1

Ecological explanations for (incomplete) speciation

Patrik Nosil[1,2], Luke J. Harmon[1,3] and Ole Seehausen[4,5]

[1]Zoology Department and Biodiversity Research Centre, University of British Columbia,
Vancouver BC, V6T 1Z4, Canada
[2]Current Address: Department of Ecology and Evolutionary Biology, University of Boulder,
Colorado, 80309, USA
patrik.nosil@colorado.edu

[3]Department of Biological Sciences, University of Idaho, Moscow ID, 83844, USA
lukeh@uidaho.edu

[4]Aquatic Ecology and Evolution, Institute of Ecology & Evolution, University of Bern,
Baltzerstrasse 6, CH-3012 Bern, Switzerland
[5]Centre of Ecology, Evolution and Biogeochemistry (CEEB), Swiss Federal Institute of Aquatic
Science and Technology (Eawag), Seestrasse 79, CH-6047 Kastanienbaum, Switzerland
ole.seehausen@eawag.ch

Corresponding author: Nosil, P.

Divergent natural selection has been shown to promote speciation in many taxa. However, although divergent selection often initiates the process of speciation, it often fails to complete it. A number of time-based, geographic and genetic factors have been recognized to explain this variability in how far speciation proceeds. We review here recent evidence indicating that variability in the completeness of speciation can also be associated with the nature of divergent selection itself, with speciation being greatly promoted by: (1) stronger selection on a given, single trait (the 'stronger selection' hypothesis) and (2) selection on a greater number of traits (the 'multifarious selection' hypothesis). However, evidence for each selective hypothesis is still scarce, and further work is required to determine their relative importance.

Variability in the completeness of ecological speciation

The causes of speciation have received much attention from biologists [1-11]. One hypothesis posits that divergent selection between ecological niches drives the evolution of reproductive incompatibility (see glossary). This process of 'ecological speciation' occurs because traits under divergent natural selection, or those genetically correlated with them, affect reproductive compatibility [1,2,4-11], and includes the special case where divergent selection operates directly on mate choice. Ecological speciation predicts that ecologically divergent pairs of populations will exhibit greater levels of reproductive incompatibility (e.g. reproductive isolation) than ecologically similar pairs of populations, because ecological divergence is a proxy for the presence of divergent selection [1,4-8]. Another prediction is that traits under divergent selection often affect reproductive compatibility [1,2,4,9]. There are now numerous examples supporting these predictions, and thus it is generally accepted that divergent selection can promote speciation [1,2,4,6-11].

1 However, as support for ecological speciation accumulated, it became evident that
2 divergence in this process often varies continuously (even if the end point is the development of
3 a discontinuity) [12-23]. For example, the degree of phenotypic divergence can vary
4 quantitatively [1,5,8], as can the completeness of reproductive isolation [2-4,7,8,11-14], the
5 degree of genotypic clustering [16-17], the sharpness of geographic clines in gene frequencies
6 [19], and the extent of lineage sorting [20,21] (Fig. 1, Table 1). These different means of
7 quantifying divergence can be used to measure arbitrary 'stages' of speciation, ranging from
8 continuous variation to population differentiation, ecotype formation, speciation, and post-
9 speciational divergence [2,8,11,13,24,25]. We hereafter use the term 'stage' of speciation to
10 refer to a point along this continuum of divergence, with the later stages being associated with
11 strong reproductive isolation and bimodal genotypic clustering (i.e. a strong discontinuity).
12 Notably, different species concepts can disagree on when speciation starts and when it is
13 complete, while still sharing the characteristic of having stages of divergence [12,23]. Thus, our
14 arguments apply across species concepts.
15 For the process of ecological speciation in particular, cases in which speciation is clearly
16 incomplete are not uncommon. For example, ecological divergence can be accompanied by
17 only weak reproductive isolation, low genotypic clustering, and little neutral genetic
18 differentiation, as observed in *Timema* walking-stick insects, *Pundamilia* cichlids, *Ostrinia*
19 *nubilalis* corn borers and many other organisms [8,10,13,20,24-29]. Thus, although divergent
20 selection often initiates speciation, it does not always complete it. In other instances, divergent
21 selection does not even initiate speciation [24-26], as observed in Trinidadian guppies (*Poecilia*
22 *reticulate*) [26], or distinct species pairs collapse [26-29], as observed in stickleback fishes
23 (*Gasterosteus aculeatus*) [29]. In short, there is abundant variability in the stage of ecological
24 speciation achieved, for both cases of speciation with gene flow and in examples of allopatric
25 divergence. This raises the central question reviewed here: what factors explain which stage of
26 ecological speciation is achieved?
27 We first discuss some well-recognized time-based, geographic and genetic factors that
28 affect the stage of speciation achieved (Fig. 2). These can all be viewed as non-exclusive
29 alternatives to the selective hypotheses that are the focus of our review. We then provide a
30 framework for testing the role of ecological divergence in the completeness of speciation, and
31 review studies indicating that speciation is most strongly promoted when an ecological shift
32 along a single niche axis is extreme or when a shift involves divergence in a large number of
33 independent axes. We describe two hypotheses for how such shifts so strongly promote
34 speciation: 'stronger selection' and 'multifarious selection' hypotheses. These hypotheses have
35 been discussed in the past, but the relationships among them have not been clearly established
36 [3,5,13,30-34]. Moreover, given the numerous 'non-ecological' explanations for variability in
37 speciation, an evaluation of empirical support for hypotheses that involve divergent selection is
38 appropriate. Thus, we outline critical differences between the selective hypotheses, and review
39 support for each. To keep the selective hypotheses ecologically-rooted, we focus on phenotypic
40 traits that mediate interactions between organisms and their environment. However, we note
41 that the hypotheses could be applied to the genetic level, for example by considering strong
42 selection on one gene versus weaker selection on many genes (the term 'selection' is hereafter
43 shorthand for divergent selection).
44
45 **Non-ecological factors promoting speciation**
46 Much theoretical and empirical work on the completeness of speciation has focused on

1 time-based, geographic, or genetic factors (Fig. 2). Speciation can be promoted by increased
2 time since beginning of divergence [2,14] and by geographic barriers to gene flow [2,3].
3 Speciation can also be promoted by pleiotropic effects on reproductive isolation of genes under
4 selection [2,3,35] and by physical linkage of genes under selection and those conferring
5 reproductive isolation, perhaps facilitated by chromosomal inversions [2-4,36-39]. Finally,
6 speciation is promoted by one-allele assortative mating mechanisms, which resolve the
7 antagonism between selection and recombination during divergence [4,40,41], and by abundant
8 standing genetic variation (sometimes created by interspecific hybridization) [42-44].
9 　　　　There are now examples of the above factors promoting speciation. A role for time is
10 exemplified by the positive relationship between genetic distance and levels of reproductive
11 isolation between species pairs of *Drosophila* and other taxa [2,14] and a role for geographic
12 factors comes from a large number of publications [2,3,45,46]. Empirical examples of genetic
13 factors affecting speciation also exist. For example, adaptation to different pollinators by
14 *Mimulus* monkeyflowers, via divergence in a flower color gene, pleiotropically affects the
15 probability of cross-pollination and thus hybridization [35]. Similarly, female mating preference
16 for male color in cichlids is affected by divergence in a color vision gene [10,13]. Examples of
17 physical linkage and chromosomal inversions promoting speciation include physical linkage of
18 host-plant preference and performance in *Acyrthosiphon* pea aphids [36], physical linkage
19 between sex ratio distorting genes and color in cichlid fish [47], and chromosomal inversions
20 harboring genetic variation for divergent diapause adaptation in *Rhagoletis* flies [39]. Evidence
21 for a one-allele assortative mating mechanism is provided by a study of *Drosophila*
22 *pseudoobscura* [41], and good examples for the role of standing genetic variation and
23 hybridization in speciation also exist [42-44,48,49]. These hypotheses have increased our
24 understanding of the factors driving and constraining the speciation process. The hypotheses
25 also provide explicit alternatives, albeit not mutually exclusive ones, to the ecological factors
26 discussed below.
27
28 **A unified framework for testing ecological speciation**
29 　　　　The study of ecological speciation involves isolating the association between ecological
30 divergence and the completeness of speciation, independent from the other factors discussed
31 above (Box 1). Examples of measures of ecological divergence are the extent of divergence
32 between taxa along one niche dimension (we hereafter use the term 'niche dimension' to refer
33 to an ecological axis, such as habitat use [1,30,34]), the number of niche dimensions that differ
34 between taxa, the strength of divergent selection on one trait, and the number of traits subject to
35 divergent selection. From a causal perspective we are most interested in selection itself,
36 although niche divergence is more easily measured (e.g. from environmental data). The second
37 factor, i.e. the completeness of speciation, can be measured using experimental estimates of
38 reproductive isolation or the extent of gene flow inferred from molecular markers (e.g. in
39 hybrid zones). Higher levels of reproductive isolation, less gene flow, and increased genotypic
40 clustering all indicate a later stage of the speciation process [1-4,10,16,17].
41 　　　　When only a few taxon pairs are available for analysis, qualitative comparisons can be
42 made between ecological divergence and the stage of speciation achieved. When numerous
43 taxon pairs are available, more quantitative analyses can be conducted, for example to control
44 for the time since beginning of population divergence [6,7]. This can be important because
45 reproductive isolation can increase through time via non-ecological processes such as random
46 genetic drift. Specifically, data on niche divergence or selection can be added to the regression

1 method used to study the relationship between reproductive isolation and time alone, where
2 time is generally inferred using genetic distance from molecular data (see Box 1 for caveats)
3 [2,7,14,50-54]. Multiple regression can thus be used to statistically isolate the association
4 between ecological divergence and the completeness of speciation, independent of time.
5 Notably, this framework might also be used to control for variation among taxon pairs in
6 genetic architecture, by adding such information to the regression analysis. With this analytical
7 framework in mind we turn to empirical data.
8
9 **The nature of ecological shifts and the completeness of speciation**
10   The nature of ecological shifts can affect the completeness of speciation (Box 1). Under
11 one scenario, slight shifts along a single niche dimension initiate speciation, but more extreme
12 shifts along that same dimension are required to complete speciation [3,5,7]. This idea has seen
13 few tests, because most speciation studies consider only two categories of ecological
14 divergence (ecologically similar and ecologically divergent), precluding a test of how
15 reproductive isolation varies with the quantitative degree of divergence along a niche dimension
16 [1,4,7-9]. Moreover, studies have not statistically isolated independent niche dimensions [1,4,7-
17 9], potentially confounding the magnitude of an ecological shift in a single direction with the
18 dimensionality of the shift. Nonetheless, indirect evidence supporting a role for extreme shifts
19 stems from a study which examined over 500 species pairs from eight plant, invertebrate, and
20 vertebrate taxa. This study used the multiple regression approach described above and revealed
21 that the magnitude of divergence in the single niche dimension of diet or of habitat was
22 sometimes significantly positively correlated with the degree of reproductive isolation between
23 species pairs, independent of time [7]. Assuming that increased divergence in the single niche
24 dimensions examined is not correlated with divergence in a greater number of dimensions, the
25 results support a role for extreme shifts along a single dimension in speciation. Another
26 example comes from ecological and phylogenetic studies of galling Australian thrips
27 (Thysanoptera), which suggest that extreme shifts in host-plant use promote speciation more
28 strongly than smaller shifts [55]. One caveat is that extreme niche shifts might be difficult to
29 implement, for example due to lack of suitable genetic variation, and thus could be relatively
30 rare [56-58]. The importance of extreme shifts in generating new species might thus reflect a
31 balance between them being rare but more likely to complete speciation when they do occur.
32   Another scenario involves the actual number of niche dimensions differing between
33 taxa, with divergence in one or a few dimensions initiating speciation, but with the completion
34 of speciation being characterized by divergence in many niche dimensions [3,5,10,13,34,].
35 Support for this idea stems from Lake Victoria cichlid fish (Box 1). In the incipient species pair
36 *Pundamilia pundamilia* and *P. nyererei*, divergence in numerous niche dimensions has been
37 quantified, including water depth, distance from shore in the lake, diet inferred from stable
38 isotopes, and parasite community. Reproductive isolation was measured using both
39 experimental estimates of sexual isolation and levels of gene flow in sympatry inferred from
40 molecular markers [10-13,59]. Both measures of reproductive isolation exhibit a positive
41 relationship with the dimensionality of niche divergence.
42   We note that the two types of ecological shifts are not independent from one another,
43 and can apply simultaneously. Thus, increased divergence in the single dimension of diet might
44 promote speciation, increased divergence in the single dimension of habitat might promote
45 speciation, and divergence in both diet and habitat might promote speciation more than
46 divergence in diet alone or habitat alone. A central remaining question is how do these types of

1 ecological shifts promote speciation?
2
3 **Mechanisms strongly promoting speciation: 'stronger' versus 'multifarious' selection**
4   We consider two mechanisms by which extreme or highly dimensional ecological shifts
5 promote speciation (Fig. 3). First, under the 'stronger selection' hypothesis, the completeness of
6 speciation is positively related to the strength of selection on a single trait, with very strong
7 selection on one or a few traits driving the completion of speciation [1-4]. Second, under a
8 'multifarious selection' hypothesis, the completeness of speciation is positively related to the
9 number of genetically independent traits subject to selection, with selection on many traits
10 required to complete speciation [10,13,30-34]. These two hypotheses can be visualized in terms
11 of the metaphor of an adaptive landscape: is the completion of speciation caused by increased
12 divergence between adaptive peaks in a single dimension (i.e. trait), or via the generation of
13 peaks that are separated in multiple dimensions [34]? Although both factors can contribute
14 simultaneously to speciation, the strength of divergent selection on a trait can vary
15 independently from the number of traits under selection, and thus it is useful to treat each
16 mechanism as a distinct hypothesis.
17   An important point is that one-to-one mapping is not expected between the nature of an
18 ecological shift (i.e. how extreme or multidimensional it is) and the nature of divergent
19 selection (i.e. its' strength and how many traits it acts upon). Thus both types of ecological shift
20 noted above might cause stronger selection on a given single trait, selection on a greater number
21 of traits, or both (as illustrated in Fig. 3a). For example, an extreme ecological shift along a
22 single niche dimension might cause stronger selection on a trait that was previously under
23 weaker selection, or it might result in more (i.e. new) traits being subject to selection. This
24 means that selection estimates, rather than environmental data, are required to distinguish the
25 stronger versus multifarious selection hypotheses (Box 2 for an approach using divergence in
26 phenotypic traits as a surrogate for selection on traits).
27
28 **Probability of speciation under stronger versus multifarious selection**
29
30   The probability of speciation under stronger versus multifarious selection can vary
31 according to the total strength of divergent selection, per trait selection coefficients, and the
32 nature of correlations between selected traits and other traits.
33
34 *Total selection strength*
35   Two arguments suggest that multifarious selection can be important for completing
36 speciation. First, multifarious selection can be required to generate increased total strength of
37 divergent selection in natural populations, because the strength of selection on any single trait is
38 dictated by the ecological setting, and thus can be low and never increase [1,15]. Therefore,
39 multifarious selection might be required to generate a total selection strength that is sufficient to
40 complete speciation. Second, even if divergent selection on one trait is strong, extreme
41 divergence in that trait can be constrained by a lack of suitable genetic variation [56-58] or
42 functional constraints [60,61]. In such a scenario, multifarious selection on many traits can be
43 required to generate an overall degree of trait divergence that is large enough to complete
44 speciation. Empirical studies of selection strength and levels of genetic variation in traits under
45 selection are required to test these ideas.
46

Per trait selection coefficients and correlated evolutionary response

We outline here critical differences in how genetic divergence is expected to occur under the stronger versus multifarious selection hypotheses, even when the total strength of divergent selection is held constant (Fig. 3b). These differences arise for two reasons. First, the hypotheses differ in the expected magnitude of per trait (gene) selection coefficients. Specifically, for a given total strength of selection, per trait selection coefficients will increase as the number of traits under selection decreases. Thus, per trait selection coefficients are expected to be higher when selection acts on one or a few traits (stronger selection hypothesis) relative to when selection acts on many traits (multifarious selection hypothesis). Divergence in a given trait is a function of its (per trait) selection coefficient and rates of gene flow [3,34,62]. The implication is that strong selection on a few traits will sometimes be more effective at causing and maintaining adaptive divergence in the face of gene flow, because the higher selection coefficients associated with it can more strongly overcome gene flow.

Second, divergent selection on a trait (gene) can cause divergence in other correlated traits, referred to as a 'correlated evolutionary response'. At the genetic level, correlated response occurs due to pleoitropy or hitchhiking [31]. Under the reasonable assumption that the number of genomic regions under selection increases with the number of traits, the correlated response resulting from multifarious selection will often cause more genomically-widespread divergence than the response caused by strong selection on one trait [1,31,63]. Thus, by sampling through the genome more widely, multifarious selection is more likely to incidentally cause divergence in a few key genomic regions that are particularly important for speciation. Examples of such regions are those affecting mating preference or causing intrinsic hybrid inviability [34 for review of this 'sampling model']. Thus, controlling for total selection strength, some predictions emerge.

1) Strong selection on one or a few traits is better at causing adaptive divergence in the face of gene flow than is multifarious selection. However, because selection on a single trait often causes little correlated response, it will often result in single trait polymorphism rather than speciation.

2) Multifarious selection will sometimes be too weak to strongly overcome gene flow, precluding divergence in the selected traits and any correlated response. However, when multifarious selection does cause divergence, its widespread correlated response may be more effective at driving speciation than the more limited correlated response caused by selection on one or a few traits.

Thus, the probability of speciation under each hypothesis will depend on a balance between total selection strength, the number of traits subject to selection, rates of gene flow, and the probability that the correlated response to selection causes reproductive isolation (Fig. 3b). For example, in the presence of high gene flow between populations, only strong selection on a few traits might be capable of causing the adaptive divergence required to drive ecological speciation. If gene flow is somehow restricted, for example in parapatry, multifarious selection might be strong enough to overcome gene flow, and drive speciation. When gene flow is very low, the correlated response caused by multifarious selection might be very effective at incidentally causing the evolution of reproductive isolation [3,63]. Due to the numerous

1 interacting factors at play, explicit theory is required to make clearer predictions about
2 speciation probabilities under each hypothesis.
3   A final point is that the importance of the two hypotheses might vary among stages of
4 the speciation process. For example, strong selection on one or a few traits may initiate
5 speciation, thereby either causing some reduction in gene flow or the evolution of a genetic
6 polymorphism, either of which in turn allows divergence in others traits that are under (weaker)
7 multifarious selection. In such a scenario, single trait polymorphisms may become converted to
8 speciation. Because most past work has focused on the early stages of ecological speciation,
9 future studies that examine multiple stages are required to avoid a bias towards understanding
10 only the early stages of the process. We now turn again to empirical data.
11
12 **Support for the 'stronger selection' hypothesis**
13   Indirect support for the stronger selection hypothesis stems from reference [7], a study
14 which, in addition to the niche dimensions of diet and habitat, also reports on divergence in one
15 phenotypic trait (body size). In some cases, body size divergence was positively correlated with
16 reproductive isolation, independent from time. Assuming that greater divergence in size arises
17 via stronger divergent selection on size, the results support the stronger selection hypothesis.
18 Similar results stem from positive associations between body size divergence and levels of
19 premating isolation in stickleback fishes [64], body size divergence and levels of intrinsic
20 postzygotic isolation in *Centrarchid* fishes [65], body shape divergence and premating isolation
21 between *Gambusia* fish ecotypes [66], and the magnitude of color pattern shifts in relation to
22 levels of premating isolation in *Heliconius* butterflies [67]. A final example concerns the
23 *Pundamilia* cichlids discussed above, where reproductive isolation is positively related to the
24 degree of divergence in opsin genes [10].
25
26 **Support for the 'multifarious selection' hypothesis**
27   This hypothesis most clearly traces its roots to a review of experimental evolution
28 studies in *Drosophila* that concluded that "Laboratory experiments collectively indicate that
29 multifarious…divergent selection can readily lead to complete reproductive isolation, but that
30 single-factor…divergent selection will typically lead to only incomplete reproductive isolation"
31 [31, p. 1647]. Despite being intuitive, there are almost no tests of this hypothesis in nature,
32 perhaps due to the difficulty of generating the required selection estimates. Nonetheless, a few
33 key systems, such as taxon pairs of herbivorous insects, provide some preliminary information
34 (Box 3). For example, in *Timema* stick insects the degree of reproductive isolation between
35 taxon pairs increases with the number of traits subject to divergent selection [34,68]. Similarly,
36 multifarious selection on diapause life-history traits creates a strong barrier to gene flow
37 between host races of *Rhagoletis* flies [33]. However, these studies did not explicitly isolate an
38 effect of multifarious selection independent from selection strength on individual traits.
39
40 **Integration of different factors affecting speciation**
41   The selective hypotheses reviewed here are refinements of the ecological speciation
42 hypothesis. Nonetheless, such refinements are important, given the abundant unexplained
43 variability in the stage of speciation achieved (Table 1). Similar refinement has been important
44 for understanding the specific role of geographic, genetic, and time-based factors in speciation
45 [2,3,69]. For example, it would be overly crude to classify levels of gene flow during
46 divergence as present versus absent; a migration rate, m, of 0.001 is very different from $m =$
47 0.20, yet both represent non-zero gene flow. Likewise, time since divergence of 1 million years

1   is relatively 'ancient', but cannot be equated with a divergence time of 20 million years.
2   Detailed empirical [2,8,14,69] and theoretical [3,19,58] treatments of the role of time, gene
3   flow, and genetics in the completeness of speciation have increased our understanding of the
4   speciation process. Thus, work on the role of ecological factors in completing speciation is
5   warranted. A particularly unexplored area is how ecological and non-ecological factors interact
6   during speciation. Treating such interactions is beyond the scope of this review, but some
7   limited data does suggest they represent an important direction for further research. For
8   example, in *Rhagoletis* flies, genetic variation for diapause traits under divergent selection in
9   sympatry originated in chromosomal inversions that arose in allopatry, so that selection,
10 geography, and genetic architecture interact [39]. Other examples of such interactions come
11 from Lake Victoria cichlid and stickleback fishes [10,13,48].
12
13 **Conclusions and Future Directions**
14   We outlined a framework for testing the role of ecology in the completeness of
15 speciation, reviewed how speciation can be strongly promoted by extreme or highly
16 dimensional ecological shifts, and outlined two hypotheses for why such shifts drive speciation:
17 the stronger and multifarious selection hypotheses. Each hypothesis has seen some limited and
18 relatively indirect support, and certainly one hypothesis does not appear more strongly
19 supported than the other. Thus, further tests are required to yield a comprehensive
20 understanding of how speciation unfolds from beginning to end.
21   In addition to the directions highlighted throughout the article, we suggest four main
22 avenues of further research. First, the existing data does not robustly identify independent axes
23 of trait divergence and selection, which is required to avoid confounding different hypotheses.
24 Second, we focused on the degree to which speciation reaches completion. Future work could
25 focus on the maintenance of species boundaries after speciation is completed. For example,
26 some types of reproductive barriers, such as intrinsic hybrid inviability, might be less reversible
27 than others, and thus particularly effective at maintaining species boundaries. Third, even in the
28 examples cited here, it is uncertain whether variation in niche divergence is the cause or the
29 consequence of variation in levels of gene flow (i.e. reproductive isolation) [10,13,34].
30 Potential solutions involve comparing allopatric taxa in which levels of gene flow cannot affect
31 niche divergence (because gene flow is absent), measuring selection itself, and using
32 experimental manipulations to infer causality [70]. Fourth, increased sophistication of
33 molecular tools will eventually allow the stronger and multifarious selection hypotheses to be
34 tested at the genetic (i.e. gene) rather than phenotypic (i.e. trait) level. Thus, systems that span a
35 range of divergences, and that can combine genetic and ecological data, hold much promise for
36 testing whether the stronger and multifarious selection hypotheses might represent 'ecological
37 rules' of speciation.
38

3
4 **Glossary**
5
6 **Correlated evolutionary response** – divergence of a trait, which itself might not be under
7 selection, which occurs because it is correlated with another trait that is under divergent
8 selection. Here we use this term primarily to refer to reproductive isolation that evolves as a
9 correlated response to selection on other traits.
10
11 **Divergent natural selection** - selection arising from environmental differences or ecological
12 interactions (e.g. competition) that acts in contrasting directions on two populations (e.g. large
13 body size confers high survival in one environment and low survival in the other) or favors
14 opposite extremes of a trait within a single population (i.e. disruptive selection)
15
16 **Ecological speciation** – a speciation process in which divergent natural selection drives the
17 evolution of reproductive incompatibility (i.e. isolation) between taxa
18
19 **Genetic variance-covariance (G) matrix** – a matrix whose diagonal elements are the additive
20 genetic variances of different traits, and the off-diagonal elements are the additive genetic
21 covariances of different traits
22
23 **Genotypic clustering** – the modality of the distribution of gene frequencies, particularly in
24 sympatry or parapatry, with strong bimodality characterizing the existence of two distinct
25 species
26
27 **Lineage sorting** – related to genotypic clustering, but generally referring to the grouping of
28 taxa in gene genealogies, which can range from polyphyletic through to reciprocally
29 monophyletic relationships
30
31 **'Multifarious selection' hypothesis** – a hypothesis predicting that the completeness of
32 speciation is positively related to the number of genetically independent traits subject to
33 divergent selection
34
35 **Niche dimension** – used here to refer to an ecological axis, such as habitat use or diet
36
37 **Pleoitropy** – Multiple phenotypic effects of a single gene
38
39 Q_{ST} – a metric of the degree of genetic differentiation among populations displayed by
40 quantitative traits, which partitions quantitative genetic variation between versus within
41 populations
42
43 **'Stronger selection' hypothesis** - a hypothesis predicting that the completeness of speciation is
44 positively related to the strength of divergent selection on a given trait
45
46

References

1 Schluter, D. (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.

2 Coyne, J.A. and Orr, H.A. (2004) *Speciation*. Sinauer Associates, Inc. Sunderland, MA.

3 Gavrilets, S. (2004) *Fitness landscapes and the origin of species*. Princeton Univ. Press, Princeton, NJ.

4 Rundle, H. and Nosil, P. (2005) Ecological speciation. *Ecol. Lett.* 8, 336-352.

5 Price, T. (2008) *Speciation in Birds*. Roberts and Company, Greenwood Village, CO.

6 Funk, D.J. (1998) Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. *Evolution* 52, 1744-1759.

7 Funk, D.J. *et al.* (2006) Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proc. Natl. Acad. Sci. USA* 103, 3209-3213.

8 Nosil, P. (2007) Divergent host-plant adaptation and reproductive isolation between ecotypes of *Timema cristinae* walking-sticks. *Am. Nat.* 169, 151-162.

9 Jiggins, C.D. *et al.* (2001) Reproductive isolation caused by colour pattern mimicry. *Nature* 411, 302-305.

10 Seehausen, O. *et al.* Speciation through sensory drive in cichlid fish. *Nature* in press

11 Drès, M. and Mallet, J. (2002) Host races in plant-feeding insects and their importance in sympatric speciation. *Phil. Trans. R. Soc. Lond. B* 357, 471-492.

12 Mallet, J. *et al.* (2007) Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evol. Biol.* 7, 28.

13 Seehausen, O. (2008) The sequence of events along a "speciation transect" in the Lake Victoria cichlid fish *Pundamilia.* pp. xx-xx. In '*Ecology and Speciation*' (R. Butlin, J. Bridle, D. Schluter, eds.) Cambridge University Press.

14 Coyne, J.A. and Orr, H.A. (1989) Patterns of speciation in *Drosophila. Evolution* 43, 362-381.

15 Nosil, P. *et al.* (2005) Perspective: Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* 59, 705-719.

16 Mallet, J. (1995) A species definition for the modern synthesis. *Trends Ecol. Evol.* 10, 294-299.

17 Jiggins, C.D. and Mallet, J. (2000) Bimodal hybrid zones and speciation. *Trends Ecol. Evol.* 15, 250-255.

18 Morjan, C.L. and Rieseberg, L.H. (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Mol. Ecol.* 13, 1341-1356.

19 Endler, J.A. (1977) Geographic Variation, Speciation and Clines. Princeton Univ. Press, Princeton.

20 Dopman, E.B. *et al.* (2005) Consequences of reproductive barriers for genealogical discordance in the European corn borer. *Proc. Natl. Acad. Sci. USA,* 102, 14706-14711.

21 Funk, D.J. and Omland, K.E. (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Ann. Rev. Ecol. Evol. Syst.* 34, 397-423.

22 Wu, C.I. (2001) The genic view of the process of speciation. *J. Evol. Biol.* 14, 851-865.

23 de Queiroz, K. (2005) Ernst Mayr and the modern concept of species. *Proc. Natl. Acad. Sci. USA* 102, 6600-6607.

24 Svensson, E.I. *et al.* (2007) Female polymorphisms, sexual conflict and limits to speciation

processes in animals. *Evol. Ecol.* in press.

25 Rueffler, C. *et al.* (2006) Disruptive selection and then what? *Trends. Ecol. Evol.* 21, 238-245.

26 Crispo, E. *et al.* (2006) The relative influence of natural selection and geography on gene flow in guppies. *Mol. Ecol.* 15, 49-62.

27 Seehausen, O. *et al.* (1997) Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science*, 277, 1808-1811.

28 Seehausen, O. et al. (2008) Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Mol. Ecol.* 17, 30-44.

29 Taylor, E.B. et al. (2006) Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Mol. Ecol.* 15, 343-355.

30 Hutchinson, G.E. (1957) Concluding remarks - *Cold Spring Harbor Symposia on Quantitative Biology* 22, 415-427.

31 Rice, W.R. and Hostert, E.E. (1993) Laboratory experiments in speciation: what have we learned in 40 years? *Evolution* 47, 1637-1653.

32 Via, S. (2001) Sympatric speciation in animals: the ugly duckling grows up. *Trends Ecol. Evol.* 16, 381-390.

33 Dambroski, H.R. and Feder, J.L. (2007) Host plant and latitude-related diapause variation in *Rhagoletis pomonella*: a test for multifaceted life history adaptation on different stages of diapause development. *J. Evol. Biol.* 20, 2101–2112.

34 Nosil, P. and Harmon, L.J. (2008) Niche dimensionality and ecological speciation. In 'Ecology and Speciation' (R. Butlin, J. Bridle, D. Schluter, eds.), pp. xx-xx, Cambridge University Press.

35 Bradshaw, H.D. and Schemske, D.W. (2003) Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers *Nature* 426, 176-178.

36 Hawthorne, D.J. and Via, S. (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* 412, 904-907.

37 Noor, M.A.F. *et al.* (2001) Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. USA* 98, 12084-12088.

38 Rieseberg, L.H. (2001) Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* 16, 351-358.

39 Feder, J.L. *et al.* (2003) Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. *Proc. Natl. Acad. Sci. USA* 100, 10314-10319.

40 Felsenstein, J. (1981) Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* 35, 124-38.

41 Ortiz-Barrientos, D. and Noor, M.A.F. (2005) Evidence for a one-allele assortative mating locus. *Science* 310, 1467-1467.

42 Barrett, R.D.H. and Schluter, D. (2008) Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23, 38-44.

43 Rieseberg, L.H. *et al.* (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* 301, 1211–1216.

44 Seehausen, O. (2004) Hybridization and adaptive radiation. *Trends Ecol. Evol.* 19, 198-207.

45 Hendry, A.P. and Taylor, E.B. (2004) How much of the variation in adaptive divergence can be explained by gene flow? An evaluation using lake-stream stickleback pairs. *Evolution* 58, 2319-2331.

46 Bolnick, D.A and Nosil, P. (2007) Natural selection in populations subject to a migration load. *Evolution* 61, 2229-22243.

47 Seehausen, O. et al. (1999) Colour polymorphism and sex ratio distortion in a cichlid fish as an incipient stage in sympatric speciation by sexual selection. *Ecol. Lett.* 2, 367-378.

48 Colosimo P.F. *et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307, 1928-1933.

49 Hey, J. (2006) Recent advances in assessing gene flow between diverging populations and species. *Curr. Opin. Gen. Develop.* 16, 592-596.

50 Kremer, A. *et al.* (1997) Multilocus and multitrait measures of differentiation for gene markers and phenotypic traits. *Genetics* 145, 1229-1241.

51 Harmon, L.J. *et al.* (2005) Convergence and the multidimensional niche. *Evolution* 59, 409-421.

52 Hine, E. and Blows, M. (2006) The effective dimensionality of the genetic variance-covariance matrix. *Genetics* 173, 1135-1144.

53 Chenoweth, S.F. and Blows, M. (2008) Q_{ST} meets the G matrix: the dimensionality of adaptive divergence in correlated quantitative traits. *Evolution* 62, 1437-1449.

54 Lande, R. and Arnold, S.J. (1983) The measurement of selection on correlated characters. *Evolution* 37, 1210-1226.

55 Crespi, B.J. *et al.* (2004) Evolution of Ecological and Behavioural Diversity: Australian Acacia Thrips as Model Organisms. Australian Biological Resources Study and CSIRO Entomology: Canberra, Australia.

56 Bush, G.L. (1969) Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera, Tephritidae). *Evolution* 23, 237-251.

57 Futuyma, D.J. *et al.* (1995) Genetic constraints on macroevolution: the evolution of host affiliation in the leaf beetle genus *Ophraella. Evolution,* 49, 797-809.

58 Gavrilets, S. and Vose, A. (2005) Dynamic patterns of adaptive radiation. *Proc. Natl. Acad. Sci. USA* 102, 18040-18045.

59 Maan M.E. *et al.* (2008) Parasite-mediated sexual selection and species divergence in Lake Victoria cichlid fish. *Biol. J. Linn. Soc.* 94, 53-60.

60 Lande, R. (1979) Quantitative genetic-analysis of multivariate evolution, applied to brain – body size allometry. *Evolution* 33, 402-416.

61 Arnold, S.J. (1992) Constraints on phenotypic evolution. *Am. Nat.* 140, S85-S107.

62 Mallet J. (2006) What does *Drosophila* genetics tell us about speciation? *Trends Ecol. Evol.* 21, 186-193.

63 Johnson, N.A. and Porter, A.H. (2000) Rapid speciation via parallel, directional selection on regulatory genetic pathways. *J. Theor. Biol.* 205, 527-542.

64 McKinnon, J.S. *et al.* (2004) Evidence for ecology's role in speciation. *Nature* 429, 294-298.

65 Bolnick, D.I. *et al.* (2006) Body size divergence promotes post-zygotic reproductive isolation in centrachids. *Evol. Ecol. Res.* 8, 903-913.

66 Langerhans, R.B. *et al.* (2007) Ecological speciation in *Gambusia* fishes. *Evolution,* 2056-2074.

67 Jiggins, C.D. *et al.* (2004) Mimicry and the evolution of premating isolation in *Heliconius melpomene* Linnaeus. *J. Evol. Biol.* 17, 680-691.

68 Nosil, P. and Sandoval, C.P. (2008) Ecological niche dimensionality and the evolutionary diversification of stick insects. *PLoS ONE* 3: e1907.

69. Ramsey, J. *et al.* (2003) Components of reproductive isolation between *Mimulus lewisii* and

M. cardinalis (Scrophulariaceae). *Evolution* 57, 1520-1534.

70 Räsänen, K. and Hendry, A.P. (2008) Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecol. Letters* 11, 624-626.

Table 1. Examples of variation in completeness of speciation, which illustrate the continuous nature of variation in divergence.

Level of variability[a]	Measure of divergence[b]	Study system(s)	Result	Reference
Among disparate systems	RI (expt.)	numerous (20 different groups)	Total reproductive isolation varied among systems from 0.08 to 1.00	[15]
Among disparate systems	RI (gene flow)	numerous (1284 studies reviewed)	F_{ST} varied among taxon pairs from 0.00 to 1.00	[18]
Among disparate systems	Lin. Sort.	numerous (2319 animal species)	Phylogenetic grouping between closely-related species ranged from polyphyly to reciprocal monophyly (with 23% of taxa being para- or polyphyletic)	[21]
Among disparate systems	Gen. Cluster.	numerous (17 hybrid zones in different taxa)	Modality of the distribution of gene frequencies in hybrid zones ranged from unimodal, to flat, to strongly bimodal	[17]
Within and among disparate systems	RI (expt.)	numerous (hundreds of taxa from eight disparate groups)	Individual forms of reproductive isolation varied from 0.00 to 1.00, both within and among study systems	[2,7]
Among populations within systems	RI (gene flow)	*Gasterosteus* sticklebacks	Species pairs exhibit variability in the degree of reproductive isolation; one previously strongly reproductively isolated species pair has collapsed back into a single interbreeding population	[29]
Among populations within systems	RI (expt.)	*Timema* walking-stick insects	Populations on different host plant species vary in total reproductive isolation according to exposure to homogenizing gene flow	[8]
Among populations within systems	RI (expt. and gene flow)	*Pundamilia* cichlids	Phenotypes adapted to opposite ends of ambient light gradients vary in RI and gene flow, depending on steepness of gradient	[10,13]
Among loci within systems	Lin. Sort.	*Ostrinia nubilalis* Corn borer strains	Genealogies for five gene regions are discordant, and only one molecular marker (one that potentially affects reproductive isolation) exhibited evidence for pheromone strain exclusivity	[20]
Among reproductive barriers within a	RI (expt.)	*Mimulus* monkeyflowers	Estimates of the strength of individual barriers to gene flow, among nine different reproductive barriers, ranged from 0.00 to 0.99	[69]

system				

[a] Various measures of speciation are considered (e.g. reproductive isolation, genotypic clustering, lineage sorting), at various levels of variability (e.g. among study systems, within study systems, etc.). The set of examples was chosen to span a wide range of criteria and taxa.

[b] RI (expt.) = level of reproductive isolation inferred from experimental data. RI (gene flow) = level of reproductive isolation inferred from molecular differentiation in sympatry or parapatry. Gen. cluster = degree of genotypic clustering. Lin. Sort. = lineage sorting.

Figure 1. The continuous nature of divergence during speciation. Divergence during speciation can vary quantitatively, for numerous types of differentiation. Thus, different means of quantifying divergence can be used to measure arbitrary 'stages' of speciation, representing stages from the initiation through to the completion of the speciation process (when two populations are depicted, one is shown in blue and the other in orange). For example, reproductive isolation can vary from absent through to complete. Likewise, the distribution of gene frequencies in individuals sampled from two populations, depicted here as genotypic clustering, can vary from unimodal through to strongly bimodal. The extent of lineage sorting can vary from weak to strong. Finally, the steepness of geographic or ecological clines in gene frequency can vary, with the latter stages of speciation being characterized by steep or stepped clines.
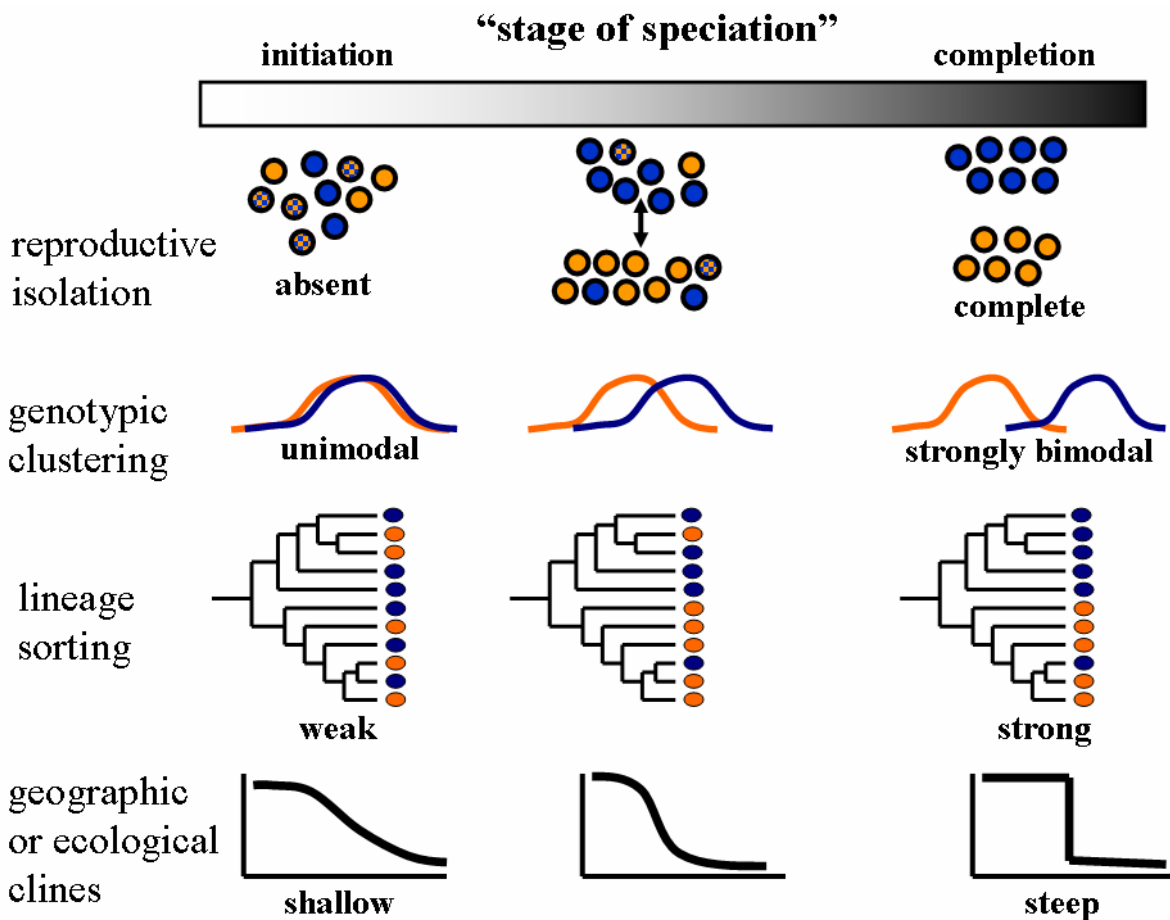
Figure 2. Time-based, geographic, and genetic factors affecting the completeness of speciation a) Time-based factors. Levels of premating isolation between taxon pairs of *Drosophila* increase with genetic distance, a proxy for time since divergence. However, much variation is unexplained by genetic distance. Thus, within a narrow window of genetic distance, sexual isolation can vary from absent to complete (highlighted by red oval). Data from [14]. b) Geographic factors. *Timema cristinae* walking-stick insects exhibit two color-pattern morphs; an unstriped morph is more cryptic on the host plant *Ceanothus* and a striped morph is more cryptic on *Adenostoma*. The degree of between-host gene flow is a function of the geographic arrangement of populations (particularly relative population sizes in parapatry). The frequency of the maladaptive morph within populations is correlated with the degree of gene flow into the population from populations of the alternative host. Thus, gene flow constrains the adaptive divergence that drives ecological speciation. Data from [46]. c) Genetic factors. Speciation is promoted by the pleiotropic effects of genes under selection on reproductive isolation (e.g. *Mimulus*, photo credit: D. Schemske) [35], physical linkage of genes under selection and those conferring reproductive isolation (e.g. *Acyrthosiphon*, photo credit: S. Via)[36], perhaps facilitated by chromosomal inversions (e.g. *Rhagoletis*, photo credit: A. Forbes)[39], and the fixation of the same allele in both of two diverging populations (e.g. *Drosophila*, photo credit: D. Ortiz-Barrientos)[41].

a) Time since divergence and speciation
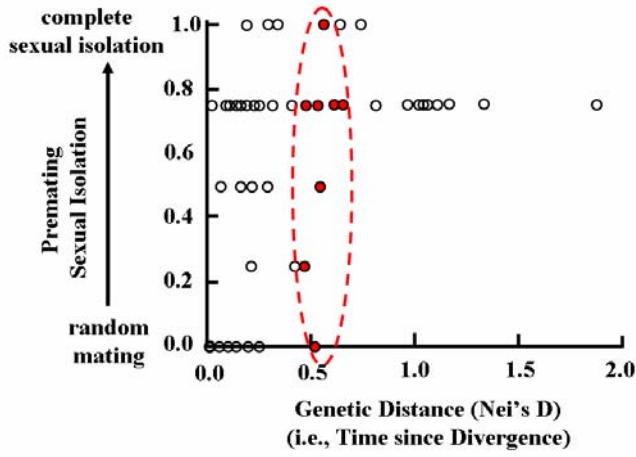
b) Geographic factors and speciation
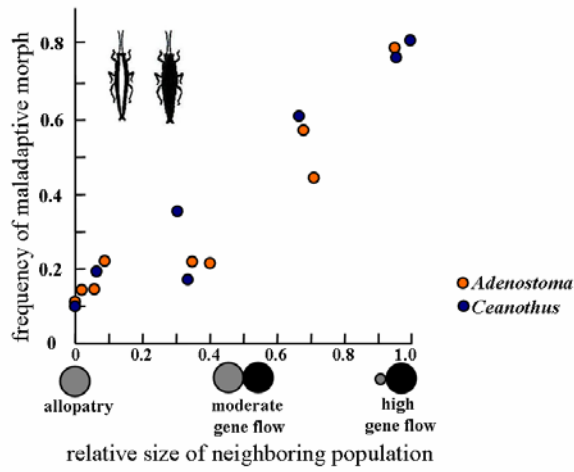
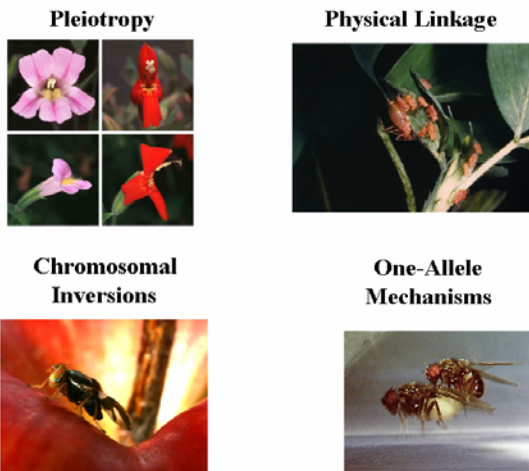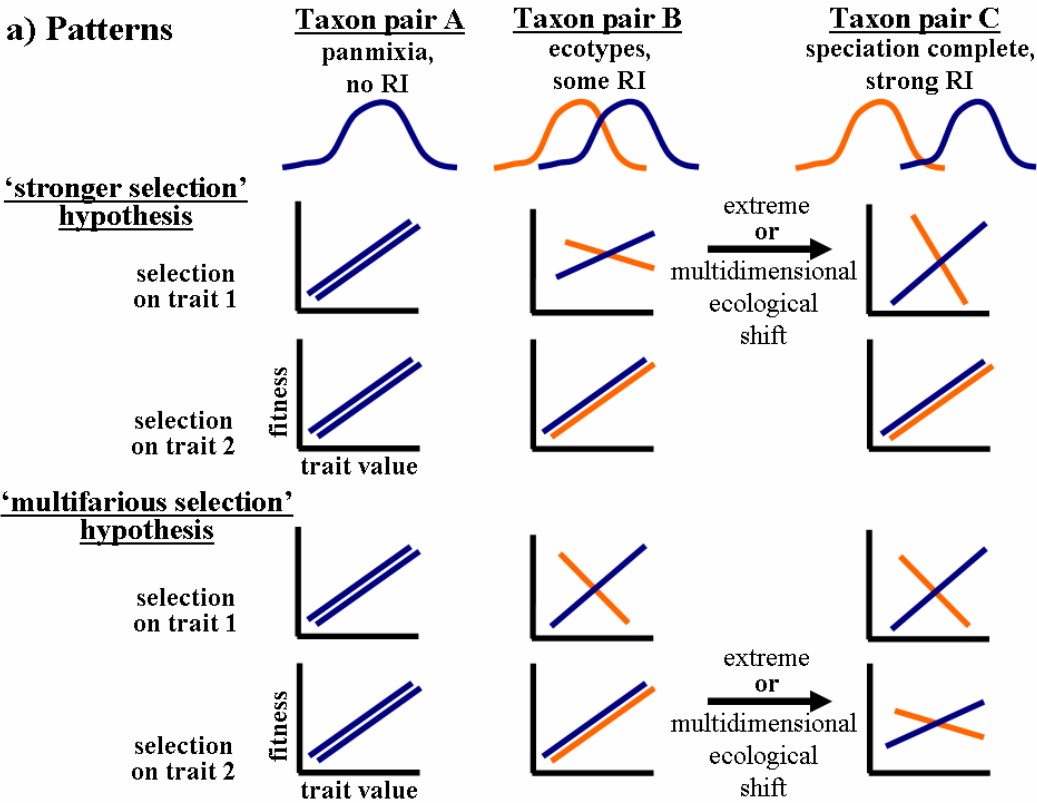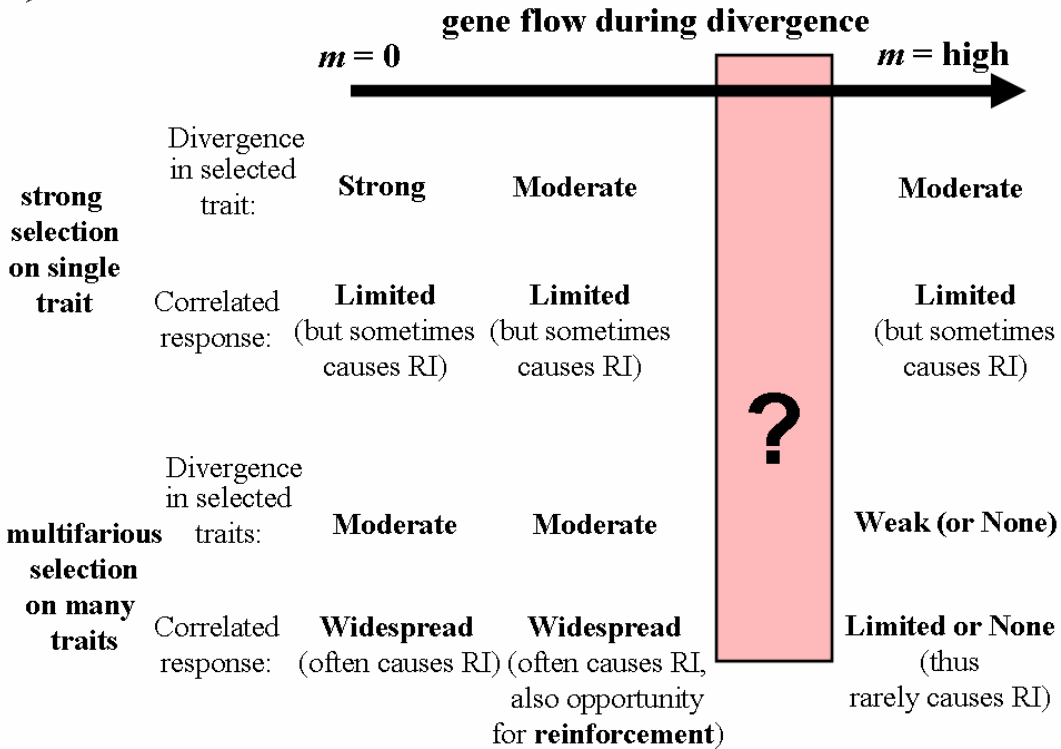**c) Genetic factors promoting speciation**

Figure 3. Patterns and predictions of the 'stronger selection' and 'multifarious selection' hypotheses. a) The three taxon pairs depicted vary in which stage of the speciation process has been achieved (RI = reproductive isolation, the distributions represent stages of the speciation process, as outlined in Fig. 1). Selection might act on two phenotypic traits (e.g. morphology and physiology). Graphs represent fitness functions, where the x-axes represent trait values and the y-axes represent fitness. Crossing fitness functions are indicative of divergent selection, with steeper lines indicating stronger divergent selection. The critical change predicting the completion of speciation under each hypothesis is labeled by an arrow. Note that both extreme shifts along one niche dimension and multidimensional niche shifts can cause either stronger selection on a given single trait ('stronger selection'), selection on a greater number of traits ('multifarious selection'), or both. The case depicted here might be extended to a multiple regression framework, as described in Box 1. b) The probability of speciation under each hypothesis depends on a balance between the total strength of selection, the number of traits subject to selection (which affects per trait selection coefficients), rates of gene flow (m, which also affect the opportunity for reinforcement), and the probability that divergent selection incidentally affects reproductive isolation (i.e. causes a 'correlated response'). Shown here is the expected divergence in selected traits and the nature of any correlated response (RI = reproductive isolation). Multifarious selection on many traits is more likely to result in reproductive isolation as a correlated response. However, speciation under strong selection on a single trait becomes more likely when gene flow is too high to allow divergence except under strong selection. The actual point at which multifarious selection loses efficacy in causing divergence in the face of gene flow is wide-ranging (denoted by the question mark), being dependent on a balance between the factors noted above.

a) Patterns

Taxon pair A
panmixia,
no RI

Taxon pair B
ecotypes,
some RI

Taxon pair C
speciation complete,
strong RI

'stronger selection'
hypothesis

selection
on trait 1

extreme
or
multidimensional
ecological
shift

fitness

selection
on trait 2

trait value

'multifarious selection'
hypothesis

selection
on trait 1

fitness

selection
on trait 2

extreme
or
multidimensional
ecological
shift

trait value

b) Predictions

gene flow during divergence

$m = 0$

$m = $ high

				?	
strong selection on single trait	Divergence in selected trait:	**Strong**	**Moderate**		**Moderate**
	Correlated response:	**Limited** (but sometimes causes RI)	**Limited** (but sometimes causes RI)		**Limited** (but sometimes causes RI)
multifarious selection on many traits	Divergence in selected traits:	**Moderate**	**Moderate**		**Weak (or None)**
	Correlated response:	**Widespread** (often causes RI)	**Widespread** (often causes RI, also opportunity for **reinforcement**)		**Limited or None** (thus rarely causes RI)

Box 1. A framework for testing ecological speciation, with supporting examples.
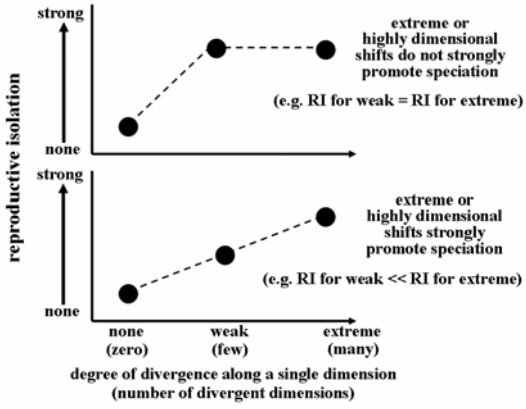
We describe here a framework for isolating the role of various types of ecological divergence in the completeness of speciation. For example, reproductive isolation might increase with the magnitude of divergence in any one niche dimension or with the number of divergent dimensions (Fig. Ia). An empirical example where the completeness of speciation increases with the number of divergent niche dimensions involves the incipient species pair *Pundamilia pundamilia* and *Pundamilia nyererei* (Fig. Ib, blue and red males, respectively). In this example, the completeness of speciation was inferred using neutral genetic differentiation at microsatellite loci, experimental data on mating preferences, and the distribution of male nuptial colouration (blue left, red right, and three intermediate classes) [10,13,59].

When a large number of taxon pairs are available for analysis, time since population separation might be controlled for in a multiple regression framework, where time is inferred using molecular genetic distance. For example, Fig. Ic depicts a hypothetical scenario where reproductive isolation increases with both genetic distance and various types of ecological divergence. Fig Id shows an analysis where extreme shifts in habitat between angiosperm taxon pairs promote speciation, independent from time (modified from [7] and reprinted with permission of the National Academy of Sciences USA). The approach of controlling for time using genetic data works best for allopatric species pairs. For hybridizing taxa, the degree of reproductive isolation confounds the estimate of divergence time because it directly affects genetic distance. A potential solution is to apply coalescent based techniques to estimate divergence time independent from gene flow, but the efficacy of these methods needs consideration [49].
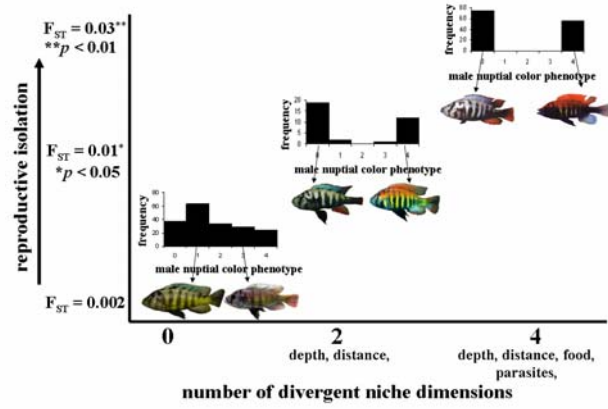
To test which specific ecological factors (e.g. Fig Ic) determine the completeness of speciation, one must avoid confounding the extent of divergence along one niche dimension with the number of divergent dimensions, and likewise avoid confounding the strength of divergent selection on a given trait with the number of traits subject to divergent selection. Thus, the first and second measures of divergence should be independent from one another, as should the third and fourth. Various multivariate statistics can generate independent dimensions of niche and trait divergence [1,34,,50-54]. For example, a multivariate analog of Q_{ST} can control for correlations among traits when calculating quantitative trait divergence [53]. When it comes to selection itself, multiple regression procedures for quantifying selection on one trait independent from other measured traits are well-established (i.e. selection gradients)[54]. Thus, the strength of divergent selection on one trait might be identified, and the number of traits under divergent selection inferred.

Figure I. Isolating a role for different types of ecological divergence in speciation. a) A hypothetical scenario where the completeness of speciation varies according to the nature of an ecological shift. b) An empirical example where the completeness of speciation in *Pundamilia* cichlids is positively related to the number of ecological dimensions that taxon pairs differ in [data from 10,13,59]. c) A hypothetical example where the completeness of speciation varies as a function of both time and ecological divergence. d) An empirical example where residual reproductive isolation (effects of time statistically removed) increases with divergence in habitat use [data from 7].
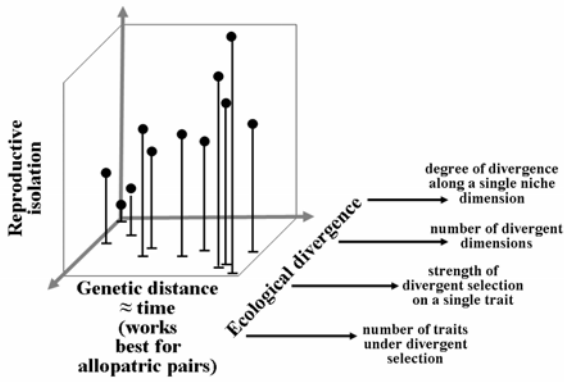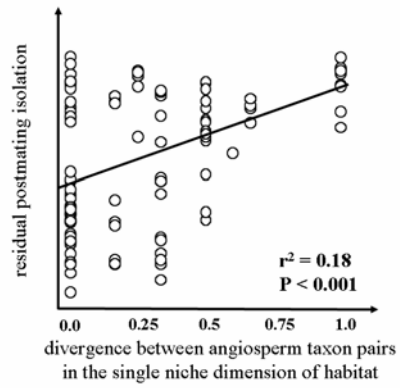
a) Hypothetical case : three taxon pairs

extreme or highly dimensional shifts do not strongly promote speciation

(e.g. RI for weak = RI for extreme)

extreme or highly dimensional shifts strongly promote speciation

(e.g. RI for weak << RI for extreme)

reproductive isolation

strong / none / strong / none

none (zero) — weak (few) — extreme (many)

degree of divergence along a single dimension (number of divergent dimensions)

b) Empirical example : three taxon pairs

$F_{ST} = 0.03$** **$p < 0.01$

$F_{ST} = 0.01$* *$p < 0.05$

$F_{ST} = 0.002$

reproductive isolation

frequency

male nuptial color phenotype

number of divergent niche dimensions

0 — 2 depth, distance, — 4 depth, distance, food, parasites,

c) Hypothetical case: many taxon pairs

Reproductive isolation

Genetic distance ≈ time (works best for allopatric pairs)

Ecological divergence

degree of divergence along a single niche dimension

number of divergent dimensions

strength of divergent selection on a single trait

number of traits under divergent selection

d) Empirical example : many taxon pairs

residual postmating isolation

$r^2 = 0.18$
$P < 0.001$

0.0 — 0.25 — 0.5 — 0.75 — 1.0

divergence between angiosperm taxon pairs in the single niche dimension of habitat

Box 2. Testing the 'multifarious selection' hypothesis using independent axes of trait divergence.

The approach uses trait divergence as a surrogate for divergent selection, with more highly multifarious trait divergence representing more highly multifarious selection. Trait divergence is quantified between multiple pairs of taxa for multiple traits, and then principle components analysis is used to determine the amount of the variance in trait divergence explained by each of x independent directions. Each direction is a composite trait, made up from a linear combination of the original traits. The dimensionality or 'evenness' of each matrix can be calculated from its eigenvalues using Levene's index [1, pp. 220-221]:
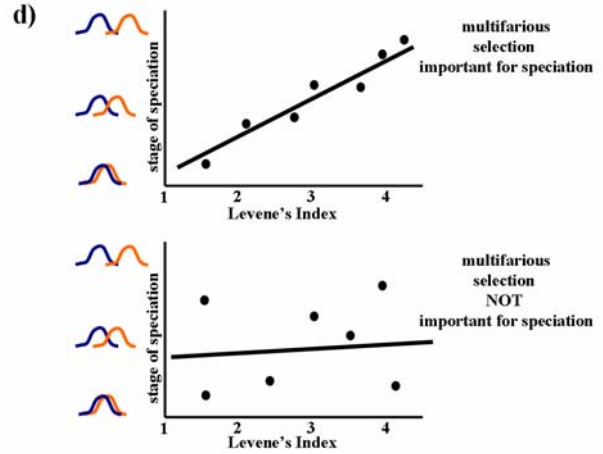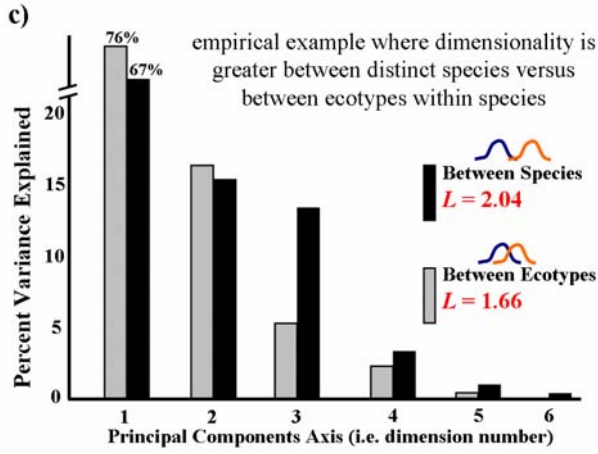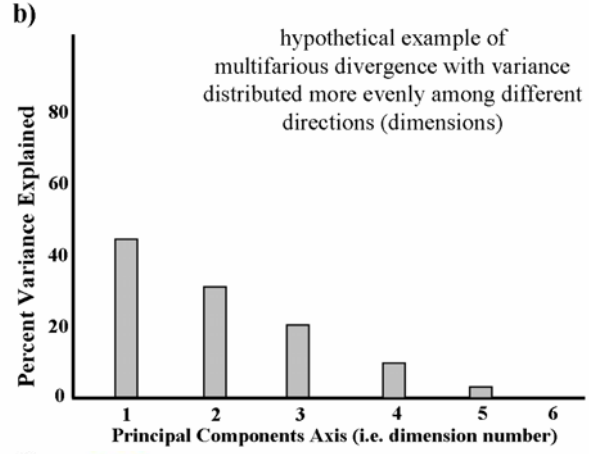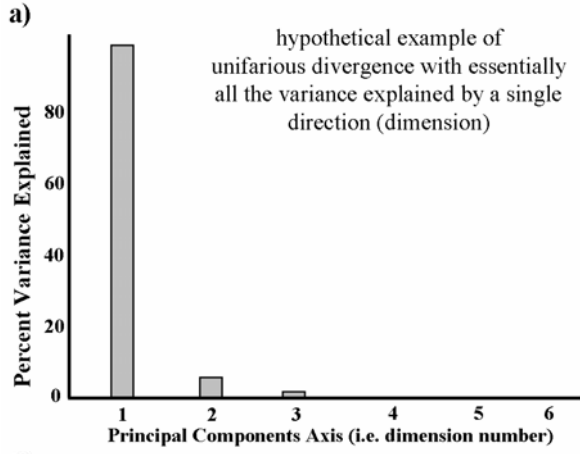
$$L = 1 / \sum p_i^2$$

where p_i is the proportion of total variance accounted for by eigenvector i. $L = 1$ if all variance is in the first direction, and L = the number of eigenvectors (i.e. traits) if variance is equitably distributed among directions. L therefore measures 'dimensionality', with higher values indicative of more multifarious divergence (Fig. Ia versus Ib below), and thus presumably more multifarious selection.

Ideally, this analysis is conducted for different classes of taxon pairs which vary in the stage of speciation achieved. Support for the multifarious selection hypothesis arises if there is a positive relationship between the stage of speciation achieved and the unevenness of the distribution. For example, in *Timema* walking-stick insects, trait divergence was more multifarious for comparisons between distinct species relative to divergence between ecotypes within species ($L = 2.04$ versus 1.66, respectively) (Fig. Ic) [34]. Ideally, many stages of divergence would be examined to test for a quantitative association, with increasing importance attached to multifarious selection as the slope and strength of the positive relationship increases (Fig. Id).

A strong assumption of this approach is that the classes of taxon pairs being compared do not differ strongly in the genetic variance-covariance (G) matrix, because trait divergence is a function of both selection and the G matrix [1,34,60]. In practice, this means that phenotypic data can only be a useful surrogate of selection when comparing very closely related taxa with similar demographic histories, or when variability in the G matrix is measured and controlled for.

Figure I. Quantifying the dimensionality of trait divergence. a) and b) depict hypothetical examples of relatively unifarious and highly multifarious divergence, respectively. c) An empirical example in *Timema* stick insects where morphological divergence between distinct species is more multifarious than divergence between ecotypes within species. d) Applying the framework to numerous sets of taxon pairs that vary in the completeness of speciation.

a) hypothetical example of unifarious divergence with essentially all the variance explained by a single direction (dimension)

b) hypothetical example of multifarious divergence with variance distributed more evenly among different directions (dimensions)

c) empirical example where dimensionality is greater between distinct species versus between ecotypes within species

Between Species
$L = 2.04$

Between Ecotypes
$L = 1.66$

d) multifarious selection important for speciation

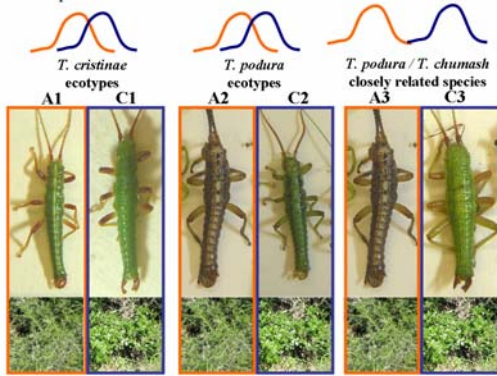multifarious selection NOT important for speciation

Box 3. Preliminary support for the multifarious selection hypothesis in herbivorous insects.
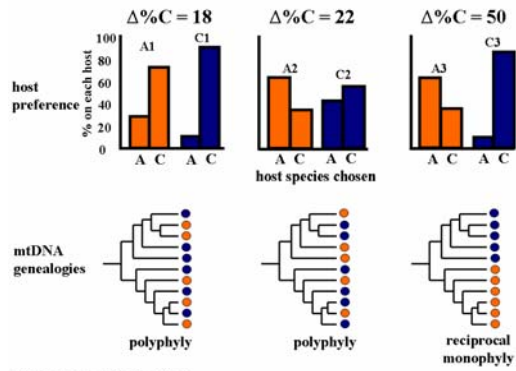
In herbivorous insects, divergent selection between populations on different host plants might act on many different types of traits, for example on cryptic colouration used to evade visual predation or on physiology used to detoxify plant chemicals. Selection was estimated on both these traits in three taxon pairs of *Timema* walking-stick insects (Figure Ia). These pairs vary in their degree of reproductive isolation and the completeness of speciation, inferred using experimental estimates of host-plant preference, levels of mtDNA differentiation, and taxonomic status (Figure Ib). The taxon pairs also differed in the number of traits subject to divergent selection (Figure Ic, y-axis measures fitness, with crossing fitness functions indicative of divergent selection). The results revealed that strong divergent selection on the single trait of cryptic colouration is associated with host ecotype formation and intermediate levels of reproductive isolation. In contrast, stronger reproductive isolation between a species pair was associated with divergent selection on both cryptic colouration and physiology, rather than on cryptic colouration alone [8,34,68]. The results are consistent with the multifarious selection hypothesis, but further replication is required for a robust test. Another potential example comes from *Rhagoletis* flies, where diapause life history traits create a strong ecological barrier to gene flow. Different diapause traits, such as initial diapause depth, timing of diapause termination and post-diapause development rate, are genetically uncoupled and are each subject to divergent selection such that the barrier to gene flow is created by multifarious selection [33].

Figure I. In *Timema* stick insects, the completeness of speciation increases with the number of traits subject to divergent selection. a) The two ecotype pairs and the species pair studied. A1 and C1 refer to ecotypes of *T. cristinae* (in all cases A refers to use of *Adenostoma* as a host, and C refers to use of *Ceanothus*)). A2 and C2 refer to ecotypes of *T. podura*. A3 and C3 refer to the species pair *T. podura* and *T. chumash*, respectively. b) Data on the completeness of speciation, here the degree of divergence in host preference and the extent of lineage sorting in mtDNA genealogies (here schematic for simplicity). Δ%C refers to the difference between each taxon pair in the percent of individuals choosing *Ceanothus* over *Adenostoma* in host preference trials. c) Estimates of selection on two traits, cryptic colour pattern and physiology. Modified from [68] and reprinted with permission of the Public Library of Science.

a) Taxon pairs

T. cristinae ecotypes

T. podura ecotypes

T. podura / T. chumash closely related species

A1 C1 A2 C2 A3 C3

b) Host preferences and mtDNA gene genealogies

$\Delta\%C = 18$ $\Delta\%C = 22$ $\Delta\%C = 50$

host preference

% on each host

host species chosen

A C A C A C A C A C A C

mtDNA genealogies

polyphyly   polyphyly   reciprocal monophyly

c) Nature of selection

crypsis

survival

A C

background plant

physiology

mean fecundity (# eggs)

to A to C

host transplanted to