Title

Genomics and the origin of species.

Journal

Nature reviews. Genetics, 15(3)

ISSN

1471-0056

Authors

Seehausen, Ole
Butlin, Roger K
Keller, Irene
et al.

Publication Date

2014-03-01

DOI

10.1038/nrg3644

Peer reviewed

Genome-scale data and the genetics of speciation

Claudia Bank[1], Janette W. Boughman[2,3], Åke Brännström[4], Alan Brelsford[5], Roger K. Butlin[6,] Chris Clarkson[7], Fabrice Eroukhmanoff[8], Jeffrey L. Feder[9], Martin Fischer[10], Andrew D. Foote[11], Paolo Franchini[12], Paul A. Hohenlohe[13], Chris D. Jiggins[14], Felicity C. Jones[15], Irene Keller[2,10,16], Anna K. Lindholm[17], Kay Lucek[2,16], Martine E. Maan[18], David A. Marques[2,16], Simon H. Martin[14], Blake Matthews[19], Joana I. Meier[2,16], Markus Möst[10,19], Michael W. Nachman[20], Etsuko Nonaka[21], Catherine L. Peichel[22], Diana J. Rennison[23], Glenn-Peter Saetre[8], Julia Schwarzer[2,16,24], Ole Seehausen[2,16*], Catherine E. Wagner[2,16], Eric T. Watson[25], Anja M. Westram[26], Alex Widmer[10]
(authorship currently in alphabetical order)

*correspondence ole.seehausen@eawag.ch

[1] School of Life Science, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

[2] Department of Fish Ecology and Evolution, EAWAG Swiss Federal Institute of Aquatic Science and Technology, Center of Ecology, Evolution and Biogeochemistry, Seestrasse 79, 6047 Kastanienbaum, Switzerland

[3] Department of Zoology; Ecology, Evolutionary Biology & Behavior Program; BEACON, Michigan State University, 203 Natural Sciences, East Lansing, Michigan 48824, USA

[4] Integrated Science Lab & Department of Mathematics and Mathematical Statistics, Umeå University, 90187 Umeå, Sweden

[5] Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland

[6] Department of Animal and Plant Sciences, The University of Sheffield, Sheffield, UK and Sven Lovén Centre – Tjärnö, University of Gothenburg, S-452 96 Strömstad, Sweden

[7] Liverpool School of Tropical Medicine, Liverpool, UK

[8] Department of Biology, Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, P. O. Box 1066, Blindern, N-0316 Oslo, Norway

[9] Department of Biological Sciences, University of Notre Dame, South Bend, IN, USA

[10] Institute of Integrative Biology, ETH Zürich, Universitätsstrasse 16, ETH Zentrum CHN, 8092 Zürich, Switzerland

[11] Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, ØsterVolgade 5-7, DK-1350 Copenhagen, Denmark

[12] Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz, Universitätstrasse 10, 78457 Konstanz, Germany

[13] Department of Biological Sciences, Institute of Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844-3051, USA

[14] Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

[15] Friedrich Miescher Laboratory of the Max Planck Society, Tübingen, Germany

[16] Department of Aquatic Ecology and Macroevolution, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland

[17] Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

[18] Behavioural Biology Group, Centre for Behaviour and Neurosciences, University of Groningen, P.O. Box 11103, 9700 CC Groningen, The Netherlands

[19] Department of Aquatic Ecology, Centre of Ecology, Evolution and Biogeochemistry, Eawag Swiss Federal Institute of Aquatic Science and Technology, Kastanienbaum, Switzerland

[20] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA

[21] Integrated Science Lab & Department of Ecology and Environmental Science, Umeå University, 90187 Umeå, Sweden

[22] Division of Human Biology, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, Washington 98109, USA

[23] Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada

[24] Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee 160, 53113 Bonn, Germany

[25] Department of Biology, The University of Texas at Arlington, TX, USA

Introduction

Speciation is a central process in evolution, distinct from adaptation and critical for understanding the origins of diversity. Major insights into mechanisms of speciation have come from a diverse range of genetic approaches: mapping of individual speciation genes, mapping of genome-wide divergence, quantitative genetic analysis of phenotypes and of reproductive isolation, genetic and functional constraints, molecular phylogenetics and population genetics[1,2]. Theoretical understanding of the genetics of speciation has also advanced markedly[3-6]. However, at a time when the field is faced with a flood of data from next generation sequencing, more integrated conceptual foundations are needed to make the most of the available opportunities, linking different types of genetic change and connecting genetic change to phenotypes, and to ecological and non-ecological drivers. In this review, we identify elements required, available and still needed to build these foundations. We start out by reviewing three areas of study in which rapid progress has recently been made through the use of population genomic methods (patterns and rates of genome-wide divergence, the evolution of intrinsic incompatibilities, genomic coupling); and two areas of study that are central to speciation genetics, but have not yet come to capitalize on the new genomic data to the full extent possible (the distribution of variant effect sizes on speciation, and genomic constraint). We make suggestions for how genome-scale data can contribute to further progress in all of these fields and to aid in the construction of a synthetic view on the genetics and genomics of speciation.

Patterns and rates of genome-wide divergence in the course of speciation

Speciation can rarely be studied in real time. However, by integrating case studies of speciation among closely related taxa that vary in their extent of differentiation, speciation researchers are provided windows into the mechanisms at play at different time-points in the process of speciation. Emerging genomic studies making use of this approach of studying the "speciation continuum" show that incipient species surprisingly quickly accumulate divergence at very many regions in the genome, even in the presence of gene flow (Box 1). In a few examples, such as *Heliconius* races[7] and sunflower ecotypes[8], early stage divergence is limited to a few genomic regions. However, in many other cases, divergence is already extensive early in the process of speciation[9-12]. Oftentimes background F_{ST} increases as phenotypic divergence increases[13,14] (Fig. 1). These studies provide evidence that genomic divergence during speciation is often heterogeneous across the genome and can be accentuated in regions of low recombination, including centers of chromosomes or near centromeres[15,16]. In addition, sex chromosomes show elevated divergence in many systems [16-18].

One of the most exciting and surprising findings of these studies of replicate cases of ecological speciation is the contribution of ancient allelic variation to recent divergence, as exemplified by stickleback[11,19], cichlids[12,20], *Rhagoletis* flies[21] and *Heliconius* butterflies[22]. The sources of such ancient allelic variation can either be **standing genetic variation**[23], defined in the narrow sense as genetic variation with a history of residence in a single population, and **hybridization** between species or distinct populations[24]. Variation arising from either of these sources is often referred to as standing genetic variation in the wider sense, but distinguishing between these sources is important albeit not trivial. The balance of evidence in all of the above mentioned cases suggest that hybridization brought ancient genetic variation into single populations. These data suggest that adaptation and speciation might often have been facilitated by hybridization introducing ancient variation into ancestral populations and providing the genetic material for adaptation and reproductive isolation in the face of gene flow.

Despite these intriguing results, caution is warranted. Heterogeneity in divergence patterns is commonly inferred to result from a mosaic of divergent selection and gene flow across the genome. However, these patterns can also be generated by correlated co-ancestry inflating the neutral variance in F_{ST}[25], heterogeneity in recombination rates[26,27] and background selection (e.g. gene density[28]). Indeed, heterogeneous divergence is observed between allopatric populations of the

same species in the absence of gene flow[29] (Fig. 1; Martin et al., in review). Thus, it is important to consider these factors when interpreting the results of genome scans[13,25].

Some studies have used alternative lines of evidence to identify regions that are likely involved in reproductive isolation between populations or species. For example, combining classic cline theory with genome-wide analyses has allowed measurements of the strength of selection at specific loci[30] (Box 2). Another approach involves identifying genomic regions that diverge repeatedly across replicated pairs of species or environmental contrasts, which provides strong evidence that the regions are involved in adaptation and/or reproductive isolation[9,11,31-34]. Finally, more parameter-rich model-based analyses, such as fitting coalescent models of divergence with gene flow to genomic data, consider the heterogeneity of demographic history across the genome to identify genomic regions that have experienced gene flow[35,36].

Speciation can be initiated either by divergent extrinsic, i.e. "ecological" selection, or by the evolution of intrinsic genic incompatibilities (Box 1). Even though studying the evolutionary continuum of the accumulation of intrinsic genic incompatibilities has a strong tradition in evolutionary biology[37], the currently available population genomic studies of multiple populations along the speciation continuum have mostly investigated candidate cases of ecological speciation. However, as replicate studies of speciation driven by both extrinsic selection and intrinsic incompatibilities (panels a and b, Box 1) accumulate, it will become possible to ask whether different mechanisms and modes of speciation can be distinguished based on patterns observed from genome-wide data.

"Speciation effect sizes" of alleles

While F_{ST} estimates from genome scans characterize the extent of divergence between populations or species, it is important to recognize that they tell us little about the effect sizes of individual loci on phenotypes, fitness or reproductive isolation[38,39] (Fig 2). Fisher's classical geometric model[40] predicts that the number of mutations contributing to adaptive evolution is numerous and that their effect sizes are small. More recent theoretical models predict that fitness effect size follows an exponential distribution[41] (Fig. 2c). As a population climbs an adaptive peak, mutations of progressively smaller effect on fitness are favoured by selection. Although several empirical studies support an exponential distribution of mutation effect size underlying phenotypic traits (e.g., [42-45]), relatively few data are available on effect sizes underlying fitness in natural diverging populations[46,47], and even less information is available concerning mutation effect sizes underlying reproductive isolation. Moreover, this body of work is based on theory of adaptation rather than speciation, and does not take into account standing variation, gene flow or changing environments. When those factors are considered the predictions change[48-50] and may even reverse[51].

Premating isolation traits, such as conspecific mate choice, habitat choice, and different reproductive timing are likely to have large effects on reproductive isolation because they directly influence mating patterns[1,6,52-56]. Unfortunately, studies exploring the molecular genetic basis of behavioral isolation traits are relatively few in number[20,57-62] as compared to morphological traits relevant to speciation. Identifying the genetic basis of behavioral isolating mechanisms is complicated by relatively high levels of plasticity, mediated by learning and phenotype-dependent habitat choice[63,64]. Loci involved in premating isolation and in extrinsic and intrinsic postmating isolation may often differ in effect size, but current data are equivocal. For example, mapping hybrid inferiority in natural environments for *Arabidopsis* has shown reproductive isolation is polygenic[65]. In contrast, hybrid inviability in *Mimulus guttatus* arising from differential adaptation to copper soil is due to linked loci of major effect[66]. Intrinsic postmating DMIs may depend on multiple interacting factors[67] or have a simple genetic basis[68,69].

Because the effect sizes of a given genotype on phenotype, fitness, or reproductive isolation are not the same, we must independently estimate the magnitudes of each of these effects. QTL mapping[15,66] and admixture mapping using large numbers of SNP markers[58,70] can link genotype to phenotype, fitness and reproductive isolation (Box 2); however, an unfortunate drawback of these approaches is their bias toward detecting loci of large effect. Estimates of fitness and reproductive isolation effects can be achieved using manipulative selection experiments which track allelic changes or genome-wide responses[29,71] and estimates of these effects can also be ascertained by measuring selection and introgression in the wild (comparatively or experimentally). All of these approaches hinge on our ability to genotype a large number of loci across the genomes of many individuals. As sequencing costs continue to decrease, this proposition becomes more and more feasible, even for non-model organisms. For example, a genotyping by sequencing method, RAD-seq (Box 2), has recently been used to map QTL underlying phenotypes and reproductive isolation (RI) in whitefish[14]. However, our knowledge about the sources of selection is still incomplete even for genes of major effect. Still, such experiments need to be rigorously coupled with studies of the ecology of speciation in natural populations in order to understand the sources of selection.

Often when large effect alleles involved in adaptation and speciation are identified, they turn out to be highly phenotypically pleiotropic (e.g., Ectodysplasin [Eda] in sticklebacks; [71,72], although we lack estimates of the effect Eda has on reproductive isolation). Pleiotropy can facilitate speciation because it connects multiple traits and can increase the chance that divergent adaptation "couples" to assortative mating[53,73,74]. Pleiotropy also exposes an allele to multifarious selection. Whether pleiotropic effects on fitness are antagonistic or synergistic (affecting all fitness traits in similar directions) and aligned with the direction of divergent selection will determine the magnitude of evolutionary change and the effect of the substitution on reproductive isolation. Synergistic pleiotropy is likely to facilitate speciation by predisposing loci to couple different barriers whereas antagonistic pleiotropy will likely slow speciation by making some loci less likely to contribute to reproductive isolation.

The discovery that ancient allelic variation plays important roles in several cases of ecological speciation (see above) may speak to the role of alleles of larger effect sizes. Alleles of large phenotypic effect are often recruited from ancient genetic variation[11,17,19,20,38]. This likely increases the rate of evolutionary response and speciation both by removing the need to wait for new mutation and because these alleles have been honed by selection over time. They are, therefore, likely to have large positive effects on both adaptation and reproductive isolation[75].

Genomic constraint

With the new population genomic data revealing divergence at many regions of the genome early in speciation, it is appropriate to take a quantitative genetic perspective to the multivariate evolution of polygenic traits during speciation. In quantitative genetic terms, standing variation is quantified by the G matrix of additive genetic (co)variance[76]. To the extent that traits are tightly correlated with each other as a result of pleiotropy or linkage disequilibria, the G matrix illustrates constraints on adaptive evolution. These constraints affect the response to directional selection[77-79]. Divergence among populations is biased along multivariate axes with greater genetic variation and constrained along axes with little variation[78,80,81]. Importantly, however, genetic constraints are not only negative. They may align with correlational selection[82,83] and thus prevent maladaptive trait combinations from arising. There is some empirical evidence that aspects of the structure of the G matrix may persist over long periods of time, maintained by mutational correlation and/or correlational selection[76], although other studies have found significant changes within populations[84]. It is less clear to what extent genes of major effect, versus the traditional assumption of many genes of small effect, may influence the structure of the G matrix[85,86], and how higher moments of the distribution of genetic variation affect the response to selection[87]. These questions can now be addressed with genomic

methods. Gene regulatory pathways can lead to transcriptional constraints on multivariate divergence that persist over multiple speciation events[88]. However, developmental pathways may also be flexible across species so that multiple mechanisms can produce similar patterns of correlation, or similar developmental mechanisms can lead to different correlations among traits[89]. Further work in this direction will help to illuminate the nature of genomic constraint that may affect speciation.

Important for facilitating speciation is standing variation in traits potentially under divergent or disruptive selection and in traits related to reproductive isolation, such as mate preference[90]. This standing variation may depend on genotype-by-environment interactions: cryptic variation may be revealed by environmental shifts or epigenetic changes and may then contribute to adaptation. Phenotypic plasticity can also lead to different phenotypes in alternative environments and may contribute to divergence through genetic assimilation[91]. To evaluate the role of standing genetic variation in speciation, it is critical to better understand its effects on the G matrix and, in particular, its role in increasing evolvability in the direction of selection.

Traditionally, gene flow and hybridization have been thought to restrict adaptive divergence by homogenizing variation across populations. However, gene flow and hybridization may also facilitate adaptation by releasing constraints caused by genetic correlations. While empirical evidence has accumulated in the last decade suggesting a strong role for selection in altering genetic architecture[84,92], the role of gene flow in aligning the G matrix in the direction of divergent or disruptive selection has been mainly investigated theoretically[51]. The emerging consensus that hybridization can introduce novel and potentially adaptive variation[75] calls for added research and empirical studies in this area. We predict that hybridization will influence the G matrix in the following ways: First, hybridization may lead to the evolution of intermediate phenotypes not present in the parental species (Figure 3a). Second, gene flow between diverging populations should lead to the alignment of genetic covariances along the axis of divergence[51]. Third, transgressive segregation could effectively lead to the emergence of novel genetic correlations between traits, which would result in new dimensions of evolvability[93,94] (Figure 3a). Hybrids might often be maladapted, but hybrid populations may benefit from increased evolvability[95]. While selection may alter patterns of genetic covariance and facilitate evolutionary response to divergent selection over longer time scales, novel patterns of genetic variation produced by hybridization may lead to bursts of evolutionary divergence and speciation. That burst would be observable at the phylogenetic level[24] (Figure 3b-d). With rapidly increasing SNP densities or whole genome resequencing, these **hypotheses regarding the impact of hybridization can now be tested by comparing the phylogenetic** history of regions in the genome that confer adaptation and reproductive isolation to those of other genomic regions[12,17].

Genomic conflict and the evolution of intrinsic postzygotic isolation

Speciation research has been divided into research on isolation due to extrinsic forces (i.e. ecological speciation) and intrinsic hybrid incompatibilities causing post-zygotic reproductive isolation. Studying these different sources of isolation has required different approaches in the past, and the integration of the resulting literatures remained fractious. Ecological speciation research has focused on closely related but phenotypically and ecologically distinct populations and species and asked if and how divergent adaptation constrains gene flow in nature either directly[6,105] or through interactions with mate choice[53,96]. Intrinsic hybrid incompatibility research on the other hand has mostly worked with somewhat older species and has used experimental hybridization in the lab outside the ecological context to study the time course of isolation[50,106-108] and identify the genes involved[97]. These approaches were only rarely combined in the same taxa[98]. Population genomic approaches to studying speciation now have the potential to change this dichotomy in the field of speciation genomics and allow a full integration of speciation processes spanning these mechanisms.

Intragenomic conflict, defined as antagonistic selection among genomic elements with different fitness interests in an individual, may be a powerful force driving the evolution of intrinsic hybrid incompatibilities[97,99-101]. A key property of speciation driven by genomic conflict is the potential for unlimited change due to ongoing co-evolution. This evolutionary change may give rise to new alleles at multiple loci that do not interact properly in hybrids and thus lead to reproductive isolation (Fig. 4). Genomic conflict may arise from competing interests of males and females[102], and such things as meiotic drivers[103,104], mobile elements[105,106], or other selfish elements and their suppressors, and organellar and nuclear genomes[107]. Similar patterns of co-evolution may also result from intragenomic conflict driven by disruptive extrinsic selection between the sexes[108] or intergenomic conflict between hosts and pathogens[109]. Thus, genomic conflict more generally (both intra- and intergenomic) may be caused by both intrinsic and extrinsic factors, although reproductive isolation arising from genomic conflict is less directly dependent on the environment, in contrast to ecological isolation.

Strong evidence for intragenomic conflict in speciation comes from specific genes known to underlie Dobzhansky-Muller incompatibilities (DMIs), many of which appear to be involved in antagonistic interactions. A well-studied example is *Ovd*, an X-linked gene that underlies both hybrid male sterility and sex-ratio distortion in crosses between *Drosophila pseudoobscura pseudoobscura* and *D. p. bogotana*[110]. In this cross, hybrid males are mostly sterile but are able to sire some offspring when aged, and then produce mostly daughters. Genomic conflict also explains many widely observed patterns in intrinsic postzygotic isolation. First, post-zygotic isolation is usually caused by negative epistatic interactions in hybrids (DMIs[1]), and genomic conflict requires interactions[99]. Second, nearly all DMI genes show evidence of positive selection[97]. This is consistent with recurring antagonistic co-evolution. Third, the high density of sex-linked DMI genes[97] can be explained by the fact that sex-linked meiotic drivers are expected to invade more easily than autosomal ones. This is because the sex chromosomes (X and Y or Z and W) are constantly in a battle over segregation whereas only small, tightly linked autosomal regions are in conflict with their homologs[100]. Moreover, there will be particularly strong selection for suppression of sex-linked compared to autosomal distorter loci because the former tend to bias sex ratios[110,111]. Finally, the observation of Haldane's rule for sterility in both male heterogametic and female heterogametic taxa is consistent with sex-linked meiotic drivers, but not with some alternative explanations, like "faster male evolution"[101].

Speciation due to genomic conflict may be facilitated or hindered under certain conditions. For example, gene flow between incipient species in the early stages of divergence may allow the spread of selfish genetic elements and thereby slow the accumulation of conflict-driven DMIs[112]. Thus, speciation driven by genomic conflict is more likely in allopatry. However, relatively brief periods of allopatry may be sufficient for the creation of conflict-driven DMIs since selfish genetic elements can invade quickly. Genomic conflict is also more likely in species with differentiated sex chromosomes, since this situation provides more potential arenas for antagonistic co-evolution.

Genomic conflict is expected to be particularly common in certain genomic regions and thus may lead to associations between loci. First, several well characterized drive systems, such as mouse t-haplotypes and SD in *Drosophila*, occur in regions of suppressed recombination, allowing for linkage disequilibrium between distorter and responder loci[113,114]. More generally, sex chromosomes are particularly susceptible to the accumulation of DMIs derived from genomic conflict and mutations with different fitness effects in males and females can establish on sex chromosomes more easily than on autosomes. These considerations lead to the prediction that genomic conflict may build up associations among loci on the sex chromosomes and in regions of suppressed recombination. Since the sex chromosomes are also attractors for sex-limited sexual traits and preferences[115-118] this can lead to associations between pre- and postzygotic isolation, including that caused by conflict-driven DMIs (see below on genomic coupling).

Sexual conflict is predicted to drive the evolution of new sex chromosome systems[119,120], and empirical evidence exists in fish [15,108]. Theoretical models and empirical data in Lake Victoria cichlids of the genus *Neochromis* suggest that the invasion of a new sex chromosome can lead to RI[118,121]. In sticklebacks, traits that contribute to RI map to a new sex chromosome, suggesting a direct role for sex chromosome turnover in the evolution of RI[15].

Discovery of DMIs used to be laborious. Genomic data are now allowing us to identify DMI loci at an increasing pace[122,123] (Box 2). The evidence for conflict-driven DMIs raises a number of important questions. What kinds of conflict are most important in causing speciation? How often is postzygotic isolation caused by conflict rather than other kinds of DMIs that may accumulate as a consequence of ecological selection or drift? What is the relative importance of extrinsic vs. intrinsic postzygotic isolation? We anticipate that the identification of many DMI loci in the near future will provide partial answers to these questions.

Genomic Coupling

Coupling is the statistical association between different traits involved in reproductive isolation (RI)[124,125]. The build-up of trait associations strengthens the barrier to gene flow between diverging populations, and is therefore important for the evolution of strong reproductive isolation.

Associations between traits can initially be generated by random processes or by divergence with limited gene flow. Multiple coinciding barriers can, for example, be produced by secondary contact between two divergent populations, by the evolution of DMIs as an accidental by-product of divergent selection[126], or via hitchhiking of intrinsic incompatibility alleles with divergently selected alleles[66]. Within populations, trait associations can be created by random drift or by mutations generating a new barrier, which will initially always be associated with a single genetic background. For trait associations to be important for speciation, however, they have to be maintained or even strengthened in the face of gene flow, and this typically involves selection.

Selection is expected to favour the coupling of multiple existing barriers if this leads to an increase in mean fitness. Coupling can involve intrinsic barriers (like DMIs)[127,128] or, across an ecotone, multifarious extrinsic selection can assemble and maintain many coinciding clines at traits involved in local adaptation[129]. These two types of barriers – intrinsic and extrinsic – can also become coupled with each other at ecotones ([75,125,130]). Finally, when hybrids have reduced fitness, selection can directly favour the evolution of increased premating isolation through reinforcement [131].

Because recombination tends to break up associations, genomic architectures that eliminate or decrease recombination are generally expected to facilitate coupling, and hence speciation [132]. Most prominently, recombination will not affect traits pleiotropically influenced by the same alleles or by 'one-allele' mechanisms where the same phenotype in different genetic backgrounds confers RI [133]. Here, it is important to point out that one-allele mechanisms do not leave a population-specific signature in the genome at the primary isolation locus; methods other than genome scans or QTL mapping of diverging populations will be needed to find genes underlying these theoretically important traits. Reduced recombination is expected between loci in close physical proximity and in certain regions of the genome, particularly sex chromosomes, centromeres, inversions and other genomic rearrangements. Therefore, loci underlying two-allele mechanisms may be concentrated in such regions but so far we have few concrete examples.

The recent advances with genome-wide studies have produced an accumulation of empirical examples for coupling between unlinked loci, but also for a role of genomic architectures that eliminate or reduce recombination between traits involved in RI. In hybrid zones, clines at many loci, including unlinked loci, often coincide, although it is not always clear exactly how these loci are implicated in RI[134]. There is also (rare) evidence for one-allele mechanisms[135], habitat matching[136] and multiple-effect traits[137-139]. Genes underlying multiple isolating traits have been found together in

inversions[140,141], on sex-chromosomes[15,116] and also in tight physical linkage[66,142]. These data further provide some evidence that reinforcement is facilitated by linkage, as in the case of the flycatcher[143], or by multiple effect traits, as in the phlox[139], while reinforcement might be constrained in other cases[144] because loci are not linked and there is extensive gene flow.

Towards a synthesis

New genomic wine makes old bottles of theory shine in new light

(i) geographical modes and gene effect sizes

Historically, speciation research had been preoccupied with an emphasis on the geographical modes of speciation, and has only quite recently moved towards a more mechanistic approach[145]. The new genomic data on speciation add a new twist. Phrased in terms of the genetics of reproductive isolation, speciation can proceed in many different ways, but these can be grouped into two different classes, that are at least in theory quite distinct, illustrated in Box 1. Both of these can readily generate reproductively isolated species in allopatry, and many species doubtlessly owe their origins to periods of allopatry. On the other hand, the majority of recent speciation genomic studies have dealt with cases of speciation with gene flow. When gene flow is present between populations, speciation is constrained to situations where divergent selection exceeds gene flow[6]. So why is such speciation apparently readily occurring in some taxa such as cichlid fish, stickleback, *Rhagoletis* flies and *Heliconius* butterflies[145]?

The first generation of population genomic speciation studies reviewed here have given us glimpses into the potential for a unifying framework for discourse over the importance of major genes[73] versus many genes with small additive effects[146] to speciation. The data suggest that divergence involves many regions in the genome already very early in speciation, but some of these have large effects on adaptation and often pleiotropic effects on isolation. It seems a distinct possibility that allelic substitution at the major loci can seal off gene flow quickly, effectively isolating the genomes so that substitutions with smaller effects can subsequently fix across the genome. A surprising number of the cases studied provide evidence that such alleles with large and pleiotropic effects tend to be ancient variants that were present as standing variation in the ancestors of emerging species pairs. Ancient allelic variants depart in their expected effect sizes from Fisher's geometric model of adaptation in important ways. Whereas the probability that a novel mutation of a given phenotypic size is favorable falls rapidly with mutational size, this is not the case for standing genetic variation[23]. This explains how adaptation and reproductive isolation can possibly proceed rapidly in populations that that are enriched for ancient adaptive genetic variation. A prediction, scrutinizable with population genomic data, is that taxa with large historical population sizes or histories of hybridization are particularly prone to undergo speciation in the face of gene flow.

(ii) types of reproductive isolation

The architecture of **intrinsic postzygotic isolation** is predicted to result from DMIs, and there is good evidence to suggest that such negative epistatic interactions do in fact often underlie intrinsic postzygotic isolation[1,97,147,148]. However, neither the absolute nor the relative time course of their accumulation is well understood[122,147,149,150]. It may vary between taxa and is expected to vary between mechanisms of DMI evolution[147]. While DMIs are certainly important, other mechanisms, like underdominance[151] or gene duplication and loss[152-154] may also contribute to intrinsic postzygotic isolation. Predictions about the distribution of fitness effect sizes expected for genes that underlie DMIs are generally lacking, and whether 2-way or more complex interactions are typical is unknown. Most studies have investigated DMI genes in fully isolated species, and it is unclear if the fixation of the underlying mutations was the cause or a consequence of speciation[97,155] (but see[15,116]).

The genetic basis of traits that cause **extrinsic postzygotic isolation** is expected to be similar to the genetic basis of traits that contribute to adaptation[40,156]; i.e. we expect that a few loci of large effect

and many more loci of smaller effect will contribute additively. However, a different pattern may be expected in cases of ecological speciation in the face of gene flow where evolutionary divergence may be possible only if divergent or disruptive selection operates on a small number of loci[6] and large-effect or pleiotropic alleles may be favoured in the long term[49]. On the same grounds, we hypothesize that effect sizes of alleles that have been "reshuffled" by hybridization (such as *Eda* in stickleback, or *LWS* in Lake Victoria cichlids) are larger than those of new alleles arising through mutation. This is because such alleles tend to be ancient and may have accumulated many mutations that were honed by selection. Because additivity is expected to be widespread, we can predict that hybrids will often be intermediate in phenotype, but we cannot necessarily predict their fitness which will depend on phenotype-environment interactions. Several studies have investigated extrinsic postzygotic isolation in young and incompletely isolated species pairs and there is positive evidence that the sources of this isolation are causaly involved in ecological speciation[2]. Few studies to date though have explicitly tested the genetic architecture of extrinsic postzygotic isolation, and we expect that experimental genomics will shortly fill this gap (Arnegard et al. in prep).

The evolution of **sexual isolation** (assortative mating) is a hallmark of speciation, yet the study of its genetic basis is in its infancy. Sexual isolation may evolve as an indirect consequence of sexual selection and would typically involve allopatric divergence in secondary sexual traits and preferences for such traits[157-159]. Where speciation by disruptive sexual selection happened in the face of gene flow, genetic mechanisms that promote coupling between trait and preference are predicted[160]. Some empirical evidence supports this prediction[116,137,142] (see genomic coupling section) although other sources of isolation probably contributed to speciation in all of these cases, and in other cases no such mechanism is apparent[161]. In general though, we might predict that genes involved in sexual isolation of sympatric hybridizing species will be found in tight linkage with each other and/or with genes for divergent adaptation or post-zygotic isolation. Genome scans in diverging populations combined with mapping of behavioural isolation traits are needed to test this.

Where rates of evolution of different components of reproductive isolation have been studied in the same taxon[37,98,162], data suggest that sexual and extrinsic postzygotic isolation evolve faster than intrinsic postzygotic isolation. Classical theory on the accumulation of DMIs suggests this too[163]. However, it does not always have to be so. Rapid evolution of genes underlying any form of isolation is expected when there is co-evolution (see genomic conflict section). Rates of evolution of multiple components of isolation, and their genetic basis, will have to be studied in many more taxa to reveal general patterns. Genome scans combined with genetic mapping of reproductive isolation provide new opportunities for unifying the genetics of speciation because they allow to identify, study and compare regions in the genome that have diverged under any kind of process, including divergent or disruptive selection, intragenomic conflict, heterogeneity in recombination rates and demographic processes (Box 2).

Because strong reproductive isolation typically involves multiple barriers, it is important to learn more about the coupling between barriers on the level of the whole genome, to understand how it develops and to determine why it is greater in some cases than in others. A full understanding of the extent of coupling will require study of multiple components of isolation in the same system at multiple points on the speciation continuum. In addition, new theory is needed that includes influences of demography, recombination rates and divergence times, and explicitly considers standing genetic variation. We believe that researchers studying speciation will soon be able to empirically test predictions about the genetic architecture of different forms of reproductive isolation, their associations, their relative and absolute rates of evolution and to what extent they leave distinctive genomic signatures in any type of organism.

New data for new theory: speciation genomics and patterns in biodiversity

Speciation researchers have begun to ask how speciation affects patterns in biodiversity[164] and we envisage that speciation genetics can make important and unique contributions here. Study of the distribution of species richness among clades provides evidence for non-uniform diversification rates among taxa, which can arise from differences in speciation and/or extinction rate (e.g. [165]). Speciation rates estimated from the fossil record are far slower than those predicted from mathematical models and observed in studies of recent diversification, and one explanation for this discrepancy is a high frequency of "ephemeral speciation," in which taxa that have recently undergone speciation have high rates of extinction[166]. This has been documented in cases of "speciation reversal"[167-169] which is possible when speciation does not reach "completion"[129,170].

A better understanding of the genomic basis of speciation might help us to understand the influence of speciation on species persistence and patterns of species diversity. For instance, ecological speciation readily and rapidly produces divergent, partially isolated ecotypes and species that may immediately be able to coexist without competitive exclusion. Ecological speciation might thereby contribute disproportionately to the buildup of biodiversity compared to non-ecological mechanisms[164]. However, isolation between young ecologically differentiated species is often extrinsically based and contingent upon the persistence of divergent selection (see Box 1). The species that arise most rapidly may therefore be those species that are most vulnerable to extinction early in their histories[170]. In contrast, speciation via intrinsic mechanisms may produce species that are less prone to ephemerality because speciation reversal may be less likely. However, speciation rates might be slower in these lineages than in lineages where ecological speciation is common, and ecological differences must evolve after speciation in order for closely related taxa to coexist. Progress in connecting speciation to broader-scale patterns of species richness will require attention to how speciation mechanisms, and their genomic basis, influence rates of speciation and the persistence and coexistence of young species. If mechanisms of speciation leave distinctive genomic signatures, correlation between genomic patterns and disparity in species richness among clades could be tested quantitatively using comparative phylogenetic approaches.

Roadmap for a more complete synthesis

We propose four directions of research as the major building blocks for synthesis.

1. We need many more studies of genome-wide divergence of replicate species pairs in the same taxon but at different stages along the speciation continuum. Future sampling design should explicitly aim at crossing the factor "stage on the speciation continuum" with the factor geographical isolation, ranging from allopatry to sympatry (see Fig. 1).
2. Genome scans should be combined with genetic mapping of traits and reproductive isolation, and with annotations of genomes in terms of the effects of alleles on different components of reproductive isolation.
3. We need development of theory to allow relatively parameter rich models capable of making testable predictions.
4. We need experimental studies of the genomics of speciation.

Until now, genome-wide association and admixture mapping have been used to identify the genetic basis of single traits and identify loci of major effect[70]. With increasing genomic data, we can better estimate the effect size distribution across many loci, and measure the influence of loci of smaller effect as well as the relationship between allelic effect sizes on phenotypes and on reproductive isolation. Genome scan and QTL/association studies may identify substitutions that contribute to reproductive isolation but, in non-model species, the functions of these loci may frequently be unknown. *OdsH* is a classic example of a gene where the contribution to isolation was known but its function has been hard to determine[171]. Approaches are needed to build the mechanistic connections between such substitutions and reproductive isolation, via phenotypic and fitness

effects. Tools such as RNAi, TALENS or functional testing of individual substitutions are becoming available for functional characterization in a wider range of taxa as illustrated, for example, by work on pheromone production in moths[62]. New methods and theory are becoming available to improve the interpretation of genome scans in the context of complex demography and genomic features such as variable mutation rates[172] and to use information from re-sequencing or high-density markers more effectively[173].

The feasibility of population genomic data collection on large numbers of individuals has opened completely new opportunities for experimental genomics of adaptation[26] and these approaches should now be extended to experimental speciation research. Specifically, the role of ancient allelic variants with large or pleiotropic effects in generating initial reproductive isolation in response to divergent selection can now be tested with experimental populations that either do or do not harbor these ancient variants.

Conclusions

New approaches for gathering large amounts of genomic data in non-model organisms has produced intriguing and unexpected results. However, it is clear that integrated studies are needed that cover multiple components of RI at multiple stages of the speciation continuum, and in geographical settings ranging from allopatry to sympatry, all within well-chosen study systems. With the latest genomic approaches, we can then construct a picture of the progressive build-up of barriers to gene flow, the way it is influenced by ecological and genetic or genomic constraints and the way it influences patterns of genomic divergence. Given study systems with contrasting histories, we can ask about variation in this process and about its consequences for diversity. We have tantalizing glimpses from the first studies to exploit the latest technologies, and although they are challenging to interpret and assemble into a coherent picture, they have uncovered novel and unexpected insights into the genetics of speciation. There is no doubt that a new phase of discovery has begun which will usher in a much increased understanding of speciation.

Glossary

Admixture mapping
Identification of genetic loci that contribute to phenotypic differences between ancestral populations, by exploring genotype-phenotype correlations in a population of mixed ancestry

Assortative mating
Non-random pattern of mating where individuals sharing particular traits or genes preferentially mate with each other.

Cline
Gradual change in frequency (e.g. of an allele) over a geographic area.

Coalescent
A statistical framework for the analysis of genetic data where the genotypes shared by populations or species are traced back in time to their most recent common ancestor.

Disruptive selection
Selection within a single population which favours extreme phenotypes over intermediates.

Divergent selection
Selection favouring different phenotypes in different populations.

Dobzhansky-Muller Incompatibility (DMI)
An intrinsic postmating barrier due to epistatic interactions between alleles at two or more loci that cause reduced fitness in hybrids but not in the parental species.

Ecological speciation
The evolution of reproductive isolation as a consequence of divergent or disruptive natural selection between populations that inhabit different environments or exploit different resources.

Effect size
Proportion of difference explained, for example the proportion of phenotypic difference attributable to a specific locus or, for isolation, the proportion of total isolation.

Extrinsic reproductive isolation
Fitness reduction in hybrids that is mediated by environmental effects

Intrinsic reproductive isolation
Fitness reduction in hybrids, resulting from developmental problems that are independent of the environment

Linkage disequilibrium
Non-random associations between alleles at different loci, i.e. the occurrence of some combinations of alleles in a population more or less often than expected based on their individual frequencies.

Pleiotropy
Effects of an allele on more than one trait.

Prezygotic reproductive isolation
Barriers acting before or after mating but before fertilisation, including the isolating effects of divergent sexual behaviour, habitat preference, reproductive timing, as well as gametic incompatibility.

Postzygotic reproductive isolation
Barriers acting after fertilisation, such as hybrid sterility and hybrid inviability. Can be extrinsic (mediated by environment) or intrinsic.

Quantitative trait locus (QTL)
Chromosomal region with a statistically significant effect on a phenotype.

Reinforcement
Strengthening of prezygotic barriers in sympatry or parapatry between taxa that have previously evolved sufficiently strong postzygotic isolation for hybrids to experience low fitness

Reproductive isolation
Absence (or restriction) of gene flow between populations.

Speciation continuum
Order of population pairs differing in time of divergence and/or the strength of reproductive isolation in order to illuminate patterns and processes at work in different stages of the speciation process.

Standing genetic variation
Defined in the narrow sense, it is genetic variation with a history of residence in a single population. Defined in the wider sense, it includes variation that results from hybridization between species or distinct populations.

Transgressive segregation
Expression of phenotypic variation in hybrids that exceeds the range of phenotypes observed in the parental taxa.

1 Coyne, J. & Orr, H. *Speciation*. (Sinauer Associates, Sunderland, M.A., 2004).

2 Nosil, P. *Ecological Speciation*. 280 (Oxford University Press, 2012).

3 van Doorn, G. S., Edelaar, P. & Weissing, F. J. On the Origin of Species by Natural and Sexual Selection. *Science* **326**, 1704-1707 (2009).

4 M'Gonigle, L. K., Mazzucco, R., Otto, S. P. & Dieckmann, U. Sexual selection enables long-term coexistence despite ecological equivalence. *Nature* **484**, 506-509, doi:10.1038/nature10971 (2012).

5 Gavrilets, S. & Losos, J. B. Adaptive Radiation: Contrasting Theory with Data. *Science* **323**, 732-737 (2009).

6 Gavrilets, S. Fitness lanscapes and the origin of species. *Princeton University Press* (2004).

7 Nadeau, N. J. *et al.* Genomic islands of divergence in hybridizing Heliconius butterflies identified by large-scale targeted sequencing. *Philosophical Transactions of the Royal Society B-Biological Sciences* **367**, 343-353, doi:10.1098/rstb.2011.0198 (2012).

8 Andrew, R. L. & Rieseberg, L. H. Divergence is focused on few genomic regions early in speciation: incipient speciation of sunflower ecotypes *Evolution* **in press**, doi:doi: 10.1111/evo.12106 (2013).

9 Hohenlohe, P. A. *et al.* Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. *Plos Genetics* **6** (2010).

10 Deagle, B. E. *et al.* Population genomics of parallel phenotypic evolution in stickleback across stream-lake ecological transitions. *Proc. R. Soc. B-Biol. Sci.* **279**, 1277-1286, doi:10.1098/rspb.2011.1552 (2012).

11 Jones, F. C. *et al.* The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55-61, doi:10.1038/nature10944 (2012).

12 Keller, I. *et al.* Population genomic signatures of divergent adaptation, gene flow, and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Mol. Ecol.*, doi: 10.1111/mec.12083 (2012).

13 Roesti, M., Hendry, A. P., Salzburger, W. & Berner, D. Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. *Mol. Ecol.* **21**, 2852-2862, doi:10.1111/j.1365-294X.2012.05509.x (2012).

14 Gagnaire, P.-A., Pavey, S. A., Normandeau, E. & Bernatchez, L. The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish pairs assessed by RAD-sequencing. *Evolution* **in press**, doi:doi: 10.1111/evo.12075 (2013).

15 Kitano, J. *et al.* A role for a neo-sex chromosome in stickleback speciation. *Nature* **461**, 1079-1083 (2009).

16 Ellegren, H. *et al.* The genomic landscape of species divergence in Ficedula flycatchers. *Nature* **491**, 756-760, doi:10.1038/nature11584 (2012).

17 Dasmahapatra, K. K. *et al.* Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94-98, doi:10.1038/nature11041 (2012).

18 Janousek, V. *et al.* Genome-wide architecture of reproductive isolation in a naturally occurring hybrid zone between Mus musculus musculus and M. m. domesticus. *Mol. Ecol.* **21**, 3032-3047, doi:10.1111/j.1365-294X.2012.05583.x (2012).

19 Schluter, D. & Conte, G. L. Genetics and ecological speciation. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 9955-9962 (2009).

20 Seehausen, O. *et al.* Speciation through sensory drive in cichlid fish. *Nature* **455**, 620-626 (2008).

21 Feder, J. L. *et al.* Allopatric genetic origins for sympatric host-plant shifts and race formation in Rhagoletis. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 10314-10319 (2003).

22 Nadeau, N. J. *et al.* Genome-wide patterns of divergence and gene flow across a butterfly radiation. *Mol. Ecol.* **22**, 814-826, doi:10.1111/j.1365-294X.2012.05730.x (2013).

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

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

25 Bierne, N., Roze, D. & Welch, J. J. Pervasive selection or is it…? why are F_{ST} outliers sometimes so frequent? *Molecular Ecology Resources* **22**, 2061–2064 (2013).

26 Michel, A. P. *et al.* Widespread genomic divergence during sympatric speciation. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 9724-9729, doi:10.1073/pnas.1000939107 (2010).

27 Noor, M. A. F. & Bennett, S. M. Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity* **103**, 439-444 (2009).

28 Scally, A. *et al.* Insights into hominid evolution from the gorilla genome sequence. *Nature* **483**, 169-175, doi:10.1038/nature10842 (2012).

29 Nosil, P. *et al.* Genomic consequences of multiple speciation processes in a stick insect. *Proc. R. Soc. B-Biol. Sci.* **279**, 5058-5065, doi:10.1098/rspb.2012.0813 (2012).

30 Parchman, T. L. *et al.* The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. *Mol. Ecol.* **in press**, doi:10.1111/mec.12201 (2013).

31 Elmer, K. R. & Meyer, A. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol. Evol.* **26**, 298-306, doi:10.1016/j.tree.2011.02.008 (2011).

32 Egan, S. P., Nosil, P. & Funk, D. J. Selection and genomic differentiation during ecological speciation: isolating the contributions of host association via a comparative genome scan of Neochlamisus bebbianae leaf beetles. *Evolution* **62**, 1162-1181, doi:10.1111/j.1558-5646.2008.00352.x (2008).

33 Nosil, P., Egan, S. P. & Funk, D. J. Heterogeneous genomic differentiation between walking-stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection. *Evolution* **62**, 316-336 (2008).

34 Campbell, D. & Bernatchez, L. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology and Evolution* **21**, 945-956 (2004).

35 Garrigan, D. *et al.* Genome sequencing reveals complex speciation in the Drosophila simulans clade. *Genome Research* **22**, 1499-1511, doi:10.1101/gr.130922.111 (2012).

36 Excoffier, L., Hofer, T. & Foll, M. Detecting loci under selection in a hierarchically structured population. *Heredity* **103**, 285-298 (2009).

37 Coyne, J. A. & Orr, H. A. ''Patterns of speciation in Drosophila'' revisited. *Evolution* **51**, 295-303 (1997).

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

39 Nosil, P., Parchman, T. L., Feder, J. L. & Gompert, Z. Do highly divergent loci reside in genomic regions affecting reproductive isolation? A test using next-generation sequence data in Timema stick insects. *BMC Evol. Biol.* **12**, doi:164 10.1186/1471-2148-12-164 (2012).

40 Fisher, R. A. *The Genetical Theory of Natural Selection*. (Oxford Univ. Press, 1930).

41 Orr, H. A. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution* **52**, 935-949, doi:10.2307/2411226 (1998).

42 Albert, A. Y. K. *et al.* The genetics of adaptive shape shift in stickleback: Pleiotropy and effect size. *Evolution* **62**, 76-85 (2008).

43 Lexer, C., Buerkle, C. A., Joseph, J. A., Heinze, B. & Fay, M. F. Admixture in European Populus hybrid zones makes feasible the mapping of loci that contribute to reproductive isolation and trait differences (September 20, 2006). *Heredity* **98**, 183-184, doi:10.1038/sj.hdy.6800922 (2007).

44 Remington, D. & Purugganan, M. Candidate genes, quantitative trait loci, and functional trait evolution in plants. *International Journal of Plant Sciences* **164**, S7-S20 (2003).

45 Flint, J., Valdar, W., Sagiv, S. & Mott, R. Strategies for mapping and cloning quantitative trait genes in rodents. *Nature Reviews Genetics* **6**, 271-286 (2005).

46 Weinig, C. *et al.* Heterogeneous selection at specific loci in natural environments in Arabidopsis thaliana. *Genetics* **165**, 321–329 (2003).

47 Rogers, S. M. *et al.* Genetic signature of adaptive peak shift in threespine stickleback. *Evolution* **66**, 2439-2450 (2012).

48 Gordo, I. & Campos, P. R. A. Evolution of clonal populations approaching a fitness peak. *Biology Letters* **9**, 20120239 (2013).

49 Yeaman, S. & Whitlock, M. C. THE GENETIC ARCHITECTURE OF ADAPTATION UNDER MIGRATION-SELECTION BALANCE. *Evolution* **65**, 1897-1911, doi:10.1111/j.1558-5646.2011.01269.x (2011).

50 Orr, H. A. The genetic theory of adaptation: a brief history. *Nature Reviews Genetics* **6**, 119-127 (2005).

51 Guillaume, F. & Whitlock, M. C. Effects of migration on the genetic covariance matrix. *Evolution* **61**, 2398-2409, doi:10.1111/j.1558-5646.2007.00193.x (2007).

52 Leary, G. P. *et al.* Single mutation to a sex pheromone receptor provides adaptive specificity between closely related moth species. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 14081-14086, doi:10.1073/pnas.1204661109 (2012).

53 Maan, M. E. & Seehausen, O. Ecology, sexual selection and speciation. *Ecology Letters* **14**, 591-602, doi:10.1111/j.1461-0248.2011.01606.x (2011).

54 Hall, M. C. & Willis, J. H. Divergent selection on flowering time contributes to local adaptation in Mimulus guttatus populations. *Evolution* **60**, 2466-2477 (2006).

55 Bradshaw, H. D. & Schemske, D. W. Allele substitution at a flower colour locus produces a pollinator shift in monkey flowers. *Nature* **426**, 176-178 (2003).

56 Weigel, D. Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiology* **158**, 2-22 (2012).

57 Klahre, U. *et al.* Pollinator Choice in Petunia Depends on Two Major Genetic Loci for Floral Scent Production. *Current Biology* **21**, 730-739, doi:10.1016/j.cub.2011.03.059 (2011).

58 Malek, T. B., Boughman, J. W., Dworkin, I. & Peichel, C. L. Admixture mapping of male nuptial colour and body shape in a recently formed hybrid population of threespine stickleback. *Mol. Ecol.* **21**, 5265-5279, doi:10.1111/j.1365-294X.2012.05660.x (2012).

59 Quattrocchio, F. *et al.* Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. *The Plant Cell Online* **11**, 1433-1444 (1999).

60 Laturney, M. & Moehring, A. J. The genetic basis of female mate preference and species isolation in Drosophila. *International Journal of Evolutionary Biology*, 1-13 (2012).

61 Niehuis, O. *et al.* Behavioural and genetic analyses of Nasonia shed light on the evolution of sex pheromones. *Nature* **494**, 345-348 (2013).

62 Lassance, J.-M. *et al.* Functional consequences of sequence variation in the pheromone biosynthetic gene pgFAR for Ostrinia moths. *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1208706110 (2013).

63 Edelaar, P., Siepielski, A. M. & Clobert, J. Matching Habitat Choice Causes Directed Gene Flow: a Neglected Dimension in Evolution and Ecology. *Evolution* **62**, 2462-2472 (2008).

64 Verzijden, M. N. *et al.* The impact of learning on sexual selection and speciation. *Trends in Ecology and Evolution* **27**, 511-519 (2012).

65 Bomblies, K. & Weigel, D. Arabidopsis and relatives as models for the study of genetic and genomic incompatibilities. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 1815-1823 (2010).

66 Wright, K. M., Lloyd, D., Lowry, D. B., Macnair, M. R. & Willis, J. H. Indirect Evolution of Hybrid Lethality Due to Linkage with Selected Locus in Mimulus guttatus. *Plos Biology* **11**, doi:e1001497 10.1371/journal.pbio.1001497 (2013).

67 Mihola, O., Trachtulec, Z., Vlcek, C., Schimenti, J. C. & Forejt, J. A Mouse Speciation Gene Encodes a Meiotic Histone H3 Methyltransferase. *Science* **323**, 373-375, doi:10.1126/science.1163601 (2009).

68 Brideau, N. J. *et al.* Two Dobzhansky-Muller genes interact to cause hybrid lethality in Drosophila. *Science* **314**, 1292–1295 (2006).

69 Turner, L. M. & Hoekstra, H. E. Reproductive protein evolution within and between species: maintenance of divergent ZP3 alleles in Peromyscus. *Mol. Ecol.* **17**, 2616-2628, doi:10.1111/j.1365-294X.2008.03780.x (2008).

70 Buerkle, C. A. & Lexer, C. Admixture as the basis for genetic mapping. *Trends Ecol. Evol.* **23**, 686-694, doi:10.1016/j.tree.2008.07.008 (2008).

71 Barrett, R. D. H., Rogers, S. M. & Schluter, D. Natural selection on a major armor gene in threespine stickleback. *Science* **322**, 255-257 (2008).

72 Barrett, R. D. H., Rogers, S. M. & Schluter, D. Environment specific pleiotropy facilitates divergence at the Ectodysplasin locus in threespine stickleback. *Evolution* **63**, 2831–2837 (2009).

73 Servedio, M. R., Van Doorn, G. S., Kopp, M., Frame, A. M. & Nosil, P. Magic traits in speciation: 'magic' but not rare? *Trends Ecol. Evol.* **26**, 389-397, doi:10.1016/j.tree.2011.04.005 (2011).

74 Kronforst, M. R. *et al.* Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *Proceedings of the National Academy of Sciences USA* **103**, 6575-6580 (2006).

75 Abbott, R. *et al.* Hybridization and speciation. *Journal of Evolutionary Biology* **26**, 229-246, doi:10.1111/j.1420-9101.2012.02599.x (2013).

76 Arnold, S. J., Burger, R., Hohenlohe, P. A., Ajie, B. C. & Jones, A. G. Understanding the Evolution and Stability of the G-Matrix. *Evolution* **62**, 2451-2461 (2008).

77 Lande, R. QUANTITATIVE GENETIC-ANALYSIS OF MULTIVARIATE EVOLUTION, APPLIED TO BRAIN - BODY SIZE ALLOMETRY. *Evolution* **33**, 402-416, doi:10.2307/2407630 (1979).

78 Schluter, D. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766-1774, doi:10.2307/2410734 (1996).

79 Leimar, O. Environmental and genetic cues in the evolution of phenotypic polymorphism. *Evolutionary Ecology* **23**, 125-135 (2009).

80 Hansen, T. F. & Houle, D. Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology* **21**, 1201-1219, doi:10.1111/j.1420-9101.2008.01573.x (2008).

81 Chenoweth, S. F., Rundle, H. D. & Blows, M. W. The Contribution of Selection and Genetic Constraints to Phenotypic Divergence. *American Naturalist* **175**, 186-196, doi:10.1086/649594 (2010).

82 Roff, D. A. & Fairbairn, D. J. A TEST OF THE HYPOTHESIS THAT CORRELATIONAL SELECTION GENERATES GENETIC CORRELATIONS. *Evolution* **66**, 2953-2960, doi:10.1111/j.1558-5646.2012.01656.x (2012).

83 Jones, A. G., Arnold, S. J. & Bürger, R. Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* **57**, 1747-1760 (2003).

84 Björklund, M., Husby, A. & Gustafsson, L. Rapid and unpredictable changes in the G-matrix in a natural bird population over 25 years. *Journal of Evolutionary Biology* **26**, 1-13 (2013).

85 Agrawal, A. F., Brodie, E. D. & Rieseberg, L. H. Possible consequences of genes of major effect: transient changes in the G-matrix. *Genetica* **112**, 33-43 (2001).

86 Scoville, A., Lee, Y. W., Willis, J. H. & Kelly, J. K. Contribution of chromosomal polymorphisms to the G-matrix of Mimulus guttatus. *New Phytologist* **183**, 803-815 (2009).

87 Jones, A. G., Arnold, S. J., Bürger, R., Hohenlohe, P. A. & Uyeda, J. C. The effects of stochastic and episodic movement of the optimum on the evolution of the G-matrix and the response of the trait mean to selection. *Journal of Evolutionary Biology* **25**, 2210-2231 (2012).

88 Innocenti, P. & Chenoweth, S. F. Interspecific divergence of transcription networks along lines of genetic variance in Drosophila: dimensionality, evolvability and constraint. *Molecular Biology and Evolution in press*, Innocenti P, Chenoweth SF. 2013. Interspecific divergence of

transcription networks along lines of genetic variance in Drosophila: dimensionality, evolvability and constraint. Molecular Biology and Evolution in press (2013).

89 Mallarino, R. *et al.* Closely related bird species demonstrate flexibility between beak morphology and underlying developmental programs. *Proceedings of the National Academy of Sciences USA* **109**, 16222-16227. (2012).

90 McGuigan, K., Van Homrigh, A. & Blows, M. W. Genetic analysis of female preference functions as function-valued traits. *American Naturalist* **172**, 194-202, doi:10.1086/588075 (2008).

91 Lande, R. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology* **22**, 1435-1446 (2009).

92 Eroukhmanoff, F. & Svensson, E. I. Evolution and stability of the G-matrix during the colonization of a novel environment. *Journal of Evolutionary Biology* **24**, 1363-1373, doi:10.1111/j.1420-9101.2011.02270.x (2011).

93 Stelkens, R. B., Schmid, C., Selz, O. & Seehausen, O. Phenotypic novelty in experimental hybrids is predicted by the genetic distance between species of cichlid fish. *BMC Evol. Biol.* **9**, doi:283 10.1186/1471-2148-9-283 (2009).

94 Renaut, S. *et al.* Genome-wide patterns of divergence during speciation: the lake whitefish case study. *Philosophical Transactions of the Royal Society B-Biological Sciences* **367**, 354-363, doi:10.1098/rstb.2011.0197 (2012).

95 Nolte, A. W. & Tautz, D. Understanding the onset of hybrid speciation. *Trends in Genetics* **26**, 54-58, doi:10.1016/j.tig.2009.12.001 (2010).

96 Boughman, J. W. How sensory drive can promote speciation. *Trends Ecol. Evol.* **17**, 571-577 (2002).

97 Presgraves, D. C. The molecular evolutionary basis of species formation. *Nature Reviews Genetics* **11**, 175-180, doi:10.1038/nrg2718 (2010).

98 Stelkens, R. B., Young, K. A. & Seehausen, O. The Accumulation of Reproductive Incompatibilities in African Cichlid Fish. *Evolution* **64**, 617-632 (2010).

99 Crespi, B. & Nosil, P. Conflictual speciation: species formation via genomic conflict. *Trends Ecol. Evol.* **28**, 48-57, doi:10.1016/j.tree.2012.08.015 (2013).

100 Frank, S. A. Divergence of meiotic drive-suppression systems as an explanation for sex-biased hybrid sterility and inviability. *Evolution* **45**, 262-267 (1991).

101 Hurst, L. D. & Pomiankowski, A. Causes of sex ratio bias may account for unisexual sterility in hybrids: a new explanation of Haldane's rule and related phenomena. *Genetics* **128**, 841-858 (1991).

102 Cocquet, J. *et al.* A genetic basis for a postmeiotic X Versus Y chromosome intragenomic conflict in the mouse. *PLOS Genetics* **8**, e1002900 (2012).

103 Malik, H. S. in *Progress in Molecular and Subcellular Biology* Vol. 48 (ed D. Ugarovic) 33-52 (Springer, 2009).

104 Campbell, P., Good, J. M. & Nachman, M. W. Meiotic sex chromosome inactivation is disrupted in sterile hybrid male house mice. *Genetics* **193**, 819-828 (2013).

105 Rebollo, R., Horard, B., Hubert, B. & Vieira, C. Jumping genes and epigenetics: Towards new species. *Gene* **454**, 1-7 (2010).

106 Watson, E. T. & Demuth, J. P. in *Speciation: Natural Processes, Genetics and Biodiversity* (ed P. Michalak) (Nova Science Publishers, 2013).

107 Burton, R. S. & Barreto, F. S. A disproportionate role for mtDNA in Dobzhansky–Muller incompatibilities? *Mol. Ecol.* **21**, 4942-4957 (2012).

108 Roberts, R. B., Ser, J. R. & Kocher, T. D. Sexual Conflict Resolved by Invasion of a Novel Sex Determiner in Lake Malawi Cichlid Fishes. *Science* **326**, 998-1001, doi:10.1126/science.1174705 (2009).

109 Bomblies, K. *et al.* Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLOS Biology* **5**, e236 (2007).

110 Phadnis, N. & Orr, H. A. A Single Gene Causes Both Male Sterility and Segregation Distortion in Drosophila Hybrids. *Science* **323**, 376-379, doi:10.1126/science.1163934 (2009).

111 Bull, J. J. *Evolution of Sex-determining Mechanisms.* . (Benjamin/Cumings, 1983).

112 Macaya-Sanz, D. *et al.* Genetic analysis of post-mating reproductive barriers in hybridizing European Populus species. *Heredity* **107**, 478-486 (2011).

113 Silver, L. The peculiar journey of a selfish chromosome: mouse t-haplotypes and meiotic drive. *Trends in Genetics* **9**, 250-254 (1993).

114 Larracuente, A. M. & Presgraves, D. C. The selfish segregation distorter gene complex of Drosophila melanogaster. *Genetics* **192**, 33-53 (2012).

115 Saetre, G. P. *et al.* Sex chromosome evolution and speciation in Ficedula flycatchers. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 53-59 (2003).

116 Saether, S. A. *et al.* Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science* **318**, 95-97 (2007).

117 Qvarnström, A. & Bailey, R. I. Speciation through evolution of sex-linked genes. *Heredity* **102**, 4-15 (2009).

118 Seehausen, O., van Alphen, J. J. M. & Lande, R. Color polymorphism and sex ratio distortion in a cichlid fish as an incipient stage in sympatric speciation by sexual selection. *Ecology Letters* **2**, 367-378 (1999).

119 van Doorn, G. S. & Kirkpatrick, M. Turnover of sex chromosomes induced by sexual conflict. *Nature* **449**, 909-912 (2007).

120 Charlesworth, D. & Charlesworth, B. Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genet. Res.* **35**, 205–214 (1980).

121 Lande, R., Seehausen, O. & van Alphen, J. J. M. Mechanisms of rapid sympatric speciation by sex reversal and sexual selection in cichlid fish. *Genetica* **112**, 435-443 (2001).

122 Matute, D. R., Butler, I. A., Turissini, D. A. & Coyne, J. A. A Test of the Snowball Theory for the Rate of Evolution of Hybrid Incompatibilities. *Science* **329**, 1518-1521, doi:10.1126/science.1193440 (2010).

123 Nolte, V., Pandey, R. V., Kofler, R. & Schlötterer, C. Genome-wide patterns of natural variation reveal strong selective sweeps and ongoing genomic conflict in Drosophila mauritiana. *Genome Research* **23**, 99-110 (2013).

124 Smadja, C. M. & Butlin, R. K. A framework for comparing processes of speciation in the presence of gene flow. *Mol. Ecol.* **20**, 5123-5140, doi:10.1111/j.1365-294X.2011.05350.x (2011).

125 Bierne, N., Welch, J., Loire, E., Bonhomme, F. & David, P. The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol. Ecol.* **20**, 2044-2072, doi:10.1111/j.1365-294X.2011.05080.x (2011).

126 Bank, C., Burger, R. & Hermisson, J. The Limits to Parapatric Speciation: Dobzhansky-Muller Incompatibilities in a Continent-Island Model. *Genetics* **191**, 845-U345, doi:10.1534/genetics.111.137513 (2012).

127 Barton, N. H. MULTILOCUS CLINES. *Evolution* **37**, 454-471, doi:10.2307/2408260 (1983).

128 Barton, N. H. & de Cara, M. A. R. THE EVOLUTION OF STRONG REPRODUCTIVE ISOLATION. *Evolution* **63**, 1171-1190, doi:10.1111/j.1558-5646.2009.00622.x (2009).

129 Nosil, P., Harmon, L. J. & Seehausen, O. Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol.* **24**, 145-156, doi:10.1016/j.tree.2008.10.011 (2009).

130 Butlin, R. K. & Ritchie, M. G. Pulling together or pulling apart: hybridization in theory and practice. *Journal of Evolutionary Biology* **26**, 294-298, doi:10.1111/jeb.12080 (2013).

131 Servedio, M. R. & Noor, M. A. F. The role of reinforcement in speciation: Theory and data. *Annual Review of Ecology Evolution and Systematics* **34**, 339-364, doi:10.1146/annurev.ecolsys.34.011802.132412 (2003).

132 Kirkpatrick, M. & Barton, N. Chromosome inversions, local adaptation and speciation. *Genetics* **173**, 419-434, doi:10.1534/genetics.105.047985 (2006).

133 Felsenstein, J. Skepticism Towards Santa Rosalia, or Why Are There So Few Kinds of Animals. *Evolution* **35**, 124-138 (1981).

134 Teeter, K. C. *et al.* THE VARIABLE GENOMIC ARCHITECTURE OF ISOLATION BETWEEN HYBRIDIZING SPECIES OF HOUSE MICE. *Evolution* **64**, 472-485, doi:10.1111/j.1558-5646.2009.00846.x (2010).

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

136 Bolnick, D. I. *et al.* Phenotype-Dependent Native Habitat Preference Facilitates Divergence between Parapatric Lake and Stream Stickleback. *Evolution* **63**, 2004-2016 (2009).

137 Merrill, R. M., Van Schooten, B., Scott, J. A. & Jiggins, C. D. Pervasive genetic associations between traits causing reproductive isolation in Heliconius butterflies. *Proc. R. Soc. B-Biol. Sci.* **278**, 511-518, doi:10.1098/rspb.2010.1493 (2011).

138 Hopkins, R. & Rausher, M. D. Pollinator-Mediated Selection on Flower Color Allele Drives Reinforcement. *Science* **335**, 1090-1092, doi:10.1126/science.1215198 (2012).

139 Hopkins, R. & Rausher, M. D. Identification of two genes causing reinforcement in the Texas wildflower Phlox drummondii. *Nature* **469**, 411-+, doi:10.1038/nature09641 (2011).

140 Lowry, D. B. & Willis, J. H. A Widespread Chromosomal Inversion Polymorphism Contributes to a Major Life-History Transition, Local Adaptation, and Reproductive Isolation. *Plos Biology* **8**, doi:e1000500 10.1371/journal.pbio.1000500 (2010).

141 Joron, M. *et al.* Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **477**, 203-U102, doi:10.1038/nature10341 (2011).

142 Shaw, K. L. & Lesnick, S. C. Genomic linkage of male song and female acoustic preference QTL underlying a rapid species radiation. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 9737-9742 (2009).

143 Saetre, G. P. & Saether, S. A. Ecology and genetics of speciation in Ficedula flycatchers. *Mol. Ecol.* **19**, 1091-1106 (2010).

144 Bimova, B. V. *et al.* Reinforcement selection acting on the European house mouse hybrid zone. *Mol. Ecol.* **20**, 2403-2424, doi:10.1111/j.1365-294X.2011.05106.x (2011).

145 Bolnick, D. I. & Fitzpatrick, B. M. Sympatric speciation: Models and empirical evidence. *Annual Review of Ecology Evolution and Systematics* **38**, 459-487 (2007).

146 Haller, B. C., De Leon, L. F., Rolshausen, G., Gotanda, K. M. & Hendry, A. P. Magic traits: distinguishing the important from the trivial. *Trends Ecol. Evol.* **27**, 4-5, doi:10.1016/j.tree.2011.09.005 (2012).

147 Presgraves, D. C. Speciation Genetics: Search for the Missing Snowball. *Current Biology* **20**, R1073-R1074 (2010).

148 Rieseberg, L. H. & Blackman, B. K. Speciation genes in plants. *Annals of Botany* **106**, 439-455, doi:10.1093/aob/mcq126 (2010).

149 Moyle, L. C. & Nakazato, T. Hybrid incompatibility “snowballs“ between Solanum species. *Science* **329**, 1521-1523 (2010).

150 Stadler, T., Florez-Rueda, A. M. & Paris, M. Testing for "Snowballing" Hybrid Incompatibilities in Solanum: Impact of Ancestral Polymorphism and Divergence Estimates. *Molecular Biology and Evolution* **29**, 31-34, doi:10.1093/molbev/msr218 (2012).

151 Rieseberg, L. H. & Burke, J. M. A genic view of species integration - Commentary. *Journal of Evolutionary Biology* **14**, 883-886 (2001).

152 Bikard, D. *et al.* Divergent evolution of duplicate genes leads to genetic incompatibilities within A. thaliana. *Science* **323**, 623-626 (2009).

153 Masly, J. P., Jones, C. D., Noor, M. A. F., Locke, J. & Orr, H. A. Gene transposition as a cause of hybrid sterility in Drosophila. *Science* **313**, 1448-1450 (2006).

154 Ting, C. T. *et al.* Gene duplication and speciation in Drosophila: Evidence from the Odysseus locus. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12232-12235, doi:10.1073/pnas.0401975101 (2004).

155 Coyne, J. A. GENETICS AND SPECIATION. *Nature* **355**, 511-515 (1992).

156 Orr, H. A. The genetic basis of reproductive isolation: Insights from Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 6522-6526, doi:10.1073/pnas.0501893102 (2005).

157 Lande, R. Models of Speciation by Sexual Selection on Polygenic Traits. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **78**, 3721-3725 (1981).

158 Panhuis, T. M., Butlin, R., Zuk, M. & Tregenza, T. Sexual selection and speciation. *Trends Ecol. Evol.* **16**, 364-371 (2001).

159 Price, T. *Speciation in Birds*. (Roberts & Company, 2008).

160 Andersson, M. & Simmons, L. W. Sexual selection and mate choice. *Trends Ecol. Evol.* **21**, 296-302 (2006).

161 van der Sluijs, I. *et al.* A test of genetic association among male nuptial coloration,female mating preference, and male aggression bias within a polymorphic population of cichlid fish. *Current Zoology* **59**, in press (2013).

162 Price, T. D. & Bouvier, M. M. The evolution of F-1 postzygotic incompatibilities in birds. *Evolution* **56**, 2083-2089 (2002).

163 Orr, H. A. & Turelli, M. The evolution of postzygotic isolation: Accumulating Dobzhansky-Muller incompatibilities. *Evolution* **55**, 1085-1094, doi:10.1554/0014-3820(2001)055[1085:teopia]2.0.co;2 (2001).

164 Butlin, R. K., Bridle, J. R. & Schluter, D. *Speciation and Patterns of Diversity*. 333 (Cambridge University Press, 2009).

165 Sanderson, M. J. & Donoghue, M. J. Reconstructing shifts in diversification rates on phylogenetic trees. *Trends Ecol. Evol.* **11**, 15-20 (1996).

166 Rosenblum, E. B. *et al.* Goldilocks Meets Santa Rosalia: An Ephemeral Speciation Model Explains Patterns of Diversification Across Time Scales. *Evolutionary Biology* **39**, 255-261, doi:10.1007/s11692-012-9171-x (2012).

167 Seehausen, O., vanAlphen, J. J. M. & Witte, F. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277**, 1808-1811 (1997).

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

169 Vonlanthen, P. *et al.* Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature* **482**, 357-U1500, doi:10.1038/nature10824 (2012).

170 Seehausen, O., Takimoto, G., Roy, D. & Jokela, J. Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Mol. Ecol.* **17**, 30-44 (2008).

171 Sun, S., Ting, C. T. & Wu, C. I. The normal function of a speciation gene, Odysseus, and its hybrid sterility effect. *Science* **305**, 81-83, doi:10.1126/science.1093904 (2004).

172 Bazin, E., Dawson, K. J. & Beaumont, M. A. Likelihood-Free Inference of Population Structure and Local Adaptation in a Bayesian Hierarchical Model. *Genetics* **185**, 587-602, doi:10.1534/genetics.109.112391 (2010).

173 Lawson, D. J., Hellenthal, G., Myers, S. & Falush, D. Inference of Population Structure using Dense Haplotype Data. . *PLoS Genetics* **8**, e1002453, doi:10.1371/journal.pgen.1002453 (2012).

174 Reed, R. D. *et al.* optix Drives the Repeated Convergent Evolution of Butterfly Wing Pattern Mimicry. *Science* **333**, 1137-1141, doi:10.1126/science.1208227 (2011).

175 Etges, W. J., de Oliveira, C. C., Noor, M. A. F. & Ritchie, M. G. GENETICS OF INCIPIENT SPECIATION IN DROSOPHILA MOJAVENSIS. III. LIFE-HISTORY DIVERGENCE IN ALLOPATRY AND REPRODUCTIVE ISOLATION. *Evolution* **64**, 3549-3569, doi:10.1111/j.1558-5646.2010.01096.x (2010).

176 Orr, H. A. The Population-Genetics of Speciation - the Evolution of Hybrid Incompatibilities. *Genetics* **139**, 1805-1813 (1995).

177 Slate, J. Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Mol. Ecol.* **14**, 363-379 (2005).

178 Gompert, Z. & Buerkle, C. A. Bayesian estimation of genomic clines. *Mol. Ecol.* **20**, 2111-2127, doi:10.1111/j.1365-294X.2011.05074.x (2011).

179 Trier, C. N., Hermansen, J. S., Sætre, G. P. & Bailey, R. I. Hybrid speciation through discordant cline movement of mito-nuclear and sex-linked incompatibilities. *PLoS Genetics* **in review** (2013).

180 Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. Mapping complex disease traits with global gene expression. *Nature Reviews Genetics* **10**, 184-194 (2009).

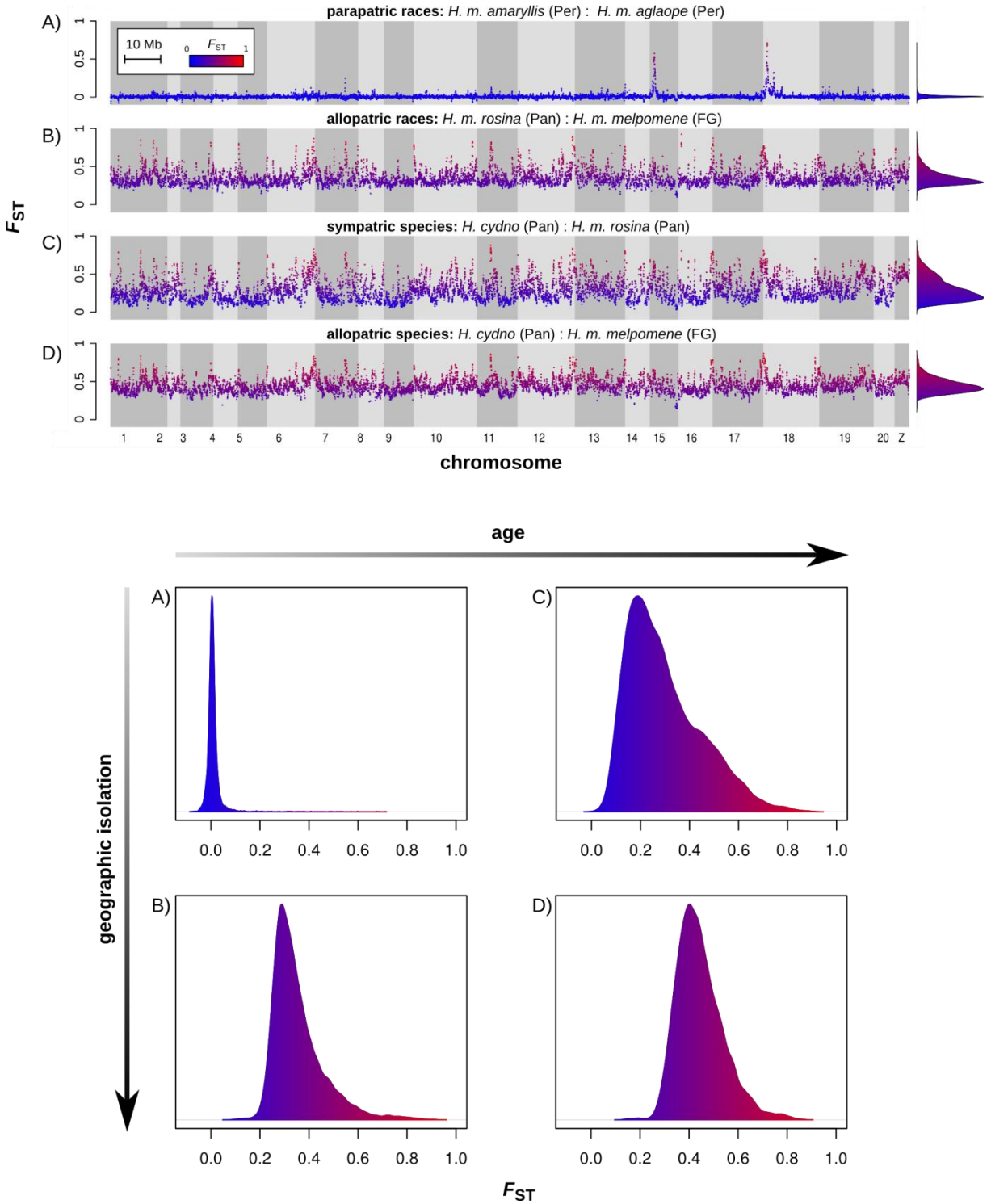


Fig. 1. Genomic patterns of divergence along the speciation continuum in *Heliconius* butterflies. The top panel shows the patterns of differentiation between hybridising races (a) and species (c) and between geographically isolated races (b) and species (d) along the genome. The pattern of divergence is highly heterogeneous even between allopatric populations of the same species (b). The bottom panels show the frequency distribution of locus-specific F_{ST} values for the four comparisons. The shape of these distributions offers a means to visualise the joint influence of selection, drift and gene flow, and the challenge is to distinguish between speciation with gene flow (a, c) versus isolation (b,d).

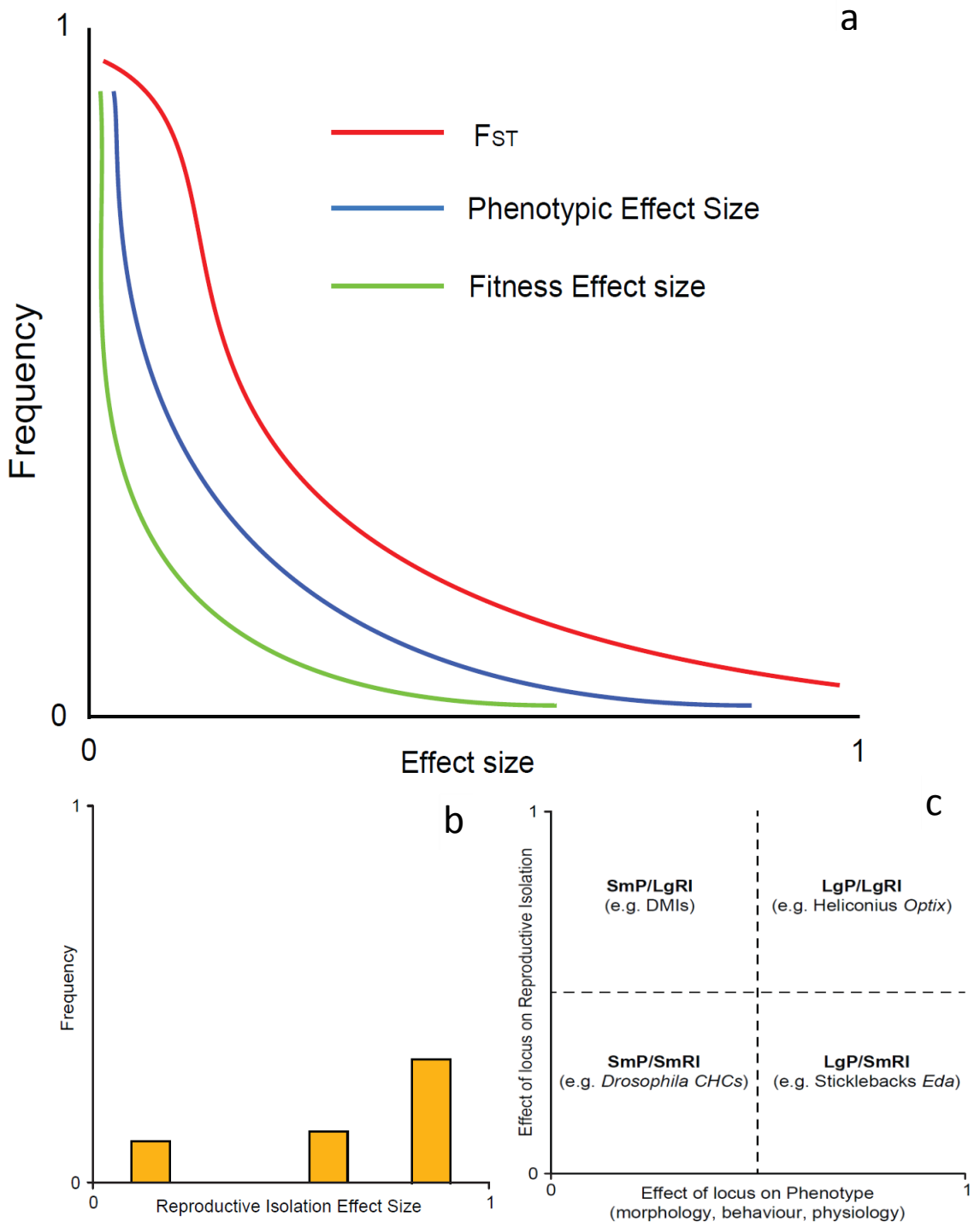


Fig. 2. **Genes of large effect do not necessarily have large impacts on RI.** a) Conceptual distributions of F_{ST} and effect sizes based on patterns seen in empirical data: frequency distribution of F_{ST} values as typically seen in genome scans (red), phenotypic effect sizes from QTL studies (blue), and fitness effect sizes from field experiments (green). b) The frequency distribution of effect sizes on reproductive isolation based on the very few published data available for natural populations. More of these data are sorely needed. c) The lack of correlation between the effect of a locus on phenotype (P) and on reproductive isolation (RI). In this latter panel, an example for each of the four relationships is shown to illustrate that phenotypic effect size does not predict RI effect size: loci with small effect on phenotype and large effect on reproductive isolation (SmP/LgRI: DMIs); loci with large effect on phenotype and large effect on reproductive isolation (LgP/LgRI: *Optix* in *Heliconius* [174]); loci with small effect on phenotype and small effect on reproductive isolation (SmP/SmRI: CHCs in *Drosophila*[175]); loci with large effect on phenotype and small effect on reproductive isolation (LgP/SmRI: *Eda* in *Gasterosteus*[38]). The relationships between phenotypic and RI effect size and F_{ST} are also unclear at present.

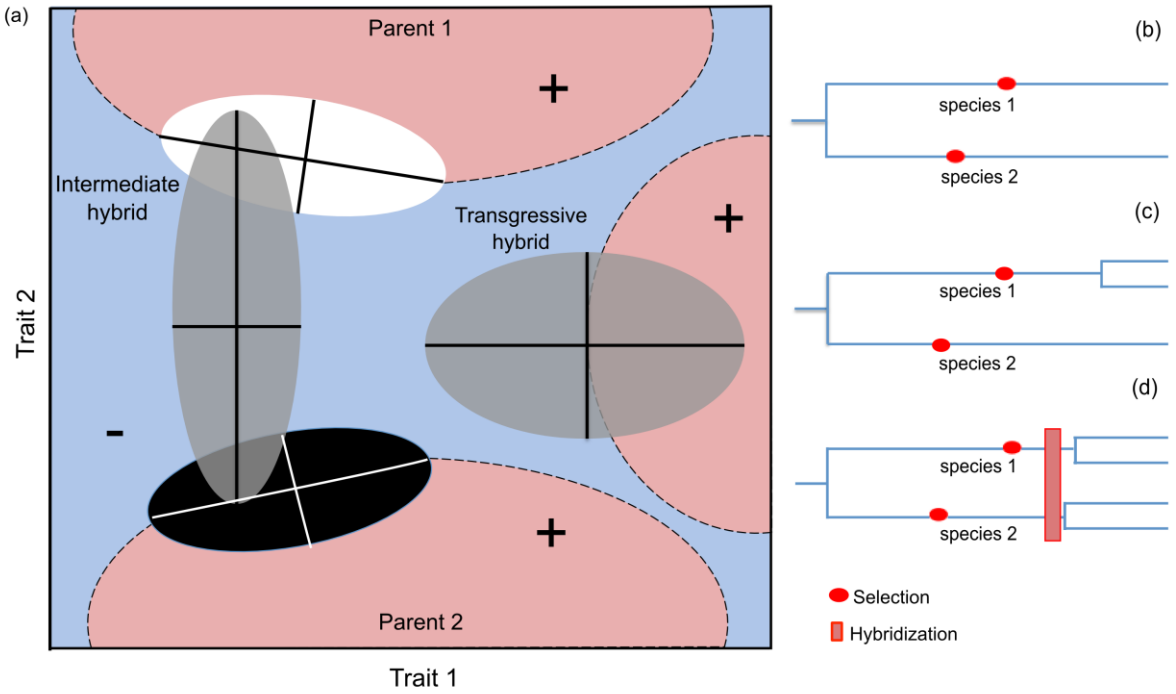


Fig. 3. Influence of genetic constraints on speciation

(a) G matrices are represented as ellipses in the space of two quantitative traits, and the adaptive landscape is represented by regions of higher (+; red) and lower (-, blue) fitness than the parental populations. Hybridization events can facilitate speciation by aligning the G-matrix in the direction of divergence between parental species (intermediate hybrid), or by giving rise to novel phenotypes (transgressive hybrid) in new regions of positive fitness that cannot be reached through gradual evolution in either of the parental species.

(b-d) The influence of genetic constraints on adaptive divergence and speciation can be seen at the phylogenetic level. (b) Constraints may persist over evolutionary time as a result of the inability of selection to change genetic architecture, restricting speciation. (c) Alternatively, other forms of selection can alter the structure and orientation of the G-matrix and potentially facilitate divergence and speciation over moderate time scales. (d) Hybridization and gene flow can dramatically alter G in just a few generations, fueling adaptive divergence and resulting in sudden bursts of speciation. Note that hybridization between sister species is shown here for illustration, but hybridization that facilitates divergence may occur more widely among related taxa.

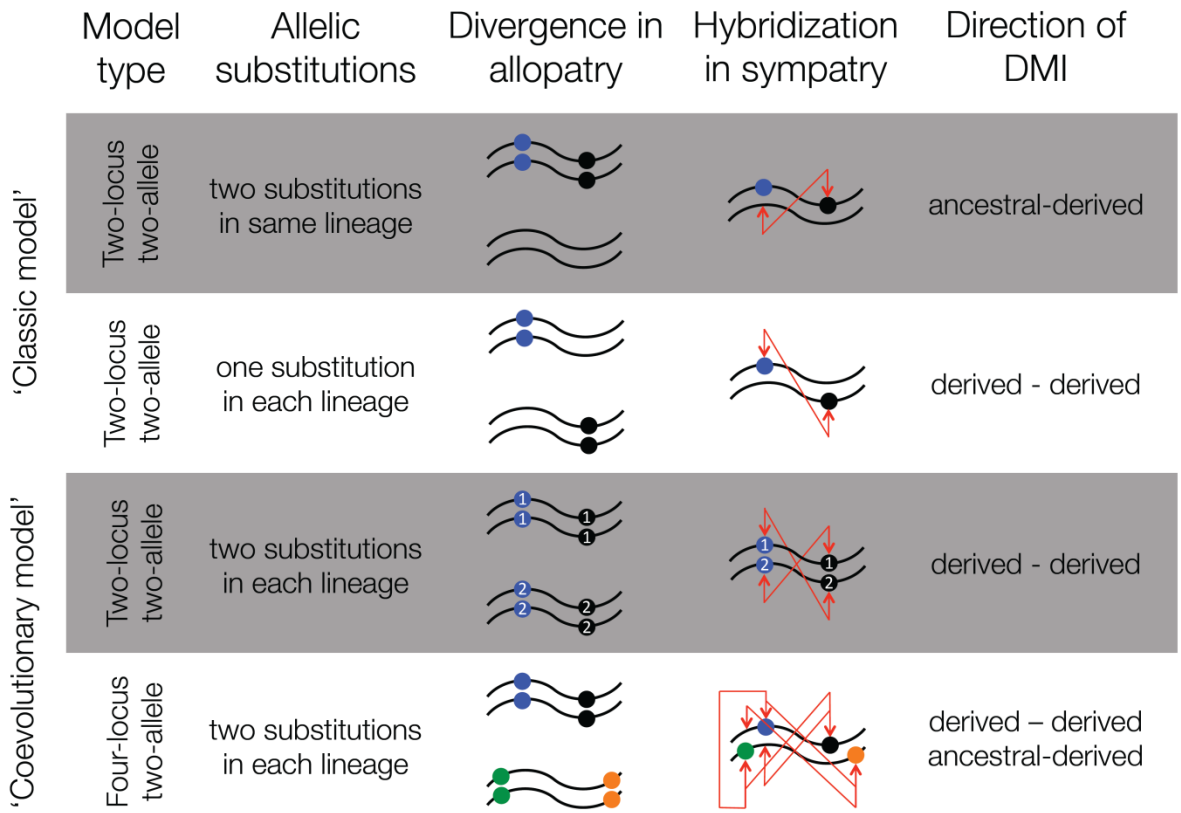
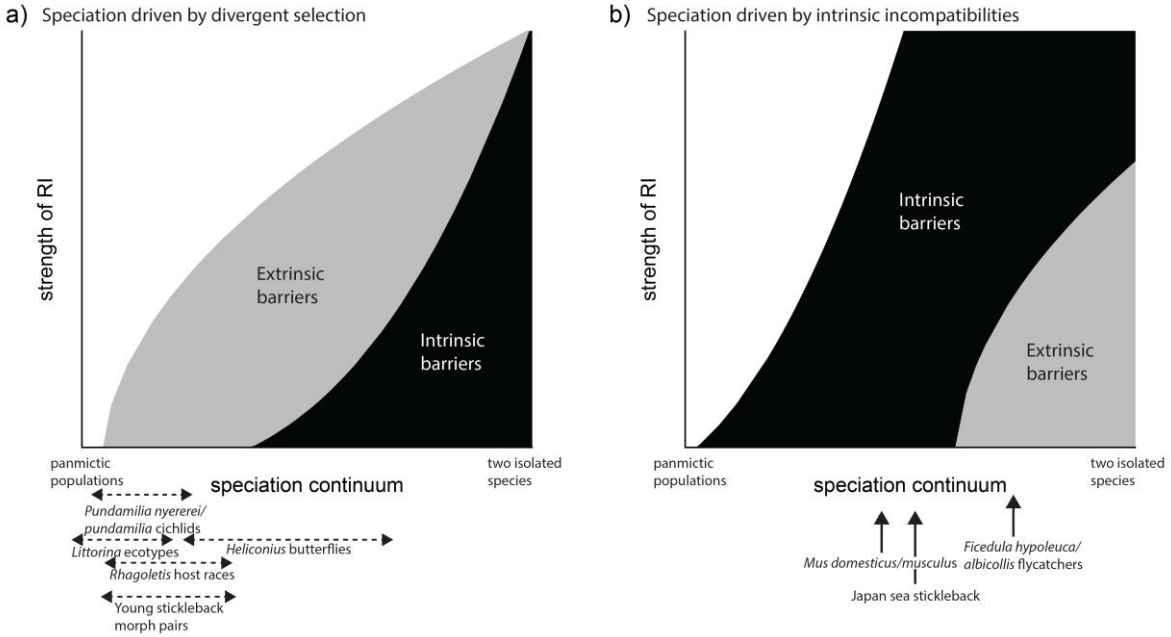


Fig 4. 'Classic' and coevolutionary models of hybrid incompatibility. In all examples with two substitutions in a lineage, the locus on the left (selfish) is fixed before the locus on the right (restorer). Model 2 is a special case that can refer to maternal-effect selfish loci since maternal "poison" and zygotic "rescue" are due to developmental expression divergence of the same locus. Red arrows indicate negative epistatic interactions between loci. In all models, the ancestral state is wild-type except for row three. In this row, the ancestral state is a coevolving selfish element-restorer system. Two-locus, two-allele models are often cited as a simple and general model for the evolution of Dobzhansky-Muller incompatibilities (DMIs). However, insight into the role of genomic conflict in speciation reveals the potential for further development of models of hybrid incompatibility. The above classical scenarios describe cases of genomic conflict driven speciation, yet models incorporating the possibility for unlimited change due to ongoing coevolution reveal the potential for additional incompatibilities. Additional theoretical work is needed in this regard.

Box 1. Extrinsic and intrinsic reproductive isolation evolve at different relative rates during speciation driven by divergent selection (Panel A) compared to speciation driven by intrinsic isolating mechanisms (Panel B). In both panels the x-axis depicts the position of a diverging taxon pair on the "speciation continuum" (in terms of relative time) and the y-axis represents the strength of reproductive isolation (RI) between sister taxa. In speciation driven by divergent selection, the evolution of extrinsic isolating barriers is expected to precede the evolution of intrinsic isolating barriers. In speciation driven by intrinsic isolating mechanisms, the evolution of intrinsic barriers leads to reproductive isolation, and extrinsic isolating barriers may accumulate simultaneously or later, facilitating ecological coexistence between sibling species. Curve shapes are hypothetical, and reflect the idea that in speciation driven by divergent selection, extrinsic barriers arise rapidly under divergent selection pressure early in speciation. In contrast, speciation driven by intrinsic barriers often results from epistatic incompatibilities, which are thought to often (though not necessarily always[147]) accumulate in an accelerating "snowball" fashion[163,176]. Intrinsic and extrinsic barriers within each panel are not necessarily additive or interactive, and the emergence of reproductive isolation via either of these classes of barriers should be viewed as independent trajectories. Movement along the speciation continuum, from weakly isolated species to strongly isolated ones, is not constant, and the average timescales for speciation via the processes contrasted here (Panels A & B) may vary. Arrows along the x-axis indicate the position(s) of model study systems (studied by the authors of this paper) along the speciation continuum. These organisms vary in the strength and types of barriers isolating populations. Studies of the genomics of speciation at different points on the speciation continuum are emerging in several systems where speciation is driven by divergent selection (as indicated by the dashed arrows showing timespans along the speciation continuum). In many cases strong reproductive isolation may never evolve, particularly in ecological speciation (e.g. [129]). Incomplete reproductive isolation may facilitate cases of "speciation reversal" (e.g.[169]) and "ephemeral" speciation (e.g. [166]). Additional theory and empirical studies on a wider range of taxa and across a greater range of points along the speciation continuum within individual taxa are needed to transform the qualitative generalizations depicted in the figure to quantitative and statistically testable predictions. For instance, we predict that speciation driven by intrinsic isolating mechanisms might consistently involve greater divergence at centromeres and sex chromosomes that are more subject to intragenomic conflict. In contrast, divergence in ecological speciation may be more distributed across autosomes. Future work should seek to determine whether different mechanistic processes and modes of speciation can be distinguished based on patterns observed from genomic data.



Box 2: Genomic tools for identifying RI-genes

Next-generation, high-throughput sequencing is rapidly expanding the tools available for identifying genes contributing to reproductive isolation and the feasibility of such studies in a wide range of organisms. For example, the power of classic approaches, such as genotype-phenotype association mapping, has dramatically increased with the advent of NGS, and researchers can now collect and analyze genome-wide datasets of high resolution at ever decreasing costs.

RAD tag sequencing, exome- and whole genome re-sequencing of incipient species pairs along the speciation continuum is currently a logical first step in the search for candidate RI genes[7,9,11,16]. Such genome scans can reveal genomic regions that show evidence of divergent selection between incipient species pairs using F_{ST}-outlier analysis or related approaches. The latest methods can take good account of demographic and other sources of variation (e.g. [172]). Genetic maps can now be generated more easily and are essential for full interpretation of genome scans. In some cases, it is possible to use a reference genome from a related model organism, such as *Drosophila*, *Arabidopsis*, house mouse, zebrafish, or chicken to identify putative candidate RI genes in these genomic regions. However, most genome scans do not specifically link the genomic regions that experience divergent selection with phenotypes, including those that contribute to RI.

As a complement to genome scans, a range of genetic mapping tools are available to identify links between specific genomic regions and the phenotypic traits that contribute to RI. Quantitative trait locus (QTL) mapping is one powerful such method[177]. In short, a genome-wide set of markers is genotyped in a phenotypically variable population with known pedigree to statistically associate QTLs with phenotypes of interest (in this case traits associated with RI). With functional information on genes in the vicinity of a QTL, candidate RI genes can be identified. If pedigree data are not available, it is also possible to take advantage of the phenotypic and genetic differences that exist between hybridizing taxa and use admixture as the basis for genetic mapping of phenotypes that contribute to RI[58,70].

Intrinsic and extrinsic postzygotic barriers involve genes that are selected against in F1 hybrids and admixed individuals. A variety of methods can be used to identify genes under negative selection (i.e. candidate RI genes) across a hybrid zone or in other situations where admixture occurs (e.g. laboratory crossing experiments). Genomic cline analysis [178] is one such method in which a Bayesian model quantifies locus-specific patterns of introgression with parameters that describe the probability of locus-specific ancestry as a function of genome-wide admixture. Candidate RI loci with low levels of introgression relative to most of the genome can thus be identified[179].

To further investigate the potential significance of candidate RI-loci various gene expression studies can be useful. A promising method is expression QTL (eQTL) analysis, which identifies genomic loci that regulate expression levels of mRNAs and proteins[180]. Systematically generated eQTL information could provide insight into a biological basis for reproductive isolation identified through genome-wide association studies, and can help to identify networks of genes, and the role of gene interaction (including epistasis in DMIs) in reproductive isolation.

Finally, for some groups of organisms advanced tools in experimental genetics are available that allows knockout or transgenic inserts of candidate RI genes to study their phenotypic effects[11,97].