29 Vines TH, Schluter D (2006) Strong assortative mating between allopatric sticklebacks as a by-
30 product of adaptation to different environments. *Proceedings of the Royal Society of*
31 *London, Series B*, **273**, 911–916.
- 32 Vines TH, Köhler SC, Thiel M, Ghira I, Sands TR, MacCallum CJ, Barton NH, Nürnberger B
33 (2003) The maintenance of reproductive isolation in a mosaic hybrid zone between the fire-
34 bellied toads *Bombina bombina* and *B. variegata*. *Evolution*, **57**, 1876–1888.
- 35 Vitalis R, Dawson K, Boursot P (2001) Interpretation of variation across marker loci as evidence
36 of selection. *Genetics*, **158**, 1811–1823.
- 37 Vitalis R, Dawson K, Boursot P, Belkhir K (2003) DetSel 1.0: a computer program to detect
38 markers responding to selection. *Journal of Heredity*, **94**, 429–431.
- 39 Wang RL, Wakeley J, Hey J (1997) Gene flow and natural selection in the origin of *Drosophila*
40 *pseudoobscura* and close relatives. *Genetics*, **147**, 1091–1106.
- 41 Whitlock MC (1999) Neutral additive genetic variance in a metapopulation. *Genetical*
42 *Research Cambridge*, **74**, 215–221.
- 43 Wilding CS, Butlin RK, Grahame J (2001) Differential gene exchange between parapatric
44 morphs of *Littorina saxatilis* detected using AFLP markers. *Journal of Evolutionary*
45 *Biology*, **14**, 611–619.
- 46 Wood HM, Grahame JW, Humphray S, Rogers J, Butlin RK (2008) Sequence differentiation in

- 1 regions identified by a genome scan for local adaptation. *Molecular Ecology*, **17**, 3123-
2 3135.
- 3 Wright S (1931) Evolution in mendelian populations. *Genetics*, **16**, 97-159.
- 4 Wright S (1940) Breeding structure of populations in relation to speciation. *American Naturalist*,
5 **74**, 232.
- 6 Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.
- 7 Wu CI (2001) The genic view of the process of speciation. *Journal of Evolutionary*
8 *Biology*, **14**, 851-865.
- 9 Wu .-I, Ting C-T (2004) Genes and speciation. *Nature Review Genetics*, **5**, 114-122.
- 10 Xie X, Rull J, Michel A, Velez S, Forbes AA, Lobo N, Aluja M, Feder JL (2007) Hawthorn-
11 infesting populations of *Rhagoletis pomonella* in Mexico and speciation mode
12 plurality. *Evolution*, **61**, 1091-1105.
- 13 Yatabe Y, Kane NC, Scotti-Saintagne C, Rieseberg LH (2007) Rampant gene exchange across
14 a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H.*
15 *petiolaris*. *Genetics*, **175**, 1883-1893.
- 16 Zayed A, Whitfield CW (2008) A genome-wide signature of positive selection
17 in ancient and recent invasive expansions of the honey bee *Apis mellifera*. *Proceedings of*
18 *the National Academy of Sciences USA*, **105**, 3421–3426.
- 19
- 20 The authors all use a combination of experimental and molecular genetic approaches to study
21 speciation. Patrik Nosil and Daniel Funk’s research investigates insect speciation in response to
22 adaptation to different host-plant species. Patrik Nosil’s work focuses on walking-stick insects in
23 California while Daniel Funk’s work focuses on leaf beetles in eastern North America. The two
24 authors now collaborate on these systems. Daniel Ortiz-Barrientos’ doctoral work treated the
25 genetic basis of reinforced mating preferences in *Drosophila*. His current work involves speciation,
26 ecological genetics, and sexual selection in flowering plants.

- 1 Glossary
- 2
- 3 Heterogeneous genomic divergence – variation among genomic regions, of any given size or
4 content, in their degree of population differentiation
- 5
- 6 Divergent selection – selection that acts in contrasting directions between two populations,
7 including the special case where selection favors both extremes within a single population (i.e.,
8 disruptive selection)
- 9
- 10 F_{ST} – a measure of population differentiation based on the proportion of genetic variation that
11 occurs between populations versus within populations
- 12
- 13 Outlier loci – loci whose level of population differentiation statistically exceeds neutral
14 expectations, where neutral expectations are often inferred using simulations
- 15
- 16 Loose linkage – weak physical linkage along a chromosome of a neutral locus to a selected locus
17 that is sufficient to yield somewhat elevated genetic differentiation via genetic hitchhiking, but not
18 strong enough to yield outlier status
- 19
- 20 Isolation-by-Adaptation (IBA) – a positive correlation between the degree of adaptive phenotypic
21 divergence between populations and their level of molecular genetic differentiation, independent
22 from geographic distance (IBA involving neutral loci that are not tightly linked to those under
23 divergent selection is of special interest)
- 24
- 25 Genomic island of divergence – a region of the genome, of any size, whose divergence exceeds
26 neutral expectations
- 27

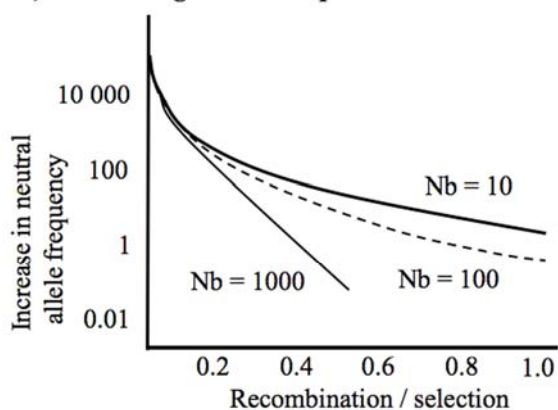
1 **Fig. 1.** Predicted patterns of genetic differentiation for different classes of loci (highlighted by grey
2 coloration in the F_{ST} distributions) when divergent selection is present and absent (ecologically
3 divergent and ecologically similar population pairs, respectively), for divergence both without and
4 with gene flow (allopatry and non-allopatry, respectively). When divergent selection occurs, the
5 processes expected to most strongly affect differentiation for each class of locus are noted above the
6 expected degree of divergence. When divergent selection is absent, only drift or gene flow affect
7 each class of locus (F_{ST} distributions would look similar but with a less extreme right-hand tail).
8 Ecological divergence yields outlier status under both geographic scenarios. Class II loci should
9 exhibit IBA in both allopatry and non-allopatry, while class III loci should only exhibit IBA under
10 the non-allopatric scenario because neutral divergence in allopatry is independent of degree of
11 ecological divergence. Patterns of heterogeneous genomic divergence are evident, such that genetic
12 divergence generally varies as follows: class I > class II > class III.

Δ = 'degree of divergence'	class of locus		
	I) selected or tightly-linked (outlier)	II) neutral loosely-linked (non-outlier)	III) neutral unlinked (non-outlier)
ecological and geographic context			
allopatry	selection	selection (weakly)	drift via physical barriers to gene flow
ecologically divergent 	↓ strong Δ	↓ moderate Δ	↓ weak Δ
(drift only) ecologically similar 	weak Δ	weak Δ	weak Δ
non-allopatry	selection, gene flow	selection (weakly), gene flow	drift via selectively restricted gene flow
ecologically divergent 	↓ moderate Δ	↓ weak Δ	↓ weak (potentially) Δ
(gene flow only) ecologically similar 	very weak / no Δ	very weak / no Δ	very weak / no Δ

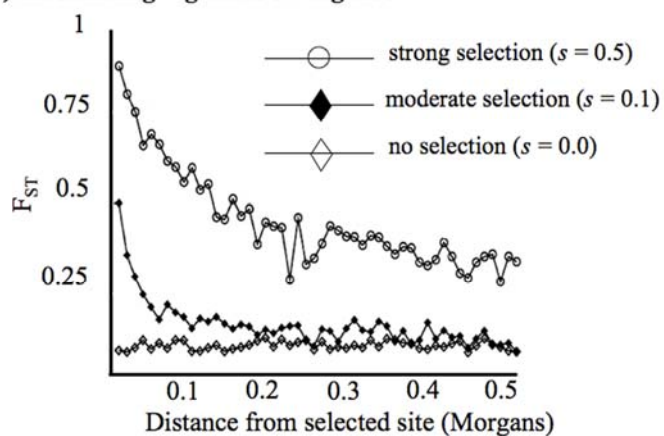
1
2
3
4

1 **Fig. 2.** Divergent selection's contributions to heterogeneous genomic divergence. A) Genetic
2 hitchhiking – allele frequencies. The net increase in allele frequency of a neutral allele (y-axis) via
3 hitchhiking, under the spread of an advantageous new mutation (Barton 2000 for further details on
4 scaling of the axes). N_b refers to neighborhood (i.e., population) size. Modified from Barton (2000)
5 and reprinted with permission of the Royal Society of London. B) Hitchhiking – genetic divergence.
6 Simulation results showing the effects of divergent selection on levels of differentiation at linked
7 neutral sites, for two demes exchanging migrants (background selection present, but similar patterns
8 observed in its absence). Modified from Charlesworth et al. (1997, see for details) with permission
9 of Cambridge University Press. C) Relationship between neutral F_{ST} and N_m across empirical
10 studies of animals (data from Morjan and Rieseberg 2004). Note that F_{ST} increases when divergent
11 selection reduces N_m (e.g., by selecting against immigrants). For example, reducing N_m from five
12 to just below one considerably increases F_{ST} (see thick arrow).

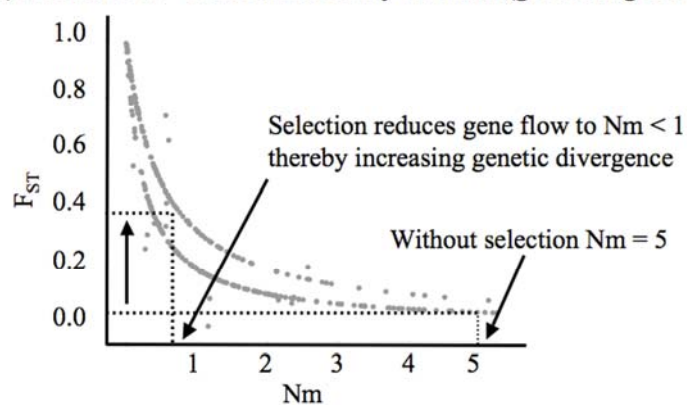
A) Hitchhiking – allele frequencies



B) Hitchhiking – genetic divergence



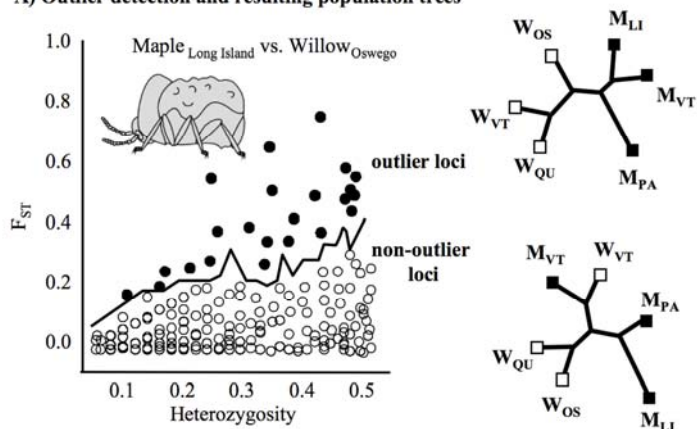
C) Neutral loci – drift facilitated by selection (generating IBA)



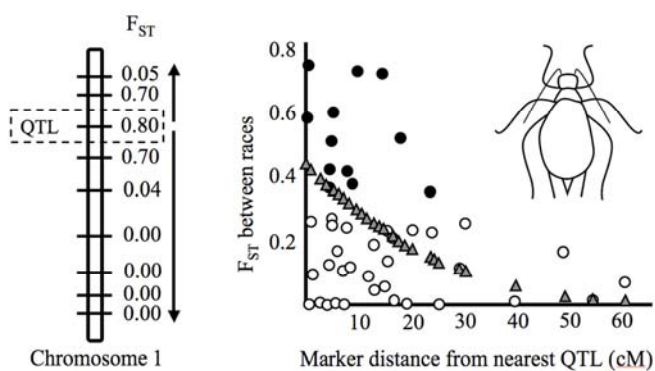
1
2
3
4
5
6

1 **Fig. 3.** Empirical patterns illustrating heterogeneous genomic divergence. A) Outlier detection using
2 the genome scan approach, for a pair of *Neochlamisus bebbianae* leaf beetle populations.
3 Simulations determine the upper level of genetic divergence expected under neutrality, and loci that
4 exceed this ‘neutrality threshold’ (solid line, in this case 95% quantile) are inferred to have evolved
5 under divergent selection. If outliers are highly replicated across population pairs that have diverged
6 ecologically, then population trees from pooled outlier loci are likely to group populations
7 according to ecology (e.g., box color indicates host plant), reflecting divergent selection. In
8 contrast, trees from putatively neutral non-outlier loci may group populations according to
9 geography, reflecting spatial opportunities for gene flow. Modified from Egan et al. (2008) and
10 reprinted with permission of the Society for the Study of Evolution. B) Genetic differentiation
11 between clover- and alfalfa-associated pea aphids as a function of distance from QTL. On the left,
12 the panel depicts a hypothetical decrease in F_{ST} along a chromosome with increasing distance from
13 a QTL. On the right, real data depict markers that were outliers (filled circles) and non-outliers
14 (open circles). Triangles show the predicted values from a logistic regression of outlier status on
15 distance to the nearest QTL. Modified from Via and West (2008) and reprinted with permission of
16 Blackwell publishing. C) Pattern depicting IBA, in which neutral genetic differentiation between
17 population pairs of *Timema cristinae* walking-stick insects is positively correlated with their degree
18 of adaptive divergence (here measured as the degree of divergence in host plant preferences),
19 independent of geographic distance. Modified from Nosil et al. (2008) and reprinted with
20 permission of the Society for the Study of Evolution.

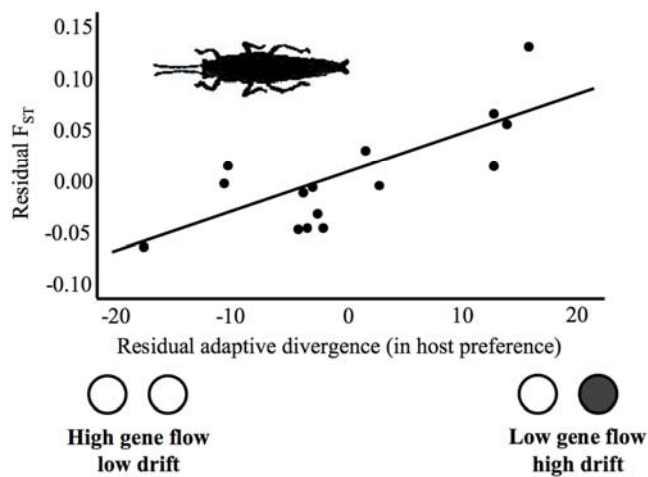
A) Outlier detection and resulting population trees



B) Genetic divergence as a function of distance from QTL



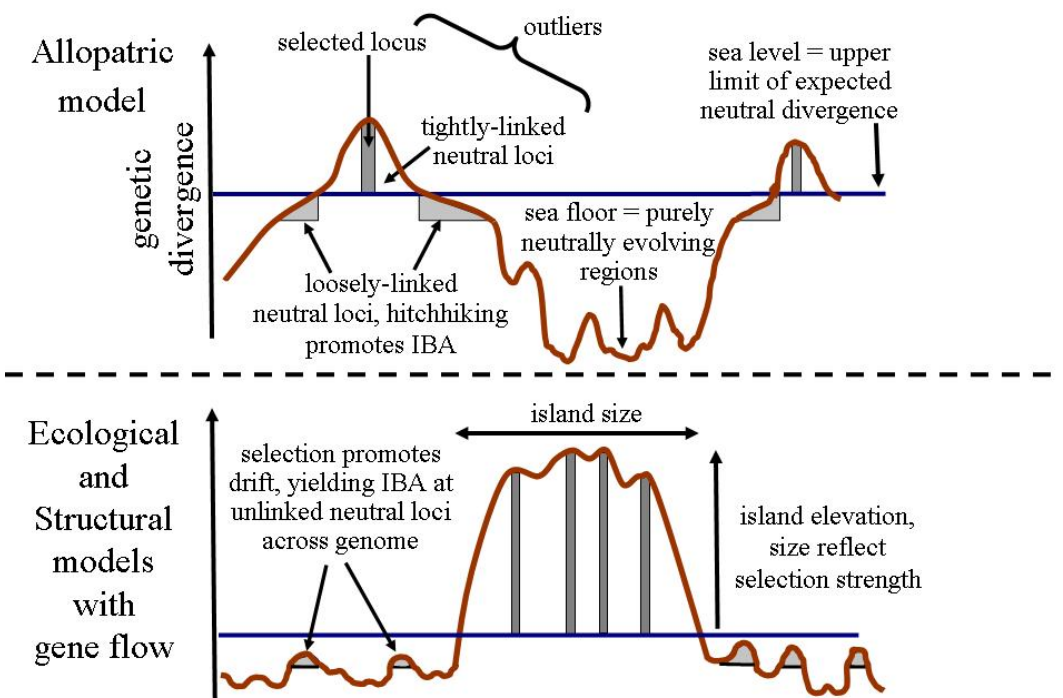
C) 'Isolation-By-Adaptation'



1 **Fig. 4.** The metaphor of genomic islands of divergence. A) Schematic illustration of expected
2 patterns of differentiation along a chromosome under models with and without gene flow. Sea level
3 represents a 'neutrality threshold'. Islands are genomic regions exhibiting greater differentiation
4 than expected under neutrality, thereby rising above sea level. See text for further details. B) An
5 empirical example involving incipient species of *Anopheles gambiae*. The bottom two panels depict
6 patterns of differentiation across chromosome two (Turner et al. 2005). Grey areas were identified
7 as highly differentiated in sliding window analyses, with differentiation further confirmed by
8 sequencing loci within these regions (red circles). A large island is evident on the left arm, near the
9 centromere. A small island is also evident on the right arm. The top panel treats a subsequent study
10 (Turner and Hahn 2007) where portions of all annotated genes within the smaller island were
11 sequenced. As predicted by Turner et al. (2005), sequence differentiation peaked within the 'island'
12 (between the grey lines). However, the fine scale data from Turner and Hahn (2007) allow more
13 detailed characterization of the nature of the island, showing, for example, that differentiation drops
14 off rapidly with distance from the region of maximum differentiation (i.e., the island is very steep).
15 Modified from the original studies and reprinted with permission of the Public Library of Science
16 and the Society of Molecular Biology and Evolution.

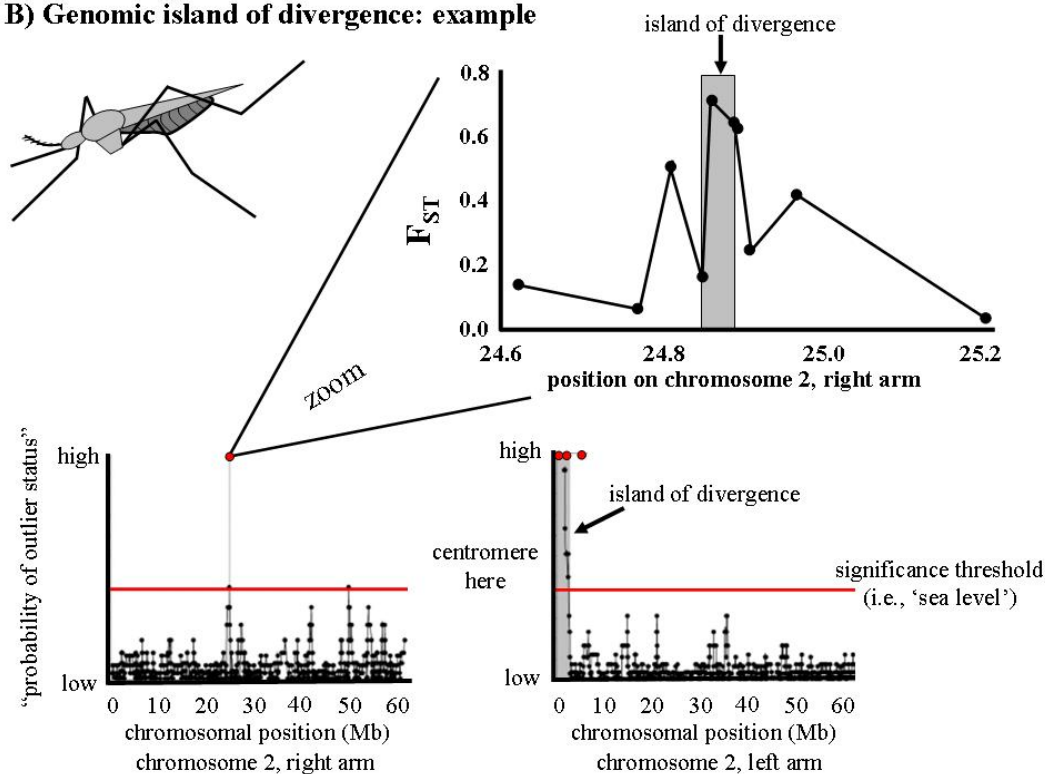
17
18
19
20
21
22
23
24
25
26
27

A) Genomic island of divergence: visual metaphor



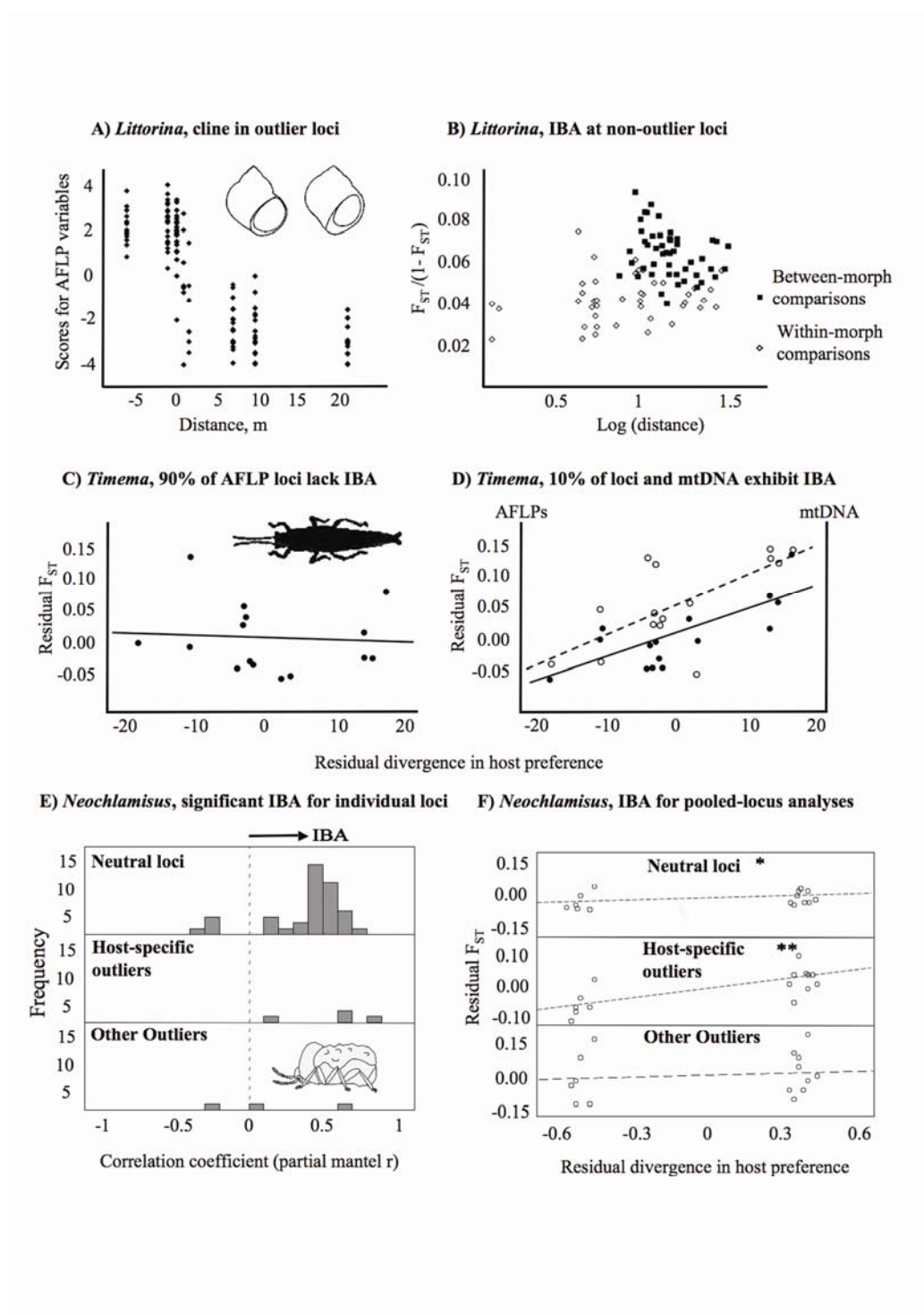
1

B) Genomic island of divergence: example



2
3
4

1 **Fig. 5.** Patterns of outlier differentiation and IBA in three study systems where both have been
2 examined. These examples illustrate the heterogeneous nature of genomic differentiation: outlier
3 loci exhibit greater divergence than non-outlier loci (A, B), and differentiation varies among non-
4 outlier loci such that (only) a fraction of them exhibit strong IBA (C-F). A) A cline in allele
5 frequencies for outlier loci from comparisons of upper and lower shore ecotypes of *Littorina*
6 *saxatilis* snails. B) Greater neutral (i.e., non-outlier) differentiation between than within the *L.*
7 *saxatilis* ecotypes, for a given geographic distance, illustrating IBA. Modified from Grahame et al.
8 (2006). C) Lack of association between adaptive phenotypic divergence and genetic differentiation
9 at 90% of AFLP loci between host plant ecotypes of *Timema cristinae*. D) Evidence for IBA in *T.*
10 *cristinae* at 10% of AFLP loci (filled circles, solid line) and mtDNA (open circles, dashed line)
11 based on analyses controlling for geographic distance. Modified from Nosil et al. (2008). E) For
12 *Neochlamisus bebbianae* host forms: the 11% of individual loci exhibiting significant associations
13 between adaptive and genetic divergence, controlling for geographic distance (shown are partial r
14 values from Mantel tests), for three classes of loci (host-specific outliers, other outliers not
15 associated with host plant, and putatively neutral non-outliers). F) Pooled analyses of these locus
16 classes in *N. bebbianae*, revealing strong and highly significant IBA for host-specific outliers, more
17 moderate but significant IBA for neutral loci, and no evidence of IBA for other outliers. Modified
18 from Funk et al. (in review). The *Littorina* and *Timema* figures are reprinted with permission of the
19 Society for the Study of Evolution.



1
2
3
4
5
6

Table 1. Predictions concerning the roles of divergent selection role in heterogeneous genomic divergence. Both effects on particular gene regions (see also Fig. 1) and models of the growth of differentiated regions (i.e., ‘islands’; Fig. 4) are treated.

Type of locus	Predictions
locus directly under selection	- strong genetic divergence at these loci (as illustrated by, e.g., outlier status)
locus tightly physically linked to those under selection	- strong genetic divergence at these loci (as illustrated by, e.g., outlier status)
locus loosely physically linked to those under selection	- moderately increased genetic divergence compared to neutrality - IBA pattern should be observed even beyond the spatial scale of gene flow (e.g., among completely allopatric populations)
neutral, unlinked loci (affected by ‘general barriers’)	- IBA expected at the spatial scale of gene flow <i>if</i> gene flow is sufficiently reduced to allow divergence via genetic drift (i.e., selection and reproductive isolation must be strong); - likelihood of IBA increases with decreasing N_e (i.e., as drift becomes more effective)
Models for growth of differentiated regions	Predictions
I. Allopatric model	- many differentiated regions (i.e., ‘genomic islands’) - islands need not be clustered within the genome, and will often be small in size - number and elevation of islands increases with time since population divergence
II. Ecological model (divergent selection with gene flow)	- large islands will occur - genomic clustering of islands, with genes affecting local adaptation and reproductive isolation residing within these clusters, perhaps fewer islands than the allopatric model - islands need not involve chromosomal inversions (e.g., supergenes) - small islands, which have not yet grown, are also possible
III. Structural model	- genomic differentiation is facilitated by chromosomal inversions and other factors that reduce recombination, dependent on the degree to which genes affecting local adaptation and reproductive isolation reside within inversions - the extent of this facilitation also depends on how far outside an inversion the introgression reducing effects of the inversion extend - islands can be larger than in the other models, and persist for longer periods of time

Table 2. Summary of genome scan studies. Provided is the study system, a description of the divergent forms being compared, the type and number of markers used, the study design (“compare divergence” = compare population divergence among loci), the main result with respect to outliers, patterns observed in population trees, and levels of linkage disequilibrium (LD) among outlier loci. More specifically, the “population trees differ” column addresses whether population trees from neutral (non-outlier) loci group populations by geography whereas trees that include (or only use) outlier loci group populations by ecology. Here, ‘Yes’ indicates that outlier-including trees result in groupings that are consistent with monophyly of the same ecological type, and ‘Partially’ indicates that outlier-including trees are more structured by ecology than neutral trees but do not provide evidence of complete monophyly by ecological type. For detailed consideration of the methodological robustness of each study, see the online supplementary materials.

Study system	Divergent forms	Data (marker, no. of loci)	Study design	Main result – outliers	Population Trees differ?	LD?	Reference
Studies examining parallel divergence / association with ecological parameters							
1. <i>Littorina saxatilis</i> (intertidal snails)	upper vs. lower shore ecotypes	306 AFLP loci	Compare divergence between sympatric ecotypes at three different shores	5% of loci consistently outliers between ecotypes at all three shores	Yes	low, except in middle of cline, implying any LD is generated by migration	Wilding et al. 2001; Grahame et al. 2006; Wood et al. 2008
2. <i>Coregonus clupeaformis</i> (whitefish)	dwarf vs. normal lake ecotypes	440 AFLP loci	Compare divergence between sympatric ecotypes from four different lakes	up to 3.2% of loci outliers between ecotypes, 1.4% outliers in all four lakes	Yes	possibly moderate, as outlier loci were non-randomly distributed among primer combinations	Campbell and Bernatchez 2004
3. <i>Salmo salar</i> (Atlantic salmon)	saltwater vs. brackish vs. freshwater habitats	95 genomic and EST derived mini- and microsatellites	Compare divergence for four population pairs occupying different	23-26% of loci outliers, 9% of EST derived loci were outliers at different spatial	N/A	N/A	Vasemagi et al. 2005

			habitats, at local and global spatial scales	scales, population pairs, and geographic regions			
4. <i>Rana temporaria</i> (common frog)	altitudinal gradient	392 AFLP loci	compare divergence between twelve population pairs differing in altitude	8-14% of loci outliers, roughly 2% of loci involved in altitudinal adaptation specifically	Yes	N/A	Bonin et al. 2006
5. <i>Hylobius abietis</i> (pine weevil)	N/A	83 AFLP loci	compare divergence among six forest regions, and relate genetic divergence to environmental variables	6% of loci outliers, two loci (2.4%) strongly correlated with variation in environmental parameters	N/A	N/A	Joost et al. 2007
Studies examining divergence within versus between ecological forms							
6. <i>Zeiraphera diniana</i> (herbivorous insect)	sympatric larch- vs. pine-feeding host forms	1291 AFLP loci	compare heterogeneity of genomic divergence between sympatric host forms (four comparisons) versus between populations of the same form	heterogeneity of genomic divergence between sympatric host forms, but not between geographic populations of the same form	N/A	possibly high, reflecting concentration of loci involved in between-race differentiation on just a few chromosomes	Emelianov et al. 2004

			(six comparisons)				
7. <i>Cryptomeria japonica</i> (coniferous tree)	slender branched, soft-leaved vs. rough branched hard-leaved variety	139 CAPS markers based on cDNA sequences	compare divergence using 29 populations, considering divergence among populations within and between varieties	10% of loci outliers, 6.5% outliers only between varieties	N/A	N/A	Tsumura et al. 2007
8. <i>Timema cristinae</i> (walking-stick insect)	<i>Adenostoma</i> and <i>Ceanothus</i> host plant ecotypes	534 AFLP loci	compare divergence between nine population pairs on different hosts and between six population pairs on the same host	8% of loci outliers in multiple comparisons, 1-2% outliers only in between-host comparisons with weak parallel divergence among pairs	partially	low, but higher than for non-outliers within both allopatric and parapatric populations, consistent with some weak physical linkage	Nosil et al. 2008
9. <i>Neochlamisus bebbinae</i> (leaf beetle)	willow and maple host forms	447 AFLP loci	compare divergence between nine population pairs on different hosts and between six population pairs on the same host	15% of loci outliers in multiple comparisons, 5% outliers only in between-host comparisons with strong parallel divergence	Yes	low	Egan et al. 2008

				among pairs			
10. <i>Phytomyza glabricola</i> (holly leafminer)	<i>Ilex glabra</i> vs. <i>I. coriacea</i> holly host forms	45 AFLP loci	compare divergence between two population pairs on different hosts and between two pairs on the same host	18% of loci outliers, 11-16% of loci outliers in a between-host pair with 9% of loci outliers in both between-host pairs	N/A; but mtDNA groups by geography whereas AFLPs group by host (i.e., ecology)	N/A	Scheffer and Hawthorne 2007
11. <i>Diabrotica virgifera</i> (western corn rootworm)	crop rotation resistant vs. wild type strains	253 AFLP loci	compare divergence for nine between-strain population pairs and between six pairs of the same strain	% overall outliers not reported, <0.5% (one locus) an outlier in between-strain comparisons only	N/A	N/A	Miller et al. 2007
12. <i>Zostera marina</i> (marine flowering plant)	tidal creek versus tidal flat	25 EST-derived and anonymous microsatellite markers	compare divergence between populations in different habitats for three independent population pairs	12% of loci outliers, with these exhibiting replicated divergence for population pairs in different habitats, but not being outliers in same-habitat population pairs	N/A	N/A	Oetjen and Reusch 2007
Other types of studies							
13. <i>Quercus</i>	high nutrient	389 markers	compare	12% of loci	N/A	low, outlier loci	Scotti-

<i>robur</i> L. and <i>Q. petraea</i> (oaks)	soil vs. drier habitat species	(isozymes, AFLPs, SCARs, microsatellites, and SNPs)	divergence between species using 7-10 population pairs, depending on the marker	outliers		distributed among many different linkage groups	Saintagne et al. 2004
14. <i>Picea abies</i> (spruce)	N/A	125 AFLPs, 25 SSRs and 2 EST mapped markers	compare divergence for three pairs of populations	6% of loci outliers, but not replicated across population pairs	N/A	low, outlier loci distributed among many different linkage groups	Acheré et al. 2005
15. <i>Anopheles gambiae</i> (African malaria mosquito)	M vs. S form	hybridized DNA of single mosquitoes from samples of M and S forms to microarray chips (1,577 probes)	compare divergence between forms using seven M samples and seven S samples	1.2% of the genome, from only three genomic regions, highly differentiated between forms	N/A	N/A	Turner et al. 2005; Turner and Hahn 2007
16. <i>Mus musculus</i> (house mouse)	<i>Mus musculus musculus</i> vs. <i>M. m. domesticus</i> .	>10,000 SNP markers	compare divergence between subspecies using 22 wild-derived inbred strains	7.5% of the autosomal genome highly differentiated, comprising eight genomic regions	N/A	N/A	Harr 2006
17. <i>Crassostrea virginica</i> (oyster)	N/A	215 AFLP loci	compare divergence between a population pair	1.4% of loci outliers	N/A	low	Murray and Hare 2006
18. <i>Fagus sylvatica</i> (beech tree)	temperature gradient	254 AFLP loci	compare divergence between five	<0.5% of loci (one locus) outliers,	N/A	N/A	Jump et al. 2006

			population pairs differing in temperature	associated with divergence in temperature			
19. <i>Howea forsteriana</i> and <i>H. belmoreana</i> (oceanic palms)	divergent flowering times and soil preferences	274 AFLP loci	Compare divergence between two species	1% of loci outliers	N/A	N/A	Savolainen et al. 2006
20. <i>Peromyscus spp.</i> (mice)	N/A	10-37 protein coding genes, depending on mouse species	Compare divergence among populations within six different species	10.5% outliers (percent outliers within datasets ranged from 0-30), parallel divergence for some loci	N/A	N/A	Storz and Nachman 2003

Table 3. Summary of studies providing information on Isolation-by-Adaptation (IBA), where IBA refers to a positive association between the degree of adaptive phenotypic (or ecological) divergence between populations and the level of genetic differentiation at putatively neutral markers. All findings of IBA provided here control for geographic distance. Detection of IBD is also noted. Three main types of studies are reported: 1) population genetic studies explicitly examining IBA, generally using distance matrices of adaptive and neutral genetic divergence, 2) population genetic studies employing an AMOVA framework, and 3) one illustrative example of IBA each in a mosaic hybrid zone and in a phylogeographic study. For a more complete version of this table, including study design, molecular markers, and much more detailed results, see the online supplementary materials.

Organism	IBA?	IBD?	Reference
Population genetic studies generally using distance matrices			
1. <i>Isodon obesulus</i> (brown bandicoot)	YES	NO	Cooper 2000
2. <i>Anolis roquet</i> (anole lizard)	YES	NO	Ogden and Thorpe 2002
3. <i>Poecilia reticulata</i> (guppies)	NO	YES	Crispo et al. 2006
4. <i>Canis lupus</i> (European grey wolf)	YES	YES	Pilot et al. 2006
5. <i>Coregonus clupeaformis</i> (whitefish)	YES	N/A	Lu and Bernatchez 1999
6. <i>Timema cristinae</i> (herbivorous stick-insect)	YES for some AFLP loci and for mtDNA	YES (but weak)	Nosil et al. 2008
7. <i>Dubautia arborea</i> and <i>D. ciliolata</i> (silversword plant)	YES	NO	Friar et al. 2007
8. <i>Littorina saxatilis</i> (intertidal snail)	YES	NO	Grahame et al. 2006
9. <i>Geum urbanum</i> (forest herb)	NO	NO	Vandepitte et al. 2007
10. <i>Canis lupus</i> (North American grey wolf)	YES	YES	Musiani et al. 2007
11. <i>Hordeum spontaneum</i> (wild barley)	YES	YES	Owuor et al. 1999
12. <i>Canis lupus</i> (North American grey wolf)	YES (for some variables)	NO (but not directly tested)	Carmichael et al. 2007
13. <i>Alopex lagopus</i> (arctic fox)	NO	NO	Carmichael et al. 2007
14. <i>Zostera marina</i> (marine flowering plant)	NO	YES	Oetjen and Reusch 2007
15. <i>Biscutella laevigata</i> (flowering plant)	YES	YES	Parisod and Christin 2008
16. <i>Neochlamisus bebbinae</i> (leaf beetle)	YES for AFLPs, NO for mtDNA	NO for AFLPs, YES for mtDNA	Funk et al., in review
Studies employing primarily an AMOVA (or similar) framework			
17. <i>Parus major</i> (great tit)	NO	N/A	Blank et al. 2007
18. <i>Parus caeruleus</i> (blue tit)	NO	N/A	Blank et al. 2007
19. <i>Osmerus mordax</i> (rainbow smelt)	NO	N/A	Curry et al. 2004

20. <i>Loxia curvirostra</i> complex (red crossbills)	YES	NO	Parchman et al. 2006
21. <i>Hesperotettix viridis</i> (grasshopper)	YES	NO (but not directly tested)	Sword et al. 2005
22. <i>Salamandra salamandra</i> (fire salamander)	YES	NO	Steinfartz et al. 2007
Phylogeographic and hybrid zone studies			
23. <i>Halichoeres</i> spp. (tropical reef fish, wrasses)	YES	NO	Rocha et al. 2005
24. <i>Bombina</i> spp. (toads)	YES	N/A	MacCullum et al. 1998; see also Vines et al. 2003

Online supplementary materials

Divergent selection, reproductive isolation, and genomic divergence

We focused on heterogeneous genomic divergence during the process of population differentiation and speciation. This requires considering the relationship between selection and reproductive isolation. Loci under divergent selection and loci causing reproductive isolation are similar in exhibiting reduced introgression (and thus greater divergence) between populations relative to other loci (Barton 1979, 1983; Barton and Hewitt 1989; Mallet 1995, 2005, 2006; Wu 2001; Wu and Ting 2004; Nosil et al. 2005). Indeed, an allele ‘a’ that confers a poor fit of the phenotype to the environment can be selected against, contributing to population divergence, whether the afflicted allele resides in the parental species (e.g., homozygote ‘aa’) or in a hybrid individual (heterozygote ‘Aa’). Yet, typically, the former scenario would be considered a case of selection and the latter an example of postmating reproductive isolation. Recognizing that selection against immigrants itself represents a form of reproductive isolation (Nosil et al. 2005), and that hybrid inviability is a manifestation of selection, helps clarify the relatedness of these concepts. Moreover, the study of gene regions differentiating under selection becomes inseparable from that of gene regions causing reproductive isolation when adaptively relevant loci pleiotropically promote reproductive isolation (Muller 1942; Funk 1998; Bradshaw and Schemske 2003; Rundle and Nosil 2005), or when selection drives the population divergence of genes causing genetic incompatibilities between populations (i.e., intrinsic postmating isolation, Presgraves et al. 2003; Orr et al. 2004; Wu and Ting 2004; Dettman et al. 2007). This may also be true when genomic regions of divergence contain loci affecting phenotypic traits under selection as well as loci affecting forms of reproductive isolation such as hybrid inviability (Noor et al. 2001; Rieseberg 2001). For simplicity, the present paper focuses on evaluating genomic differentiation in terms of divergent selection per se. Thus, selection and reproductive isolation are treated conceptually together, while recognizing that: (1) they might act at different stages in life history, with consequences for heterogeneous genomic divergence, and (2) some forms of reproductive isolation will evolve due to processes other than divergent selection, such as genetic drift.

More detailed summary of the general genetic barrier to neutral gene flow caused by selection

We draw heavily on a summary by Gavrilets (2004 p. 147-148) based on his own work and that of Barton and colleagues (e.g., Bengtsson 1985; Barton and Bengtsson 1986; Pialek and Barton 1997; Gavrilets and Cruzan 1998; Navarro and Barton 2003; Gavrilets and Vose 2005).

The basic scenario is one in which a population is subject to continuous immigration. Due to divergent local adaptation, immigrants have lower fitness than residents, yielding selection against immigrants. In this case, the spread of neutral alleles between immigrant and resident populations will be slowed, to some extent, by selection against both immigrants and subsequent immigrant-resident hybrids. In this fashion, selection against incoming locally adapted alleles will – through the death of immigrants or hybrids – act as a general, if partial, genetic barrier to the spread of neutral alleles between populations. To describe this effect, Bengtsson (1985) introduced the notion of the ‘gene flow factor’, η , defined as the probability

that a neutral allele carried by immigrants makes it into the local genetic background. The inverse of η is known as the ‘strength of genetic barrier’; Barton and Bengtsson 1986; Pialek and Barton 1997). If the migration rate (i.e., the proportion of the local population replaced by immigrants each generation) is m , then with a genetic barrier the proportion of resident neutral alleles replaced by immigrant neutral alleles per generation is:

$$(1) \quad m_e = \eta m$$

Thus, equation (1) defines the effective migration rate of neutral alleles. To better characterize η , now assume that immigrating adults differ from residents at two genes: a gene reducing the viability of F_1 hybrids to $1 - s$ (where the viability of residents is 1) and a neutral gene unlinked to the selected gene. Assuming random mating, the probability that the neutral allele makes it to the next generation is $(1 - s) / 2$. The probability that the allele survives to the next generation but remains associated with the deleterious allele is also $(1 - s) / 2$. After many generations, the probability of inclusion of the neutral allele into the local genetic background is (see Gavrilets 2004 for derivation):

$$(2) \quad \eta = \frac{1 - s}{1 + s}$$

The effective immigration rate of neutral alleles is slowed even further if there is assortative mating (equation 4 in Gavrilets 2004, p. 148).

Table S1. Robustness of studies listed in Table 2 of the main text. ‘Multiple methods’ indicates whether more than one primary type of analysis was used to detect outliers (N = No; Y = Yes, one program/simulation method was run in different ways; Y+ = Yes, more than one program or statistical procedure was used). This column refers to the actual implementation of different programs, whereas other ways of confirming outlier status (e.g., replication of outlier status across different population pairs) are dealt with in subsequent columns. ‘Quantiles’ indicates the threshold of expected neutral differentiation used to determine whether a locus was an outlier. ‘Replication across population pairs’ indicates whether the replication of outlier status across multiple population pairs was evaluated (N = No; Y = Yes; Y+, direct = Yes, and the study also explicitly evaluated if outliers were associated with a specific ecological variable, for example by being outliers only in comparisons between population pairs that differ in that variable, and never outliers in population pairs similar for that variable; Y+, indirect = same as Y+, but although an association of outliers with an ecological variable was evident in the data, it was not explicitly noted). ‘Type I error’ indicates whether this type of error was accounted for (N = No; Y = Yes, by correcting for the number of loci within a comparison (e.g., via Bonferroni correction); Y+ = Yes, via additional consideration of the number of population pairs in which the locus was an outlier). ‘Mutation rate variation’ refers to whether this possible confounding factor was discussed (N = No; Y = Yes, by arguing that gene flow negates the effects of mutation rate variation, or by comparing differentiation between regions known to differ in mutation rate; Y+ = Yes, by evaluating outliers that were associated with a specific ecological variable, a pattern unlikely to arise via mutation rate variation. We note here only studies that actually discussed this issue, but any study examining parallel divergence, and particularly those that noted associations between outliers and ecological variables, indirectly argue that mutation rate variation to be an unlikely cause of outlier behavior). Past studies suggest that divergence-based methods for detecting divergent selection are robust to demographic variability (Beaumont and Balding 1996), but we further note whether the potential confounding effects of ‘demography’ were discussed. ‘Background selection’ refers to whether it was considered. When the discussion of a particular factor was particularly explicit, we note the relevant page number.

Study system from Table 2	Multiple methods	quantiles	replication across population pairs	Type I error	Mutation rate variation	Demography	Background selection
1.	Y+	0.99	Y+, indirect	Y+, p. 616	N	N	N
2.	N	0.95	Y	Y+	N	N	N
3.	Y, Y+, p. 1069	0.99	Y	Y+, p. 1074	Y+	N	N
4.	Y, Y+, p. 775	0.95, 0.99	Y+, indirect	Y+	N	Y, p. 775	N
5.	Y+	0.95, 0.99	Y+, direct	Y	N	N	N
6.	N	0.95	Y+, direct	N	Y+, p. 98	N	N
7.	Y+, p. 2396	0.99	Y+, direct	Y, p. 2401	N	N	N
8.	Y, Y+	0.95, 0.99	Y+, direct, p. 321	Y+, p. 323	Y, Y+, p. 322	N	N
9.	Y, Y+	0.95, 0.99	Y+, direct	Y+, p. 1167	Y, Y+, p. 1167	N	N

10.	Y+	0.95	Y+, direct	Y	N	N	N
11.	Y	0.95	Y+, indirect	N	N	N	N
12.	Y+	0.95, 0.99	Y+, direct	Y+, p.5163	N	Y, p.5160	
13.	N	0.95	N	N	N	N	N
14.	N	0.95	Y	N	N	N	N
15.	Y+	0.95	Y	Y, p. e285	N	Y, p. e285	Y, p. e285
16.	Y+	0.985	Y	N	Y, p. 734	N	N
17.	Y	0.99	N	Y	N	N	N
18.	Y,Y+, p.3473	0.99	Y+, direct	Y+	N	N	N
19.	N	0.95	N	N	N	N	N
20.	Y	0.05	Y	Y	N	N	N

Table S2. Summary of studies providing information on Isolation-by-Adaptation (IBA), where IBA refers to a positive association between the degree of adaptive phenotypic (or ecological) divergence and the level of genetic differentiation (here, at putatively neutral genetic markers). Provided is the study system, a description of the taxa or ecological forms examined, the type and number of markers used, the study design, and the main result with respect to IBA and IBD (where applicable). When reporting findings for IBA, the results reported are independent of geographic distance. Three main types of studies are reported: 1) population genetic studies explicitly examining IBA, generally using distance matrices of adaptive and neutral genetic divergence, 2) population genetic studies using an AMOVA framework (Excoffier et al. 1992), where we report the percent of total genetic variation observed between ecological types versus that observed among populations within ecological types, and 3) illustrative examples of a pattern consistent with IBA in a mosaic hybrid zone and in a phylogeographic study.

Organism	Divergent forms	Marker	Study design	IBA results	IBD results	Reference
Population genetic studies generally using distance matrices						
1. <i>Isoodon obesulus</i> (brown bandicoot)	gradient in rainfall and swamp vs. forest habitat types	39 RAPDS	genetic distance was related to habitat divergence and to geographic distance using 36 population pairs	IBA detected for both annual rainfall and habitat type	IBD not detected	Cooper 2000
2. <i>Anolis roquet</i> (Carribbean lizard)	gradient from xeric coastal woodland to transitional forest to montane rainforest	7 microsatellite loci	genetic distance was compared among pairs of adjacent localities from three different transects, one of which cut through the ecological gradient ('habitat transect') and two of which did not (seven to ten localities per transect)	IBA detected, strong genetic differentiation observed only in the habitat transect, at habitat boundaries; population structuring by habitat further supported by AMOVA	IBD not detected	Ogden and Thorpe 2002
3. <i>Poecilia reticulata</i> (guppies)	high vs. low predation habitats	7 microsatellite loci	genetic distance was related to ecology (predation regime)	no evidence for IBA	IBD detected (and also an effect of the biogeographical	Crispo et al. 2006

			and geography (distance, waterfalls) using from 54 to 190 population pairs		barrier of waterfalls)	
4. <i>Canis lupus</i> (grey wolf in Europe)	various ecological factors	14 microsatellite loci, mtDNA	genetic distance was related to numerous ecological variables and to distance, using from 16 to 59 populations	IBA detected, genetic differentiation among local populations was correlated with climate, habitat type, and wolf diet composition.	IBD detected, topographic barriers nor past fragmentation could explain spatial genetic structure	Pilot et al. 2006
5. <i>Coregonus clupeaformis</i> (freshwater fish ecotypes)	dwarf vs. normal lake ecotypes (but quantitative indices of morphological divergence analyzed)	6 microsatellite loci	the correlation between genetic distance and adaptive morphological differentiation between six sympatric pairs was examined (thus there is no geographic distance between any of the six pairs)	IBA detected when results are pooled across loci, 5 of 6 loci exhibit fairly strong evidence of IBA individually ($r = 0.72-0.84$)	N/A	Lu and Bernatchez 1999
6. <i>Timema cristinae</i> (herbivorous insect)	<i>Adenostoma</i> and <i>Ceanothus</i> host plant ecotypes (but quantitative indices of adaptive divergence analyzed)	209 AFLP loci, mtDNA	genetic distance was related to quantitative indices of host-associated adaptive divergence and to geographic distance using 15 population	IBA not significant when AFLP loci were pooled; 10% of putatively neutral (i.e., non-outlier)	weak IBD detected, stronger for mtDNA than for AFLPs	Nosil et al. 2008

			pairs	individual AFLP loci show significant IBA; mtDNA shows IBA		
7. <i>Dubautia arborea</i> and <i>D. ciliolate</i> (Hawaiian silverswords)	gradient in leaf characteristics	7 microsatellite loci	using two species, genetic distance between ten population pairs was related to indices of morphometric divergence in leaf traits and to geographic distance	IBA detected	IBD not detected	Friar et al. 2007
8. <i>Littorina saxatilis</i> (intertidal snails)	upper vs. lower shore ecotypes	275 AFLP loci	genetic distance was related to ecology (shore ecotype) and geographic distance, at two different shores	IBA detected, for a given geographic distance, stronger differentiation between ecotypes relative to between samples within ecotypes	IBD not detected	Grahame et al. 2006
9. <i>Geum urbanum</i> (forest herb)	herb-layer community similarity	6 microsatellite loci	genetic distance was related to ecology (herb community) and geographic distance using 18 populations (153 pairs)	IBA not detected	IBD not detected	Vandepitte et al. 2007
10. <i>Canis lupus</i> (grey wolf in	tundra vs. forest types	14 microsatellite loci	genetic distance was related to habitat type (tundra, taiga or	IBA detected	IBD detected	Musiani et al. 2007

North America)			boreal coniferous forest) and geographic distance using 11 population groupings			
11. <i>Hordeum spontaneum</i> (wild barley)	basalt and terra rossa soil types	117 RAPD loci	genetic distance was related to habitat type (basalt and terra rossa soil types) and transect position, using two topographically separated transects	44% of loci exhibit IBA	36% of loci exhibit IBD (i.e., a correlation with transect)	Owuor et al. 1999
12. <i>Canis lupus</i> (grey wolf in North America)	various ecological variables including annual temperature, rainfall, vegetation, behavior and species of primary prey	15 microsatellite loci	genetic variation was related to a variety of ecological variables, to water barriers to gene flow, and to spatial position	IBA detected, at least for some ecological variables	not explicitly tested, although there were spatial components to genetic structure	Carmichael et al. 2007
13. <i>Alopex lagopus</i> (arctic foxes)	NA	13 microsatellite loci	genetic structure was analyzed in STRUCTURE, due to the detection of only a single genetic cluster, further matrix analyses were not conducted	no evidence for IBA, across a broad geographic area, only a single genetic cluster was detected	no evidence for IBD	Carmichael et al. 2007
14. <i>Zostera marina</i>	tidal creek versus tidal flat	25 EST-derived and	genetic distance was related to ecology	no evidence for IBA	IBD detected	Oetjen and Reusch

(marine flowering plant)		anonymous microsatellite markers	(tidal ecotype) and geographic distance using 15 population pairs			2007
15. <i>Biscutella laevigata</i> (terrestrial plant)	continuous population varying in habitat characteristics	102 AFLP loci	genetic distance was related to ecology (habitat type) and geographic distance in a continuous population	IBA detected	IBD detected	Parisod and Christin 2008
16. <i>Neochlamisus bebbinae</i> (leaf beetle)	willow and maple host forms	381 AFLP loci, mtDNA	genetic distance was related to quantitative indices of host-associated adaptive divergence and to geographic distance using 15 population pairs	IBA detected for pooled AFLP loci; 10% of putatively neutral (i.e., non-outlier) individual AFLP loci show significant IBA; mtDNA does not show IBA	IBD detected for mtDNA, but not for AFLPs	Funk et al., in review
Studies employing primarily an AMOVA (or similar) framework						
17. <i>Parus major</i> (great and, respectively)	deciduous vs. mixed-coniferous forests	9 microsatellite loci and 4 allozyme loci	partitioning of genetic structure within and among habitat types was examined using AMOVA	no evidence for IBA, genetic variance among habitats: 0.19%, $p > 0.05$ among populations within habitats: 1.60% $p < 0.05$	not directly tested	Blank et al. 2007
18. <i>Parus caeruleus</i> (blue tit)	deciduous vs. mixed-coniferous forests	9 microsatellite loci and 4	partitioning of genetic structure within and among	no evidence for IBA, genetic variance	not directly tested	Blank et al. 2007

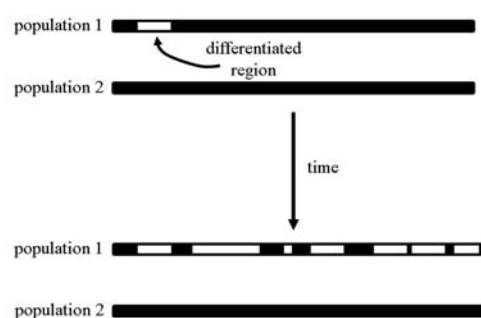
		allozyme loci	habitat types was examined using AMOVA	among habitats: <math><0.01\%, p > 0.05</math> among populations within habitats: 7.93%, $p < 0.05$		
19. <i>Osmerus mordax</i> (rainbow smelt)	giant, normal, and dwarf morphotypes	5 microsatellite loci	partitioning of genetic structure among morphotypes versus among populations within morphotypes was examined using AMOVA	no evidence for IBA, no genetic structure among morphotypes (0%, $p > 0.05$), despite appreciable structure among populations within types (10%, $p < 0.05$)	not directly tested	Curry et al. 2004
20. <i>Loxia curvirostra</i> complex (red crossbills)	eight morphologically and vocally differentiated 'call types'	440 AFLP loci	AMOVA models used to examine genetic variation between call types and among populations within call types	IBA detected, between call-type differentiation is greater (7.0%, $p < 0.05$) than that found among different geographic locations within call types (3.5%, $p < 0.05$)	IBD not detected, despite explicit tests using distance matrices	Parchman et al. 2006
21. <i>Hesperotettix viridis</i> (grasshopper)	<i>Solidago mollis</i> vs. <i>Gutierrezia sarothrae</i> host plant forms	222 AFLPs	partitioning of genetic structure among host plant forms versus among populations within	IBA detected, strong (20%) and significant variance among host forms,	no evidence for IBD, although not directly tested, insignificant (1%) variation among	Sword et al. 2005

			host plant forms was examined using AMOVA	insignificant (1%) variation among localities within forms	different geographic localities within host forms	
22. <i>Salamandra salamandra</i> (fire salamander)	stream vs. pond form	11 microsatellite loci and mtDNA sequences	partitioning of genetic structure among 33 sites was examined, in relation to geographic distance and pond type (ecology)	IBA detected, two ecologically (pond vs. stream types) differentiated groups within a relatively small forest showed signs of genetic differentiation (i.e., two main genetic clusters were correlated to larval habitat type)	IBD not detected, analysis of a large forest area (neighboring the smaller one exhibiting IBA) where all salamanders use streams showed no genetic differentiation, gene flow between ecologically similar types occurs over large distances	Steinfartz et al. 2007
Phylogeographic and hybrid zone studies						
23. <i>Halichoeres spp.</i> (tropical reef fish, wrasses)	habitat types, such as warm versus cold water habitats	mtDNA sequences	phylogeographic patterns within and among five species were attributed to effects of habitat, distance, and biogeographic barriers	IBA detected, concordance of phylogenetic partitions with habitat types	Little evidence for IBD, high genetic connectivity between similar habitats separated by thousands of kilometers	Rocha et al. 2005
24. <i>Bombina spp.</i> (toads)	pond-adapted <i>B. bombina</i> vs. puddle-adapted <i>B. variegata</i>	5 unlinked, diagnostic allozyme loci	correlates of genetic structure examined within a hybrid zone	IBA detected, genetic structure strongly associated with habitat type	N/A	MacCullum et al. 1998; see also Vines et al. 2003

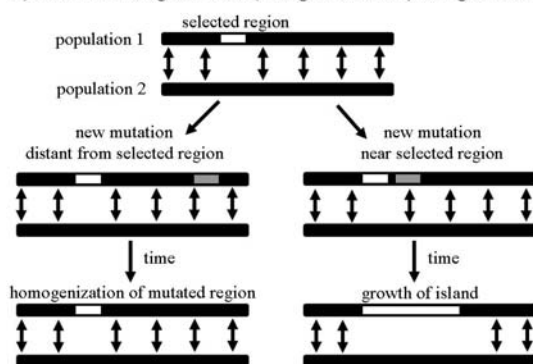
				rather than distance		
--	--	--	--	-------------------------	--	--

Figure S1. Conceptual models for the growth of genomic islands of divergence. In all cases, bars represent chromosomes, white boxes within them represent differentiated regions of the genome, and filled, black areas represent undifferentiated regions. Two-headed arrows represent regions of the genome where genetic exchange between populations is high. A) Allopatric model. Divergence proceeds unimpeded by gene flow, with the proportion of the genome differentiated between two populations being positively related to time since divergence. B) Ecological model. A new mutation (grey box within chromosome) arising near genomic regions under selection, and thus undergoing reduced introgression, has a higher likelihood of differentiating between populations than a new mutation arising in a region distant from those under selection. C) Structural model. A new mutation (grey box within chromosome) arising near an inversion has a higher likelihood of differentiating between populations than a new mutation arising in a region distant from the inversion.

A) Model I: Allopatric model



B) Model II: Ecological model (divergent selection) with gene flow



C) Model III: Structural model with gene flow

