

## Genomic divergence during speciation: causes and consequences

Patrik Nosil and Jeffrey L. Feder

*Phil. Trans. R. Soc. B* 2012 **367**, 332-342

doi: 10.1098/rstb.2011.0263

---

### References

[This article cites 72 articles, 15 of which can be accessed free](#)

<http://rstb.royalsocietypublishing.org/content/367/1587/332.full.html#ref-list-1>

[Article cited in:](#)

<http://rstb.royalsocietypublishing.org/content/367/1587/332.full.html#related-urls>

### Subject collections

Articles on similar topics can be found in the following collections

[evolution](#) (461 articles)

[genetics](#) (5 articles)

[genomics](#) (11 articles)

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

---

To subscribe to *Phil. Trans. R. Soc. B* go to: <http://rstb.royalsocietypublishing.org/subscriptions>

---

*Introduction*

# Genomic divergence during speciation: causes and consequences

Patrik Nosil<sup>1,2,\*</sup> and Jeffrey L. Feder<sup>2,3</sup>

<sup>1</sup>*Department of Ecology and Evolutionary Biology, University of Boulder, Boulder, CO 80309, USA*

<sup>2</sup>*Institute for Advanced Study, Wissenschaftskolleg, Berlin 14193, Germany*

<sup>3</sup>*Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA*

Speciation is a fundamental process responsible for the diversity of life. Progress has been made in detecting individual ‘speciation genes’ that cause reproductive isolation. In contrast, until recently, less attention has been given to genome-wide patterns of divergence during speciation. Thus, major questions remain concerning how individual speciation genes are arrayed within the genome, and how this affects speciation. This theme issue is dedicated to exploring this genomic perspective of speciation. Given recent sequencing and computational advances that now allow genomic analyses in most organisms, the goal is to help move the field towards a more integrative approach. This issue draws upon empirical studies in plants and animals, and theoretical work, to review and further document patterns of genomic divergence. In turn, these studies begin to disentangle the role that different processes, such as natural selection, gene flow and recombination rate, play in generating observed patterns. These factors are considered in the context of how genomes diverge as speciation unfolds, from beginning to end. The collective results point to how experimental work is now required, in conjunction with theory and sequencing studies, to move the field from descriptive studies of patterns of divergence towards a predictive framework that tackles the causes and consequences of genome-wide patterns.

**Keywords:** ecology; gene flow; genomics; natural selection; recombination; speciation continuum

## 1. INTRODUCTION

Speciation is a fundamental process responsible for creating the diversity of life on the Earth. In general, speciation involves the splitting of one reproductive community (or genotypic cluster) of organisms into two [1–3]. Conceptualizing speciation in this manner leads to a clear research programme: to understand speciation, one must understand how genetically based barriers to gene flow (i.e. reproductive isolation) evolve between populations. Much progress has been made on discerning the importance of different factors and traits (including ecological adaptation) in generating reproductive isolation [4–8]. In addition, individual ‘speciation genes’ contributing to reproductive isolation, and, in particular, genes causing intrinsic post-zygotic inviability and sterility in hybrids, have been identified [1,9–13].

In contrast, we lack a good understanding of how these speciation genes are embedded and arrayed within the genome, and thus of how genomes evolve during population divergence. Thus, major questions about the genomic architecture of speciation and how it facilitates or impedes further divergence remain. Generally speaking, our empirical and theoretical understanding of speciation is still largely

dominated by what Ernst Mayr described as ‘beanbag thinking’ focused on one or a few individual genes [3,14]. There are exceptions where interactions among multiple loci have been considered, e.g. the ‘snow-ball’ effect for the accumulation of the number of post-zygotic incompatibilities through time [15–17]. However, in general, detailed studies of the genetics of speciation have been restricted to examining one or a few loci [13]. These approaches have worked well enough so far because, until recently, empirical studies were limited to such a gene-centred focus. But as represented by the articles in this issue, we are now capable of rapidly scanning large portions of the genome of both model and non-model organisms for differentiation [18–22]. Consequently, the field of evolutionary genetics is beginning to move away from studies focused on a few genes to those that tackle genome-wide patterns. This theme issue is dedicated to exploring this genomic perspective of speciation. Given recent high-throughput sequencing and computational advances that allow genome-wide analyses, our goal is for this theme issue to be timely in helping move the field towards a more integrative approach for understanding speciation.

Here, in this introductory article, we discuss the major elements, questions and challenges leading us to a more integrative understanding of the genomics of speciation. Our narrative highlights how the 12 papers in this theme issue contribute to building and achieving this genomic synthesis. We begin by laying

\* Author for correspondence ([patrik.nosil@colorado.edu](mailto:patrik.nosil@colorado.edu)).

One contribution of 13 to a Theme Issue ‘Patterns and processes of genomic divergence during speciation’.

Table 1. Definitions of some newer terms being used to describe patterns and processes of genomic divergence.

term	definition
genomic island of divergence	a region of the genome, of any size, whose divergence exceeds neutral background expectations
divergence hitchhiking	a term used to describe the process by which physical linkage to a divergently selected gene(s) increases genomic divergence for regions adjacent to a selected site on a chromosome. The process is based on a site under divergent selection between populations creating a localized region of reduced gene flow around it in the genome, enhancing the potential to maintain or accumulate differentiation (both neutral and selected) at linked sites
genome hitchhiking	a term used to describe the process by which genetic divergence across the genome is facilitated, even for loci unlinked to those under selection, by a global reduction in average genome-wide gene flow that divergent selection causes
outlier loci	exceptionally differentiated gene regions, usually those that are differentiated beyond background levels/neutral expectations. Outlier loci are generally interpreted as being affected by divergent selection and/or causing reproductive isolation, but this has rarely been directly tested

out the metaphoric foundation on which the study of genomic architecture is currently based, introducing the concept of ‘genomic islands of divergence’ and the processes generating such divergence (see table 1 for definitions). In this section, we articulate the importance of gene flow and the theoretical principle of ‘selection/recombination antagonism’ that has greatly shaped considerations of which types of genetic architectures are needed for reproductive isolation to evolve in the face of gene flow [23]. These questions, their counterpoints and their resolution form the basis for the papers of Feder *et al.* [24] and Via [25] in the issue.

We next describe broad-scale patterns of genomic divergence observed in nature. These patterns quantified by ‘genome scans’ of numerous loci provide the empirical observations on which our current understanding of genomic divergence is based. Numerous studies in this issue present new empirical data or summarize the literature on how genetic differentiation is distributed throughout the genome [25–33]. The authors relate these patterns to the island metaphor to try and draw inferences concerning the significance of genome architecture and the processes facilitating speciation.

Although the description of patterns of genomic divergence seems straightforward, the papers in this issue highlight the difficulties we still face in accurately resolving genome-wide patterns of genetic differentiation and interpreting their meaning. For example, although highly divergent regions tend to be interpreted as being affected by divergent selection or harbouring genes causing reproductive isolation, other factors such as retention of ancestral polymorphism or recombination rate variation could contribute to generating similar patterns [34,35]. There has been little explicit study to date of the degree to which divergence and reproductive isolation are actually related. The study in this theme issue on genomic isolation in hybrids begins to tackle this issue [36].

Additionally, studies tend to be done on different scales and take different analytical approaches to identify exceptionally differentiated ‘outlier loci’. Thus, the statistical and methodological power to differentiate outlier loci from neutral background

differentiation varies among studies, and without critical details concerning the natural history and biology of population divergence (e.g. geographical context, time of divergence, rate of migration, strength of divergent selection) it is difficult to place individual studies within a broader comparative framework. Thus, tests for general relationships between genome structure and speciation will require more accurate information on where particular systems reside along the continuum from freely interbreeding populations, to newly formed and partially isolated populations or ecotypes, to largely reproductively isolated species. Another consideration is whether speciation was initiated in the face of gene flow or whether there was an initial period of allopatry in which differences accumulated in the absence of gene flow [37–40]. In the latter case, genome architectures maintained following secondary contact might differ somewhat from those created *de novo* in the face of gene flow. Consequently, the processes of divergence and genome hitchhiking that we discuss below may differ in their relative importance between primary and secondary modes of speciation with gene flow.

After considering studies of empirical patterns, we then turn to theoretical expectations for how processes such as divergence and genome hitchhiking promote speciation. Much of our conceptual understanding of how genome structure should facilitate speciation has been dominated by verbal arguments and metaphors based on limited theory. Thus, while basic principles may be known, our understanding of the relative importance of different processes in generating genomic divergence, and how this architecture may feedback to influence further divergence during the speciation process, is surprisingly unresolved. The theoretical papers by Feder *et al.* [24], Guerrero *et al.* [41] and Gompert *et al.* [38] work to fill this void, by examining the processes affecting the establishment of new mutations during genomic divergence, expected patterns of genetic divergence in chromosomal inversions, and the patterns of genomic isolation created by different genetic architectures, respectively. Development of such theoretical expectations allows clearer interpretations of specific empirical patterns.

Finally, we examine some of the new challenges associated with analysing and interpreting the great influx of raw genomic data we are now capable of generating. We conclude by discussing how all these aspects of genomic study might be tied together in future work, emphasizing the importance of now integrating manipulative experiments with observational studies to correctly interpret empirical patterns of genomic divergence.

## 2. THE METAPHOR OF GENOMIC ISLANDS OF DIVERGENCE

To aid thinking about divergence in the genome, evolutionary biologists have developed the metaphor of 'genomic islands of divergence' [42], where a genomic island is any gene region, be it a single nucleotide or an entire chromosome, which exhibits significantly greater differentiation than expected under neutrality [18]. The metaphor thus draws parallels between genetic differentiation observed along a chromosome and the topography of oceanic islands and the contiguous sea floor to which they are connected. Following this metaphor, sea level represents the threshold above which observed differentiation is significantly greater than expected by neutral evolution alone. Thus, an island is composed of both directly selected and tightly linked (potentially neutral) loci. Factors such as physical proximity between selected and other loci, rates of recombination, and strength of selection then each affect the height and the size of genomic islands. Under this metaphor, the few genes under or physically linked to loci experiencing strong divergent selection can diverge, whereas gene flow will homogenize the remainder of the genome (or insufficient time for genetic drift will preclude divergence of regions that are not divergently selected), resulting in isolated genomic islands (but see [34,35]).

An important consideration for expected patterns of divergence is the geographical context of speciation, specifically, whether and when gene flow accompanied the divergence process. Gene flow is important, because strictly allopatric divergence, be it via selection or genetic drift, proceeds unfettered by the homogenizing effects of migration [17]. Thus, the extent of genetic linkage and recombination among genes relative to the strength of selection is not a major constraint on divergence in allopatry. Although much has been learned about specific reproductive barriers and speciation genes from studying allopatric taxa [1], it is difficult to ascribe any special significance to a particular genetic change or genetic architecture in such systems. Genomic architecture is less relevant to allopatric speciation, as divergence across the genome is inevitable. In contrast, physical linkage relationships and recombination rates among genes, along with levels of gene flow and the strength of selection, are critical considerations with respect to speciation with gene flow, where gene flow constantly introduces the wrong combination of genes into a local population and recombination breaks-up associations between genes under selection and those causing reproductive isolation [17,43,44]. As described in the classic work by Felsenstein [23], there is an antagonism between

selection and recombination during divergence with gene flow. Selection must overcome this antagonism for reproductive isolation and widespread genomic divergence to evolve in the face of gene flow.

A key take-home message is that details concerning the natural history of speciation, which are often lacking for many systems, are important. For example, current rates of migration between taxa can affect the course of change to come, but may differ from those that existed when present patterns of genomic differentiation were generated. To correctly interpret empirical patterns, it is thus critical not only to conduct exhaustive genetic surveys, but also to resolve the history of population divergence. For example, the clustering of divergently selected genes in the genome might not have been created owing to new mutations sequentially establishing around already diverged sites in the face of gene flow. Instead, it could be the result of the enhanced retention of such an architecture following secondary contact and introgression. The empirical papers in the current issue underscore the importance of resolving this potential difference.

## 3. DIVERGENCE HITCHHIKING AND GENOME HITCHHIKING

How might genomic islands form in the face of gene flow, and then grow in size? The verbal theory of divergence hitchhiking posits that physical linkage to divergently selected loci generates a mechanism by which genomic islands form and can be (or grow to be) of relatively large size, and by which speciation in the face of gene flow may be easier than previously thought [45,46]. As outlined in the contribution to this issue by Via [25], the premise is that divergent selection reduces interbreeding between populations in different habitats [45–47]. This reduces interpopulation recombination, and even if recombination occurs, selection reduces the frequency of immigrant alleles in advanced generation hybrids [17]. This localized reduction in effective gene flow at or near genes subject to divergent selection might allow large regions of genetic differentiation to build up in the genome around the few loci subject to divergent selection. The idea rests on the assumption that a site under divergent selection will create a relatively large window of reduced gene flow around it, enhancing the potential to accumulate differentiation (both neutral and selected) at linked sites.

As an alternative to divergence hitchhiking on a few loci, selection acting on many loci distributed throughout the genome could reduce gene flow to drive speciation [48–50]. This process also produces variable patterns of genomic divergence, owing to differences among loci in selection intensities, linkage relationships and recombination rates. However, in the case of selection on many loci, genomic regions displaying weaker differentiation may not all be neutrally evolving, but rather represent regions more weakly affected by selection. In this case, many loci are diverged beyond neutral, 'sea-level' expectations such that genomes differ by many 'archipelagoes' or even 'continents' of divergence (or at very least,

numerous small and somewhat interconnected islands). We stress that the island versus continent views represent ends of a continuum, rather than mutually exclusive hypotheses (figure 1). For example, continents can be conceptualized as large islands with variable topography (e.g. mountain tops and lowland continental plains all above neutral sea level). In our own contribution to this issue [24], we define the term ‘genome hitchhiking’ to describe the process by which genetic divergence across the genome is facilitated, even for loci unlinked to those under selection, by the reductions in average genome-wide gene flow that selection causes. This process, unlike divergence hitchhiking, does not invoke a role for physical linkage, and can facilitate divergence across the genome. These considerations lay the foundation for thinking about patterns of genomic divergence, and the processes causing them.

#### 4. PATTERNS OF GENOMIC DIVERGENCE

Numerous questions remain about the empirical patterns of genomic divergence during speciation. For example, how numerous, large and genomically clustered are regions of genomic differentiation? As highlighted in this issue by Renaut *et al.* [33] and Via [25], the answer to this question will depend, in part, on how regions of differentiation are delimited. Are a few adjacent regions of differentiation along a chromosome considered one large or several small and isolated ‘islands’ of divergence? Below, we address these questions in as standardized a manner as possible, but the contributions in the theme issue clearly define a need for further work on how patterns of genomic differentiation are best quantified, and on the degree to which they represent reproductive isolation (see earlier works [36,38,51–53]).

##### (a) *How numerous?*

How many genomic regions differentiate during speciation? The available evidence suggests that the answer to this question is variable, but that multiple regions tend to be differentiated, and each to a different degree. For example, Strasburg *et al.* [26] reviewed numerous studies of the genomic basis of plant speciation and concluded that multiple regions tend to differentiate both closely and distantly related plant populations. This result is consistent with an older review in which mostly animal taxa were represented [18]. Likewise, as discussed in the article by Hahn *et al.* [29], initial work on molecular forms of *Anopheles* mosquitoes detected multiple, yet few (i.e. three), regions of differentiation [42], but subsequent finer scale sequencing detected numerous other regions of differentiation [54].

Although genome scans sometimes report only very few regions of differentiation, genome scans that have poor genomic coverage and that are conducted without complimentary selection experiments can be biased towards supporting a view that divergence occurs in only a few regions. This is because, inevitably, only the most diverged regions will be identified as statistical outliers. Other loci affected by selection, but more weakly, will go unnoticed and be considered part of the

mostly ‘undifferentiated’ and neutral genome. In short, although empirical genome scans have usefully identified candidate regions strongly affected by divergent selection and are a good starting point to characterize patterns of genomic architecture, they cannot readily detect selection on less-differentiated regions. Thus, direct experimental measurements of selection on the genome are required to detect both weak and strong selection, and determine the fraction of the overall genome that differentiates during speciation.

Such an experimental test of the genomic island scenario was conducted by Michel *et al.* [49] in the apple and hawthorn host races of *Rhagoletis pomonella*, a model for sympatric ecological speciation initiated with gene flow. Contrary to expectations, they reported numerous lines of evidence for widespread divergence and selection throughout the *Rhagoletis* genome, with the majority of loci displaying latitudinal clines, associations with an ecologically important trait (adult eclosion time), within-generation responses to selection in a manipulative over-wintering experiment and host differences in nature despite substantial gene flow (4–6% per generation).

The results, coupled with linkage disequilibrium (LD) analyses, provide field-based and experimental evidence that divergence was driven by selection on numerous independent genomic regions, suggesting that ‘continents’ of multiple differentiated loci, rather than isolated islands of divergence, can characterize even the early stages of speciation. Their results also illustrate continental topography. The divergence observed throughout the *Rhagoletis* genome was clearly more accentuated in some regions, such as those harbouring chromosomal inversions. A final point is that standard outlier analyses in this same study were consistent with the genomic island hypothesis: only two independent gene regions were detected as statistical outliers. Thus, experimental data and biological information on gene flow in nature were critical for detecting weaker yet widespread divergence across the genome. Until further such studies emerge, it will be impossible to know if genomic continents are the exception, or the norm.

The contribution to this theme issue by Nadeau *et al.* [32] on genomic divergence among hybridizing *Heliconius* butterflies further illustrates the points exemplified by the *Rhagoletis* study. Nadeau *et al.* [32] applied a cutting-edge genomic capture methodology to a non-model organism for the first time to sequence through two genomic regions known to harbour genes affecting divergent wing-colour patterns. These colours also play a role in speciation via contributing to selection against immigrants and hybrids and divergent mating preferences. Even within just these two genomic regions, they find multiple peaks of differentiation, but with some regions nonetheless more differentiated than others.

##### (b) *How large?*

How large are regions of divergence in the genome? Relatedly, how does genetic divergence decay away from a selected site, and does it decay to zero? There is much evidence that independent regions of differentiation in the genome can be small. This claim is

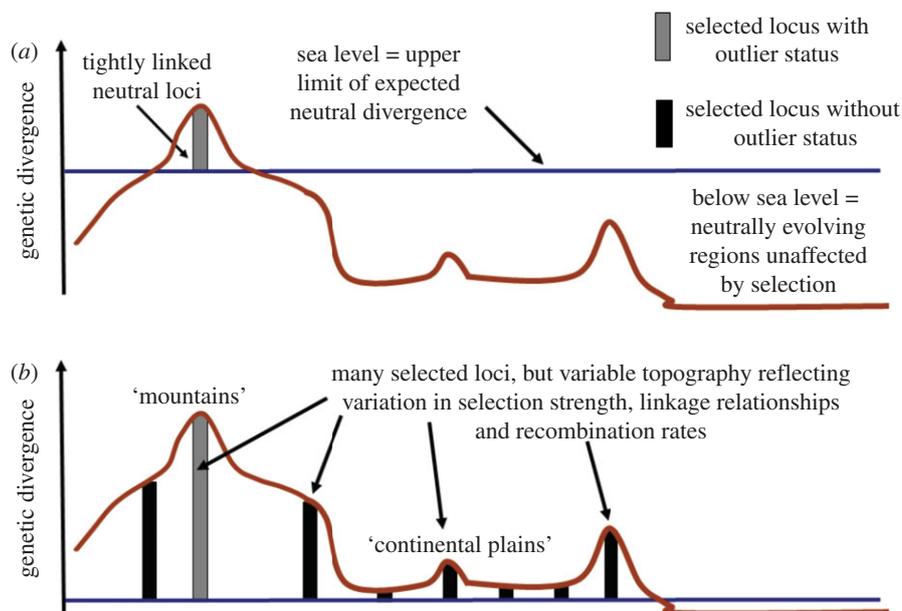


Figure 1. Schematic of the (a) island versus (b) continent views of genomic divergence. These views represent ends of a continuum, rather than mutually exclusive hypotheses. For example, ‘continents’ of divergence can be conceptualized as very large islands with variable topography. Outlier status refers to whether a locus would exhibit statistical evidence for unusually high levels of genetic differentiation in an observational genome scan. See text for details. Reproduced with permission from Michel *et al.* [49], National Academy of Sciences USA.

supported by general reviews of genomic divergence in animals [18] and plants [26]. Likewise, the genomic capture study in *Heliconius* mentioned above reports that regions of differentiation tend to span only a few hundred kilobases [32]. However, there are some strong exceptions to this trend, where regions of differentiation, measured for example by the distance between an outlier locus and the nearest quantitative trait locus (QTL) for an ecologically relevant trait, appear large. Examples from this theme issue stem from pea aphid host races [25] and lake whitefish ecotypes [33]. Likewise, in the first quantification of genome-wide levels of LD in the stickleback genome, Hohenlohe *et al.* [30] report relatively large LD blocks in the genomes of both freshwater and oceanic populations. We consider some of the causes of this variability in the subsequent section on processes generating genomic divergence.

### (c) *How dispersed?*

Are regions of divergence concentrated on just a few genomic regions, or widely spread across chromosomes? Reviews of genome scans in both plants and animals indicate that regions of differentiation often map to different chromosomes [18,26]. Although regions of differentiation might sometimes cluster [25,27,31], for example in regions of low recombination such as within chromosomal inversions, divergence in other (e.g. collinear) regions nonetheless occurs. This is exemplified by the contribution of McGaugh & Noor [31] to this theme issue, where it was documented that regions of particularly strong differentiation between *Drosophila* species often lie within inversions, but collinear regions are nonetheless differentiated. Similar results are reported in mice and rabbits in the contribution by Nachman & Payseur

[27], but in the context of recombination rate variation among different collinear genomic regions.

### (d) *What types of genomic regions?*

What types of genes or gene regions tend to be differentiated? In particular, to what extent does genomic divergence involve coding versus regulatory regions [55]? Much more data are needed, but it is clear that both can be involved [56]. In their review of the genomics of cichlid speciation, Fan *et al.* [28] also present new transcriptomic data suggesting that selection on coding regions contributes to genomic differentiation. The whitefish study by Renaut *et al.* [33] reports that mapped expression differences between dwarf normal ecotypes localize to a few ‘expression QTL hotspots’, but that these hotspots are not associated with elevated genetic divergence between natural populations (whereas, in contrast, regular ‘phenotypic’ QTLs are). Much more work is needed to determine which types of genomic regions differentiate during speciation. Specifically, further studies explicitly comparing coding and non-coding regions [32] are needed.

## 5. CAUSES OF PATTERNS OF GENOMIC DIVERGENCE

The results above demonstrate that patterns of genomic divergence can be highly variable, both among gene regions within study systems and among different study systems. This variability raises the obvious question of which processes are generating the observed patterns. For example, what are the roles of selection and recombination in generating genomic divergence? In terms of selection itself, how important is the strength of selection acting directly on gene regions, relative to the effects of selection causing hitchhiking of linked regions and overall reductions in gene flow?

Other questions concern the importance of reduced recombination. To what extent can reduced recombination facilitate genetic divergence, in a manner analogous to that played by strong selection? In addition, there are issues concerning variation in the evolutionary process itself. For example, even in genomic regions containing divergently selected loci, there can be pronounced differences in the levels of associated neutral differentiation, as probably seen by Nadeau *et al.* [32] in *Heliconius*. This can be owing to differences in the age of segregating neutral variants, to the stochastic nature of the drift process itself, and to vagaries in the recombination history of neutral sites with the specific targets of selection. Further work on how this variation be summarized when attempting to describe and compare the topology of genomic divergence is required.

An important message here is that we have a good metaphorical conception of the major population genetic processes causing patterns of genomic differentiation. However, we lack theoretical details of the relative importance of different processes as speciation unfolds over time, particularly when gene flow levels vary temporally owing to biogeography (but see [49]). Metaphors may be insufficient, and vagaries in experimental methodology and statistical analysis too great, to currently distinguish among competing factors shaping genome structure during speciation. Caution is therefore urged in attaching too much significance at this time to verbal associations made between observed empirical patterns and underlying process. These issues raise the need to strengthen the linkage between data and theory to further our understanding of the genomics of speciation, as exemplified by the theme issue contribution by Gompert *et al.* [38].

#### (a) *The roles of direct selection, divergence hitchhiking and genome hitchhiking*

The theoretical contribution to this theme issue by Feder *et al.* [24] illustrates how the relative importance of the different selective processes for generating genomic divergence can begin to be disentangled. Specifically, Feder *et al.* [24] report new analytical and simulation results which estimate the probability that new beneficial mutations will establish and differentiate populations diverging in the face of gene flow when the new mutation: (i) is the first mutation to arise in a completely undifferentiated genome, (ii) arises in physical linkage to a locus already diverged via selection, and (iii) arises unlinked to any selected loci, but within a genome that has some diverged loci. This sequential approach allows the partitioning of how various mechanisms aid the establishment of new mutations. For example, the effect on a new mutation of arising in linkage to a diverged region represents the effect of ‘divergence hitchhiking’. In contrast, the effect of arising unlinked to selected regions but in an already diverged genome represents the effects of genome-wide average gene flow reductions caused by selection (i.e. ‘genome hitchhiking’).

Feder *et al.* [24] find that the strength of selection acting directly on a new mutation is an important

predictor of establishment, with both forms of hitchhiking having smaller effects in comparison. This result is consistent with past theoretical results focused on the maintenance, rather than the origin, of differentiation [17,57,58] and population-genetic theory considering single populations [59,60]. Nonetheless, divergence hitchhiking aided mutation establishment under certain conditions, in particular, when selection coefficients favouring the new mutation were less than the migration rate (see also [25]). Genome hitchhiking also sometimes promoted mutation establishment, particularly if multiple loci have already diverged prior to the emergence of the new mutation. A key message here is that a rapid transition (in terms of the number of differentiated genes) may occur from early phases of divergence where the selective benefits of new mutations themselves are the primary factor affecting their establishment to a stage where multiple (but not necessarily numerous) loci generate the potential for widespread genomic divergence via genome hitchhiking [24]. Divergence hitchhiking can fortuitously aid this transition, but may not be necessarily vital for it.

Empirical studies have just begun to try and home in on the relative importance of these different processes for generating genomic divergence. Elucidating the specific targets of divergent selection in the first place, and distinguishing them from neutral regions that simply hitchhike to divergence, is a difficult task. The focused sequencing of gene regions implicated in colour-pattern diversification of *Heliconius* butterflies demonstrates how steps in the direction towards finding the specific targets of selection can be made [32]. Once the specific targets of selection are determined, work can then focus on the consequences for neutral regions linked and unlinked to the selected ones.

#### (b) *Recombination rate variation*

As described above, the antagonism between selection and recombination can impede genomic divergence. It therefore follows that factors that reduce recombination, such as chromosomal inversions, can facilitate genomic divergence. Inversions may therefore help create more elevated oceanic islands and broader, more mountainous continents of divergence between taxa. Similar arguments can be made for other features of the genome that result in reduced recombination, such as proximity of genes under selection to centromeres.

The basic premise for a role of inversions in speciation is that they reduce introgression across the regions of the genome they encompass and protect favourable genotypic combinations within them from being broken up by recombination [61–63]. Essentially, the favourable genes within the inversions are more favourable together in their natal habitat than they would be individually and less favourable in the habitats of other populations than they would be alone. Hence, gene flow is reduced. Moreover, in addition to preserving blocks of adaptively diverged genes, inversions also provide larger targets in the genome for divergence hitchhiking to work on; by suppressing recombination, inversions enlarge the area of the

genome in which new favourable mutations might arise linked to already diverged genes.

A number of articles in this theme issue focus on reduced recombination and genomic divergence. For example, the articles by McGaugh & Noor [31] and Nachman & Payseur [27] describe elevated divergence in regions of reduced recombination (inversions and proximity to centromeres, respectively), relative to more freely recombining areas. Nonetheless, both studies report divergence in collinear regions as well. These results are qualitatively consistent with some past work, such as the study by Strasburg *et al.* [64] examining divergence between hybridizing sunflower species, who reported that genetic divergence is not accentuated within inversions, except perhaps near chromosomal breakpoints, where recombination is particularly reduced. Moreover, widespread adaptive divergence in collinear regions is being increasingly documented [18,26,49,54,64]. Indeed, theoretical work by Feder & Nosil [65] showed that there is no reason for the vast majority of loci contributing to reproductive isolation to reside in inversions; they should also be commonly found in collinear regions. Thus, it may be that factors that reduce recombination are not essential for genomic divergence, but they can facilitate it, particularly if recombination is strongly reduced.

A key remaining and difficult task is thus separating the relative roles of strength of selection, rates of gene flow and recombination rates in generating a particular pattern of genomic divergence. This point is exemplified in the empirical patterns described in this issue by Hahn *et al.* [29] and the new theoretical results on expected coalescent patterns for neutral loci within chromosomal inversions reported by Guerrero *et al.* [41]. Rather than casting despair, this simply highlights a clear avenue for further work.

### (c) *Other remaining questions*

A number of other questions concerning genomic divergence during speciation remain unanswered. For example, what are the relative roles of new mutations versus pre-existing standing genetic variation in speciation [66]? Likewise, what geographical arrangement of populations tends to most strongly affect genomic divergence? As noted above, genomic divergence in allopatry can proceed relatively easily such that divergence in many regions of the genome can occur. Interestingly, such a pattern of multiple and relatively small islands of divergence was reported in this theme issue in two systems which appear to vary widely in levels of gene flow between them [29,32]. This highlights the need for further work on how geography and gene flow affect patterns of genomic divergence.

Indeed, recent theory demonstrates how there are some instances where gene flow might facilitate, rather than constrain, certain types of genomic changes. For example, Kirkpatrick & Barton [67] demonstrated that new inversions that originate in sympatric populations exchanging genes that fortuitously happen to trap a combination of genes all adapted to one habitat versus another can be selectively favoured over collinear arrangements, because

the fit genotypic combinations within the inversions are not broken up by recombination. This process can result in inversions rising to high frequency and differentiating populations.

Feder *et al.* [37] extended this work by showing how such adaptive spread of inversions is facilitated if the inversions arise in allopatry such that they are the most likely to contain the perfect complement of locally adapted alleles. Such inversions might be maintained at low frequency in allopatric populations, but then subsequently rise to high frequency when gene flow ensues upon secondary contact. Such a 'mixed mode' of geographical divergence allows for the establishment of inversions under a wider range of conditions than pure sympatric divergence, where, for example, ongoing gene flow and recombination makes it difficult for new inversions to capture the perfect complement of locally adapted alleles. Such a mixed mode of divergence could potentially explain the results reported in this issue by McGaugh & Noor [31], where chromosomal inversions appear to have originated prior to speciation between two closely related *Drosophila* species, and then may have spread to high frequency upon secondary contact and gene flow. Additionally, the role of effective population sizes can have important impacts on patterns of genomic divergence [41,58], and this topic warrants more focused empirical study. Finally, traditional topics such as gene duplication and polyploidization require attention in future work as well.

## 6. CONSEQUENCES OF GENOMIC DIVERGENCE

Even once patterns of genomic divergence are well described, and the processes causing the observed patterns inferred, a major question remains: what are the consequences of genomic divergence for speciation? If selection and divergence are concentrated on just a few genomic regions, will this more effectively overcome gene flow to drive speciation? Is selection on many genomic regions more likely to incidentally cause reproductive isolation (i.e. divergence in mate preference, hybrid dysfunction, reductions in neutral gene flow) than selection on a few regions? Such questions are central to our understanding of speciation, but have only begun to be addressed [38]. In some cases, it is known that certain divergent genomic regions harbour genes affecting reproductive isolation. For example, the inverted chromosomal regions described by McGaugh & Noor [31] contain genes causing intrinsic dysfunctions in hybrids between species of *Drosophila*. In addition, regions containing loci involved in divergent adaptation, which thus likely cause extrinsic selection against migrants and hybrids, have been described in several systems, including those in this issue [25,32,33]. Nonetheless, even in such examples, it remains unknown how the specific patterns of divergence and genetic architecture observed truly promoted or constrained the evolution of reproductive isolation. Further work more directly connecting patterns of genomic divergence to reproductive isolation is required [38]. Reaching true causality on this front will probably require the integration of ecological and functional genomic studies.

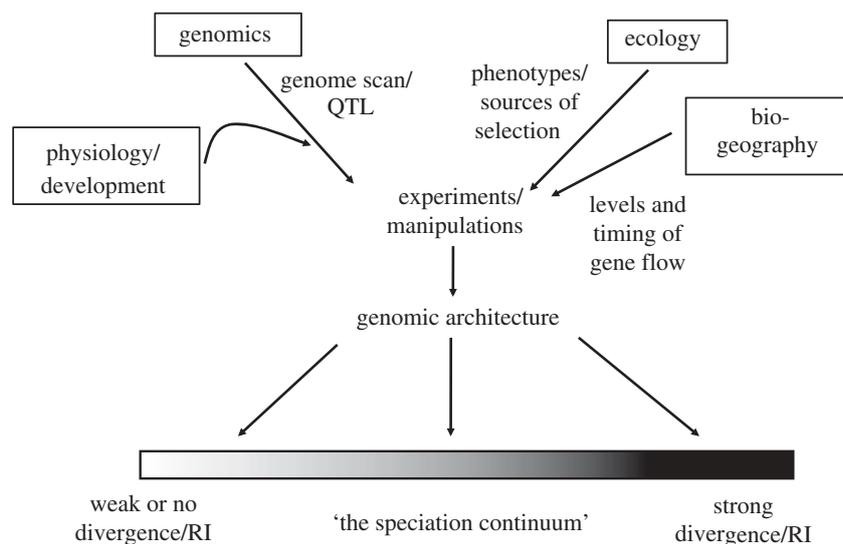


Figure 2. A diagrammatic depiction of how integrative approaches might yield an understanding of the ‘natural history of the genome’, and of how speciation and genomic divergence unfolds over time. RI, reproductive isolation (note that further work on the relationship between divergence and RI is sorely needed) [38].

## 7. ‘GROWING PAINS’: TECHNICAL AND COMPUTATIONAL CHALLENGES

The recent increases in our understanding of genomic divergence during speciation have not come without challenges. For example, although costs of sequencing have decreased rapidly, and will continue to do so, the amount of data that can be obtained, and then analysed, are nonetheless finite. Thus, initial explorations of genomic divergence using cutting edge methods have (necessarily) relied on small sample sizes [31,32]. These studies have been important in paving the way for larger and more integrative work in the future, for example, studies based on larger sampling schemes and which consider changes in both gene expression and nucleotide divergence [56]. Additionally, more flexible and powerful analytical approaches are required. The contribution by Gompert *et al.* [38] to this theme issue provides a step in this direction, describing a method for examining the genomic architecture of isolation in species hybrids. Finally, most theory of speciation to date has focused on one or a few loci. Further theoretical work examining the causes and consequences of genome-wide divergence are also needed, as represented by other contributions in this issue [24,41]. As progress in our ability to deal with genomic data continues to increase, it will be important to further develop conceptual and theoretical frameworks for understanding such data.

## 8. CONCLUSIONS: TOWARDS A NATURAL HISTORY OF THE GENOME

... these forms may still be only ... varieties; but we have only to suppose the steps of modification to be more numerous or greater in amount, to convert these forms into species ... thus species are multiplied. ([68], p. 120)

Speciation is often an extended and quantitative process, during which reproductive isolation and

genomic divergence builds up [1,17,44,50,69]. Ultimately, we would like to know how these processes unfold, and thus how speciation proceeds from beginning to end. Indeed, different points in this continuum of divergence could involve very different processes [24]. For example, speciation initiated in the face of gene flow may often begin via divergence in the few specific gene regions directly subject to divergent selection. This period may then transition to a second phase where gene flow is reduced in localized regions of the genome surrounding selected sites, and divergence hitchhiking may then act to facilitate differentiation of regions physically linked to those under selection under certain circumstances. As further loci diverge, and perhaps new mutations come to differentiate populations, effective gene flow then gets further reduced across the genome. As this proceeds further, loci diverge, and a transition to widespread genomic divergence, facilitated by genome hitchhiking, occurs. The end result then is strong reproductive isolation and widespread genetic divergence. In addition, these different ‘phases’ of the speciation continuum might differ when divergence evolves between populations in sympatry, allopatry or a mixture of these modes [37]. Thus, it may be critical to discern what aspects of genome structure originated in allopatry versus in the face of gene flow.

We can see glimpses of these phases now, by comparing the results from different taxa lying at different points in this speciation continuum. For example, this theme issue describes results from some systems that have proceeded very far in the speciation process (or even completed it), such as the molecular forms of *Anopheles* mosquitoes [29], the host races of pea aphids [25] and *Drosophila* species pairs [31]. In other systems considered in the issue, gene flow appears higher, with speciation having proceeded less far, such as in races of *Heliconius* butterflies [32] and stickleback and whitefish [30,33]. Although making comparisons among the results of these disparate systems is a starting point, it is somewhat

like comparing apples and oranges. What is required now are detailed studies of genomic divergence between very closely related taxa (e.g. population pairs within species that vary strongly in their degree of reproductive isolation or different ecotype and species pairs within a single genus) that span the speciation continuum and that have well-characterized natural and biogeographic histories. Such work is increasing at the phenotypic level [50,69–74], but has yet to be fully applied to the genomic level. Nonetheless, the few cases that have examined genomic divergence at different points in the speciation continuum, such as the study of *Heliconius* in this theme issue, support the model above where direct selection, divergence hitchhiking and genome hitchhiking are differentially important during different phases of the speciation process [32].

Ideally, genomic studies will be increasingly conducted within an experimental framework that directly tests the processes driving and constraining genomic divergence. Some cutting-edge experiments of this type have now been conducted in the laboratory using microbes [75–77], but these do not address reproductive isolation or speciation in natural populations. Future experimental work in natural populations will probably allow us to measure more directly how selection acts on the genome (note that selection within a generation might be measured even in long-lived organisms where evolution between generations cannot be readily addressed). In turn, this will allow us to reconstruct how genomic divergence unfolds across the speciation continuum in single study systems. We can then compare study systems to search for generalities. If such work is conducted using an integration of ecological and molecular genomic approaches, it will probably yield a new understanding of the ‘natural history of the genome’, and of its relevance for understanding the origins of diversity (figure 2). Although much progress remains to be made, it is often clear what needs to be done, and the tools and expertise to make progress exist.

We thank all the authors for their involvement in the Theme Issue and reviewers of the articles for improving the quality of the contributions.

## REFERENCES

- Coyne, J. A. & Orr, H. A. 2004 *Speciation*. Sunderland, MA: Sinauer Associates.
- Mallet, J. 1995 A species definition for the modern synthesis. *Trends Ecol Evol.* **10**, 294–299. (doi:10.1016/0169-5347(95)90031-4)
- Mayr, E. 1963 *Animal species and evolution*. Harvard, MA: Harvard University Press.
- Schluter, D. 2001 Ecology and the origin of species. *Trends Ecol Evol.* **16**, 372–380. (doi:10.1016/S0169-5347(01)02198-X)
- Schluter, D. 2009 Evidence for ecological speciation and its alternative. *Science* **323**, 737–741. (doi:10.1126/science.1160006)
- Rundle, H. D. & Nosil, P. 2005 Ecological speciation. *Ecol. Lett.* **8**, 336–352. (doi:10.1111/j.1461-0248.2004.00715.x)
- Feder, J. L., Xie, X. F., Rull, J., Velez, S., Forbes, A., Leung, B., Dambroski, H., Filchak, K. E. & Aluja, M. 2005 Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in *Rhagoletis*. *Proc. Natl Acad. Sci. USA* **102**, 6573–6580. (doi:10.1073/pnas.0502099102)
- 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. (doi:10.2307/2411347)
- Presgraves, D. C. 2007 Speciation genetics: epistasis, conflict and the origin of species. *Curr. Biol.* **17**, R125–R127. (doi:10.1016/j.cub.2006.12.030)
- Wu, C. I. & Ting, C. T. 2004 Genes and speciation. *Nat. Rev. Genet.* **5**, 114–122. (doi:10.1038/nrg1269)
- Orr, H. A. 2005 The genetic basis of reproductive isolation: insights from *Drosophila*. *Proc. Natl Acad. Sci. USA* **102**, 6522–6526. (doi:10.1073/pnas.0501893102)
- Rieseberg, L. H. & Blackman, B. K. 2010 Speciation genes in plants. *Ann. Bot.* **106**, 439–455. (doi:10.1093/aob/mcq126)
- Nosil, P. & Schluter, D. 2011 The genes underlying the process of speciation. *Trends Ecol. Evol.* **26**, 160–167. (doi:10.1016/j.tree.2011.01.001)
- Mayr, E. 1942 *Systematics and the origin of species*. New York, NY: Columbia University Press.
- Orr, H. A. & Turelli, M. 2001 The evolution of post-zygotic isolation: accumulating Dobzhansky–Muller incompatibilities. *Evolution* **55**, 1085–1094.
- Orr, H. A. 1995 The population-genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* **139**, 1805–1813.
- Gavrilets, S. 2004 *Fitness landscapes and the origin of species* (eds S. A. Levin & H. S. Horn). Princeton, NJ: Princeton University Press.
- Nosil, P., Funk, D. J. & Ortíz-Barrientos, D. 2009 Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* **18**, 375–402. (doi:10.1111/j.1365-294X.2008.03946.x)
- Noor, M. A. F. & Feder, J. L. 2006 Speciation genetics: evolving approaches. *Nat. Rev. Genet.* **7**, 851–861. (doi:10.1038/nrg1968)
- Nielsen, R. 2005 Molecular signatures of natural selection. *Ann. Rev. Genet.* **39**, 197–218. (doi:10.1146/annurev.genet.39.073003.112420)
- Beaumont, M. A. 2005 Adaptation and speciation: what can  $F_{ST}$  tell us? *Trends Ecol. Evol.* **20**, 435–440. (doi:10.1016/j.tree.2005.05.017)
- Ellegren, H. 2008 Sequencing goes 454 and takes large-scale genomics into the wild. *Mol. Ecol.* **17**, 1629–1631. (doi:10.1111/j.1365-294X.2008.03699.x)
- Felsenstein, J. 1981 Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* **35**, 124–138. (doi:10.2307/2407946)
- Feder, J. L., Gejji, R., Yeaman, S. & Nosil, P. 2012 Establishment of new mutations under divergence and genome hitchhiking. *Phil. Trans. R. Soc. B* **367**, 461–474. (doi:10.1098/rstb.2011.0256)
- Via, S. 2012 Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Phil. Trans. R. Soc. B* **367**, 451–460. (doi:10.1098/rstb.2011.0260)
- Strasburg, J. L., Sherman, N. A., Wright, K. M., Moyle, L. C., Willis, J. H. & Rieseberg, L. H. 2012 What can patterns of differentiation across plant genomes tell us about adaptation and speciation? *Phil. Trans. R. Soc. B* **367**, 364–373. (doi:10.1098/rstb.2011.0199)
- Nachman, M. W. & Payseur, B. A. 2012 Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Phil. Trans. R. Soc. B* **367**, 409–421. (doi:10.1098/rstb.2011.0249)
- Fan, S., Elmer, K. R. & Meyer, A. 2012 Genomics of adaptation and speciation in cichlid fishes: recent

- advances and analyses in African and Neotropical lineages. *Phil. Trans. R. Soc. B* **367**, 385–394. (doi:10.1098/rstb.2011.0247)
- 29 Hahn, M. W., White, B. J., Muir, C. D. & Besansky, N. J. 2012 No evidence for biased co-transmission of speciation islands in *Anopheles gambiae*. *Phil. Trans. R. Soc. B* **367**, 374–384. (doi:10.1098/rstb.2011.0188)
- 30 Hohenlohe, P. A., Bassham, S., Currey, M. & Cresko, W. A. 2012 Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Phil. Trans. R. Soc. B* **367**, 395–408. (doi:10.1098/rstb.2011.0245)
- 31 McGaugh, S. E. & Noor, M. A. F. 2012 Genomic impacts of chromosomal inversions in parapatric *Drosophila* species. *Phil. Trans. R. Soc. B* **367**, 422–429. (doi:10.1098/rstb.2011.0250)
- 32 Nadeau, N. J. *et al.* 2012 Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. *Phil. Trans. R. Soc. B* **367**, 343–353. (doi:10.1098/rstb.2011.0198)
- 33 Renaut, S., Mailliet, N., Normandeau, E., Sauvage, C., Derome, N., Rogers, S. M. & Bernatchez, L. 2012 Genome-wide patterns of divergence during speciation: the lake whitefish case study. *Phil. Trans. R. Soc. B* **367**, 354–363. (doi:10.1098/rstb.2011.0197)
- 34 Turner, T. L. & Hahn, M. W. 2010 Genomic islands of speciation or genomic islands and speciation? *Mol. Ecol.* **19**, 848–850. (doi:10.1111/j.1365-294X.2010.04532.x)
- 35 Noor, M. A. F. & Bennett, S. M. 2010 Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species (Correction to *Heredity* 2009, vol. 103, p. 434). *Heredity* **104**, 418. (doi:10.1038/hdy.2010.13)
- 36 Gompert, Z. & Buerkle, C. A. 2009 A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Mol. Ecol.* **18**, 1207–1224. (doi:10.1111/j.1365-294X.2009.04098.x)
- 37 Feder, J. L., Geji, R., Powell, T. H. Q. & Nosil, P. 2011 Adaptive chromosomal divergence driven by mixed geographic mode of evolution. *Evolution* **65**, 2157–2170. (doi:10.1111/j.1558-5646.2011.01321.x)
- 38 Gompert, Z., Parchman, T. L. & Buerkle, C. A. 2012 Genomics of isolation in hybrids. *Phil. Trans. R. Soc. B* **367**, 439–450. (doi:10.1098/rstb.2011.0196)
- 39 Nosil, P. & Flaxman, S. M. 2011 Conditions for mutation-order speciation. *Proc. R. Soc. B* **278**, 399–407. (doi:10.1098/rspb.2010.1215)
- 40 Barton, N. H. 2001 The role of hybridization in evolution. *Mol. Ecol.* **10**, 551–568. (doi:10.1046/j.1365-294x.2001.01216.x)
- 41 Guerrero, R. F., Rousset, F. & Kirkpatrick, M. 2012 Coalescent patterns for chromosomal inversions in divergent populations. *Phil. Trans. R. Soc. B* **367**, 430–438. (doi:10.1098/rstb.2011.0246)
- 42 Turner, T. L., Hahn, M. W. & Nuzhdin, S. V. 2005 Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biol.* **3**, e285. (doi:10.1371/journal.pbio.0030285)
- 43 Kirkpatrick, M. & Ravigné, V. 2002 Speciation by natural and sexual selection: models and experiments. *Am. Nat.* **159**, S22–S35. (doi:10.1086/338370)
- 44 Gavrillets, S. 2003 Perspective. Models of speciation: what have we learned in 40 years? *Evolution* **57**, 2197–2215.
- 45 Via, S. & West, J. 2008 The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Mol. Ecol.* **17**, 4334–4345. (doi:10.1111/j.1365-294X.2008.03921.x)
- 46 Via, S. 2009 Natural selection in action during speciation. *Proc. Natl Acad. Sci. USA* **106**, 9939–9946. (doi:10.1073/pnas.0901397106)
- 47 Nosil, P., Vines, T. H. & Funk, D. J. 2005 Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* **59**, 705–719.
- 48 Rice, W. R. & Hostert, E. E. 1993 Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* **47**, 1637–1653. (doi:10.2307/2410209)
- 49 Michel, A. P., Sim, S., Powell, T. H. Q., Taylor, M. S., Nosil, P. & Feder, J. L. 2010 Widespread genomic divergence during sympatric speciation. *Proc. Natl Acad. Sci. USA* **107**, 9724–9729. (doi:10.1073/pnas.1000939107)
- 50 Nosil, P., Harmon, L. J. & Seehausen, O. 2009 Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol.* **24**, 145–156. (doi:10.1016/j.tree.2008.10.011)
- 51 Gompert, Z. & Buerkle, C. A. 2011 Bayesian estimation of genomic clines. *Mol. Ecol.* **20**, 2111–2127. (doi:10.1111/j.1365-294X.2011.05074.x)
- 52 Gompert, Z. & Buerkle, C. A. 2011 A hierarchical Bayesian model for next-generation population genomics. *Genetics* **187**, 903–917. (doi:10.1534/genetics.110.124693)
- 53 Gompert, Z., Forister, M. L., Fordyce, J. A., Nice, C. C., Williamson, R. J. & Buerkle, C. A. 2010 Bayesian analysis of molecular variance in pyrosequences quantifies population genetic structure across the genome of *Lycæides* butterflies. *Mol. Ecol.* **19**, 2455–2473.
- 54 Lawniczak, M. K. N. *et al.* 2010 Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. *Science* **330**, 512–514.
- 55 Hoekstra, H. E. & Coyne, J. A. 2007 The locus of evolution: evo devo and the genetics of adaptation. *Evolution* **61**, 995–1016. (doi:10.1111/j.1558-5646.2007.00105.x)
- 56 Pavey, S. A., Collin, H., Nosil, P. & Rogers, S. 2010 The role of gene expression in ecological speciation. *Ann. NY Acad. Sci.* **1206**, 110–129. (doi:10.1111/j.1749-6632.2010.05765.x)
- 57 Charlesworth, B., Nordborg, M. & Charlesworth, D. 1997 The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet. Res.* **70**, 155–174. (doi:10.1017/S0016672397002954)
- 58 Feder, J. L. & Nosil, P. 2010 The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution* **64**, 1729–1747. (doi:10.1111/j.1558-5646.2009.00943.x)
- 59 Yeaman, S. & Otto, S. P. 2011 Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution* **65**, 2123–2129. (doi:10.1111/j.1558-5646.2011.01277.x)
- 60 Yeaman, S. & Whitlock, M. C. 2011 The genetic architecture of adaptation under migration-selection balance. *Evolution* **65**, 1897–1911. (doi:10.1111/j.1558-5646.2011.01269.x)
- 61 Rieseberg, L. H. 2001 Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* **16**, 351–358. (doi:10.1016/S0169-5347(01)02187-5)
- 62 Butlin, R. K. 2005 Recombination and speciation. *Mol. Ecol.* **14**, 2621–2635. (doi:10.1111/j.1365-294X.2005.02617.x)
- 63 Faria, R. & Navarro, A. 2010 Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends Ecol. Evol.* **25**, 660–669. (doi:10.1016/j.tree.2010.07.008)
- 64 Strasburg, J. L., Scotti-Saintagne, C., Scotti, I., Lai, Z. & Rieseberg, L. H. 2009 Genomic patterns of adaptive divergence between chromosomally differentiated sunflower species. *Mol. Biol. Evol.* **26**, 1341–1355. (doi:10.1093/molbev/msp043)
- 65 Feder, J. L. & Nosil, P. 2009 Chromosomal inversions and species differences: when are genes affecting adaptive divergence and reproductive isolation expected to

- reside within inversions? *Evolution* **63**, 3061–3075. (doi:10.1111/j.1558-5646.2009.00786.x)
- 66 Barrett, R. D. H. & Schluter, D. 2008 Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44. (doi:10.1016/j.tree.2007.09.008)
- 67 Kirkpatrick, M. & Barton, N. 2006 Chromosome inversions, local adaptation and speciation. *Genetics* **173**, 419–434. (doi:10.1534/genetics.105.047985)
- 68 Darwin, C. 1859 *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London, UK: John Murray.
- 69 Hendry, A. P., Bolnick, D. I., Berner, D. & Peichel, C. L. 2009 Along the speciation continuum in sticklebacks. *J. Fish. Biol.* **75**, 2000–2036. (doi:10.1111/j.1095-8649.2009.02419.x)
- 70 Peccoud, J., Ollivier, A., Plantegenest, M. & Simon, J. C. 2009 A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. *Proc. Natl Acad. Sci. USA* **106**, 7495–7500. (doi:10.1073/pnas.0811117106)
- 71 Mallet, J., Beltran, M., Neukirchen, W. & Linares, M. 2007 Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evol. Biol.* **7**. (doi:10.1186/1471-2148-7-28)
- 72 Seehausen, O. *et al.* 2008 Speciation through sensory drive in cichlid fish. *Nature* **455**, 620–623. (doi:10.1038/nature07285)
- 73 Nosil, P. & Sandoval, C. P. 2008 Ecological niche dimensionality and the evolutionary diversification of stick insects. *PLoS ONE* **3**, e1907. (doi:10.1371/journal.pone.0001907)
- 74 Berner, D., Grandchamp, A. C. & Hendry, A. P. 2009 Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. *Evolution* **63**, 1740–1753. (doi:10.1111/j.1558-5646.2009.00665.x)
- 75 Paterson, S. *et al.* 2010 Antagonistic coevolution accelerates molecular evolution. *Nature* **464**, 275–278. (doi:10.1038/nature08798)
- 76 Araya, C. L., Payen, C., Dunham, M. J. & Fields, S. 2010 Whole-genome sequencing of a laboratory-evolved yeast strain. *BMC Genom.* **11**, 88. (doi:10.1186/1471-2164-11-88)
- 77 Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E. & Kim, J. F. 2009 Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1274. (doi:10.1038/nature08480)