Genomics and the origin of species

Ole Seehausen, Roger K. Butlin, Irene Keller, Catherine E. Wagner, Janette W. Boughman, Paul A. Hohenlohe, Catherine L. Peichel and Glenn-Peter Saetre* et al.

Abstract | Speciation is a fundamental evolutionary process, the knowledge of which is crucial for understanding the origins of biodiversity. Genomic approaches are an increasingly important aspect of this research field. We review current understanding of genome-wide effects of accumulating reproductive isolation and of genomic properties that influence the process of speciation. Building on this work, we identify emergent trends and gaps in our understanding, propose new approaches to more fully integrate genomics into speciation research, translate speciation theory into hypotheses that are testable using genomic tools and provide an integrative definition of the field of speciation genomics.

Reproductive isolation
The absence or restriction of gene flow between populations beyond that caused by spatial separation.

Gene flow
The movement of alleles between populations. For gene flow to occur, individuals must disperse between populations and successfully reproduce with local individuals. Therefore, gene flow can be reduced not only by dispersal barriers but also by either intrinsic or extrinsic reproductive isolation.

**Department of Fish Ecology and Evolution, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland; and Division of Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland. Correspondence to O.S. e-mail: ole.seehausen@eawag.ch*
doi:10.1038/nrg3644

Major insights into the genetics of speciation have come from various approaches (BOX 1), which range from the mapping of individual genes that cause reproductive isolation to the characterization of genome-wide differentiation patterns, and from quantitative genetic approaches to admixture analyses that associate phenotypes with reduced gene flow between populations[1–3]. These empirical approaches have a long history that started with the work of Dobzhansky[4] and Muller[5]. The theoretical understanding of the genetics of speciation has advanced markedly[6–10]. However, the deluge of empirical data from next-generation sequencing (NGS), along with the emergence of new analytical approaches, necessitates the integration of this theoretical work to strengthen the conceptual foundations of the nascent field of speciation genomics. Such integration will help to elucidate the relationships between evolutionary processes and genomic divergence patterns on the one hand, and between genomic properties and speciation processes on the other hand, and it will help to unify research on both the ecological and non-ecological causes of speciation.

In this Review, we first discuss areas in which genomic approaches have begun to make important contributions to speciation research (BOX 1), which include elucidating patterns and rates of genome-wide divergence, improving our understanding of both the genomic basis and the evolution of intrinsic and extrinsic reproductive barriers, and identifying mechanisms by which different barriers become genomically coupled. We also highlight areas that would benefit from further attention, such as the distributions of locus effect sizes, pleiotropy and genomic constraint. We conclude by discussing how NGS data and innovative population genomic analyses — which use genome-wide data to make inferences about evolutionary processes in natural populations — could contribute to further progress in integrating these study areas into a more comprehensive and coherent understanding of speciation genomics.

Speciation: theory and classical evidence

In line with others[1,3], we define speciation as the origin of reproductive barriers among populations that permit the maintenance of genetic and phenotypic distinctiveness of these populations in geographical proximity. These reproductive barriers can be initiated either by divergent selection (that is, 'ecological' or sexual selection that creates extrinsic reproductive isolation) or by the evolution — through genetic drift, as an indirect consequence of selection, or through genomic conflict — of genetic incompatibilities that cause intrinsic reproductive isolation (BOX 2). The study of the accumulation of intrinsic isolation has a strong tradition in evolutionary biology[1,11]. However, most recent population genomic studies of divergence across the genomes of incipient and sister species have investigated cases of putative ecological speciation and have focused on divergent adaptation and extrinsic isolation (but see REF. 12, discussed below, for an important role for genomic conflict in generating reproductive incompatibilities).

Extrinsic postzygotic isolation arises as a consequence of either divergent or disruptive selection when the viability or the fertility of migrants or of individuals with intermediate genotypes is reduced[2]. Prezygotic sexual isolation and also extrinsic postzygotic isolation (when hybrids have reduced mating success[13]) may evolve as

The contributors

This Review was initiated during a workshop funded by the European Science Foundation networking programme Frontiers in Speciation Research (FroSpects). O.S. organized and coordinated the workshop, and led the preparation of the manuscript with assistance from C.E.W., I.K. and R.K.B. These four authors constitute the core writing team. J.W.B., P.A.H., C.L.P., G.-P.S., C.E.W. and I.K. led discussion groups and initial drafting of sections of the Review. E.T.W., C.D.J., C.S.C., S.H.M., J.W.B., J.S. and C.E.W. drafted the figures. O.S. and R.K.B. drafted the general sections. Other authors contributed during the workshop and commented on drafts.

Affiliations of additional core writing team and discussion leaders

- Roger K. Butlin is at the Department of Animal and Plant Sciences, the University of Sheffield, Sheffield S10 2TN, UK; and the Sven Lovén Centre — Tjärnö, University of Gothenburg, S-452 96 Strömstad, Sweden.
- Irene Keller is at the Department of Fish Ecology and Evolution, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland; the Division of Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland; and the Institute of Integrative Biology, ETH Zürich, ETH Zentrum CHN, 8092 Zürich, Switzerland.
- Catherine E. Wagner is at the Department of Fish Ecology and Evolution, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland; and the Division of Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland.
- Janette W. Boughman is at the Department of Fish Ecology and Evolution, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland; and the Department of Zoology; Ecology, Evolutionary Biology and Behavior Program; BEACON Center, Michigan State University, 203 Natural Sciences, East Lansing, Michigan 48824, USA.
- Paul A. Hohenlohe is at the Department of Biological Sciences, Institute of Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, Idaho 83844–3051, USA.
- Catherine L. Peichel is at the Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA.
- Glenn-Peter Saetre is at the Department of Biosciences, Centre for Ecological and Evolutionary Synthesis, University of Oslo, PO BOX 1066, Blindern, N-0316 Oslo, Norway.

Affiliations of additional authors (listed alphabetically)

- Claudia Bank is at the School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland.
- Åke Brännström is at the Integrated Science Laboratory and the Department of Mathematics and Mathematical Statistics, Umeå University, 90187 Umeå, Sweden.
- Alan Brelsford is at the Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland.
- Chris S. Clarkson is at the Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK.
- Fabrice Eroukhmanoff is at the Department of Biosciences, Centre for Ecological and Evolutionary Synthesis, University of Oslo, PO BOX 1066, Blindern, N-0316 Oslo, Norway.
- Jeffrey L. Feder is at the Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556–0369 USA.
- Martin C. Fischer is at the Institute of Integrative Biology, ETH Zürich, ETH Zentrum CHN, 8092 Zürich, Switzerland.
- Andrew D. Foote is at the Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, DK-1350 Copenhagen, Denmark. Present address: the Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, SE-752 36 Uppsala, Sweden.
- Paolo Franchini is at Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz, 78457 Konstanz, Germany.
- Chris D. Jiggins is at the Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK.
- Felicity C. Jones is at the Friedrich Miescher Laboratory of the Max Planck Society, 72076 Tübingen, Germany.
- Anna K. Lindholm is at the Institute of Evolutionary Biology and Environmental Studies, University of Zurich, CH-8057 Zurich, Switzerland.
- Kay Lucek is at the Department of Fish Ecology and Evolution, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland; and the Division of Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland.
- Martine E. Maan is at the Behavioural Biology Group, Centre for Behaviour and Neurosciences, University of Groningen, PO BOX 11103, 9700 CC Groningen, The Netherlands.
- David A. Marques is at the Department of Fish Ecology and Evolution, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland; the Division of Aquatic Ecology and Evolution, and the Computational and Molecular Population Genetics Laboratory, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland.
- Simon H. Martin is at the Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK.
- Blake Matthews is at the Department of Aquatic Ecology, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland.
- Joana I. Meier is at the Department of Fish Ecology and Evolution, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland; the Division of Aquatic Ecology and Evolution, and the Computational and Molecular Population Genetics Laboratory, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland.
- Markus Möst is at the Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK; and the Department of Aquatic Ecology, Eawag: Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland.
- Michael W. Nachman is at the Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, California 94720–3160, USA.
- Etsuko Nonaka is at the Integrated Science Laboratory and Department of Ecology and Environmental Science, Umeå University, 90187 Umeå, Sweden.
- Diana J. Rennison is at the Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada.
- Julia Schwarzer is at the Department of Fish Ecology and Evolution, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland; the Division of Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland; and Zoologisches Forschungsmuseum Alexander Koenig, 53113 Bonn, Germany.
- Eric T. Watson is at the Department of Biology, The University of Texas at Arlington, 76010–0498 Texas, USA.
- Anja M. Westram is at the Department of Animal and Plant Sciences, the University of Sheffield, Sheffield S10 2TN, UK.
- Alex Widmer is at the Institute of Integrative Biology, ETH Zürich, ETH Zentrum CHN, 8092 Zürich, Switzerland.

Box 1 | **Genomic tools for studying speciation**

Next-generation sequencing is rapidly expanding the tool box for the study of speciation.

Patterns of genomic divergence
Several methods can be used to investigate genome-wide divergence along the 'speciation continuum'. These methods include genome scans using single-nucleotide polymorphism (SNP) arrays[78], restriction-site-associated DNA sequencing (RAD-seq)[72,77] or related genotyping-by-sequencing (GBS) methods, whole-exome or transcriptome sequencing[76] and whole-genome resequencing[112] of population samples. Patterns in genome-wide divergence can be visualized and compared using, for example, F_{ST} kernel density plots and Manhattan plots[98] (FIG. 1).

Testing for signatures of introgression
Various approaches are available to assess whether the sharing of genetic variants between incipient species is a result of hybridization or incomplete lineage sorting[90]. The ABBA–BABA test[183] is particularly applicable to genome-scale data sets. It relies on the frequencies of two specific patterns of allele sharing among a group of four species.

Identifying signatures of selection
Genome scans can reveal genomic regions that show evidence of divergent selection between incipient species using F_{ST}-outlier analyses or related approaches, which can be applied either to individual SNPs[77] or to smoothed average F_{ST} values[72] within windows (that is, regions of a defined size) of the genome. The latest methods can account for demographic variation and other sources of variation[104,184] and make improved use of high-density marker information[185].

Mapping genes that are involved in reproductive isolation
A logical first step in the search for candidate genes that are involved in reproductive isolation is to carry out genome scans of incipient species pairs at several different stages along the speciation continuum[69,72,74,98]. A range of genetic mapping tools are available for identifying links between divergent genomic regions and the phenotypic traits that contribute to reproductive isolation. Quantitative trait locus (QTL) mapping is one such method that is powerful for doing so[186]. In short, a genome-wide set of markers is genotyped in a phenotypically variable population that has known pedigree data, and statistical associations are identified between the genetic markers (in this case, QTLs) and phenotypes of interest (in this case, traits related to reproductive isolation). With functional information on genes that are in the vicinity of a QTL, candidate reproductive isolation genes can be identified.

Admixture mapping
If pedigree data are not available, then it is possible to take advantage of the phenotypic and genetic differences between hybridizing taxa and use admixture as the basis to genetically map phenotypes that contribute to reproductive isolation[108,187] using samples from wild hybrid populations. Both intrinsic and extrinsic postzygotic barriers involve alleles that are selected against in hybrids, and various methods can be used to identify such alleles in hybrid zones or in other situations in which admixture occurs. Genomic cline analysis[188] is one such method that can identify candidate reproductive isolation loci with low levels of introgression relative to most of the genome[79,189].

Manipulative selection experiments
Both QTL and admixture mapping have an unfortunate bias towards detecting loci of large effect[147]. Alternatively, alleles that affect fitness and reproductive isolation can be located using manipulative selection experiments, which track allelic changes or genome-wide responses[86,190]. Estimates of these effects can be ascertained by measuring selection and introgression in the wild. So far, few studies have taken this approach and none has measured effects on reproductive isolation.

Gene expression studies
To further investigate the functional importance of candidate loci that are involved in reproductive isolation, expression QTL (eQTL) analysis can be useful. It identifies genomic loci that regulate expression levels of mRNAs[191]. Systematically generated eQTL information can provide insights into the mechanism that underlies reproductive isolation in regions that have been identified through genome-wide association studies, and such eQTL information can help to identify networks of genes and the role of genetic interaction (including epistasis in Bateson–Dobzhansky–Muller incompatibilities) in reproductive isolation.

Speciation genomics
The field of speciation research that addresses the influence of genomic properties on the evolution of reproductive barriers and the signatures of speciation processes that are observable in genomic patterns (for example, processes of diversity and divergence). Its aim is a conceptual and methodological integration of genomic approaches with other empirical and theoretical speciation research.

Effect sizes
The magnitude of the influence of a locus or a specific allele on a phenotypic trait. This can be expressed, for example, as the proportion of phenotypic variation due to a specific locus or as the phenotypic difference between genotypes with and without a specific allele.

Pleiotropy
Effect of an allele on more than one trait.

Divergent selection
Selection that favours different phenotypes in different populations.

Extrinsic reproductive isolation
Fitness reduction in hybrids that is dependent on the environment and that is mediated by genotype–environment interactions.

Genomic conflict
Conflict that arises between genes or genetic elements within the same genome either when they are not transmitted by the same rules (for example, biparental versus uniparental inheritance) or when a gene causes its own transmission to the detriment of the rest of the genome. The presence of elements that bias transmission (that is, distorter loci) is expected to lead to the evolution of loci that restore Mendelian segregation (that is, restorer loci).

Intrinsic reproductive isolation
Fitness reduction in hybrids that is independent of the environment.

a consequence of divergent sexual selection[3,14] that is often, but not always, mediated by differences in environments[15,16]. Prezygotic sexual isolation and extrinsic postzygotic isolation are therefore dependent on genotype–environment interactions in the wider sense, in which mating partners are part of the external environment. By contrast, intrinsic postzygotic isolation is independent of the external environment. Consequently, different types of genes and gene networks, and different evolutionary processes may be involved in generating these classes of isolation. Extrinsic postzygotic isolation and sexual isolation can rapidly evolve[17], and they often interact with each other[16] and with the evolution of intrinsic postzygotic isolating barriers[18] (BOX 2). Selection can initiate speciation in situations both with and without gene flow between populations, whereas intrinsic incompatibilities are less likely to accumulate when gene flow is present[6]. However, adaptive divergence and ecological speciation are not the same. Divergent adaptation alone rarely causes sufficient reproductive isolation to allow the accumulation or the persistence of species differences in geographical proximity: this typically requires the evolution of prezygotic isolation[1,3] (BOX 2), although it is possible that this varies between major taxonomic groups such as insects compared with vertebrates or plants.

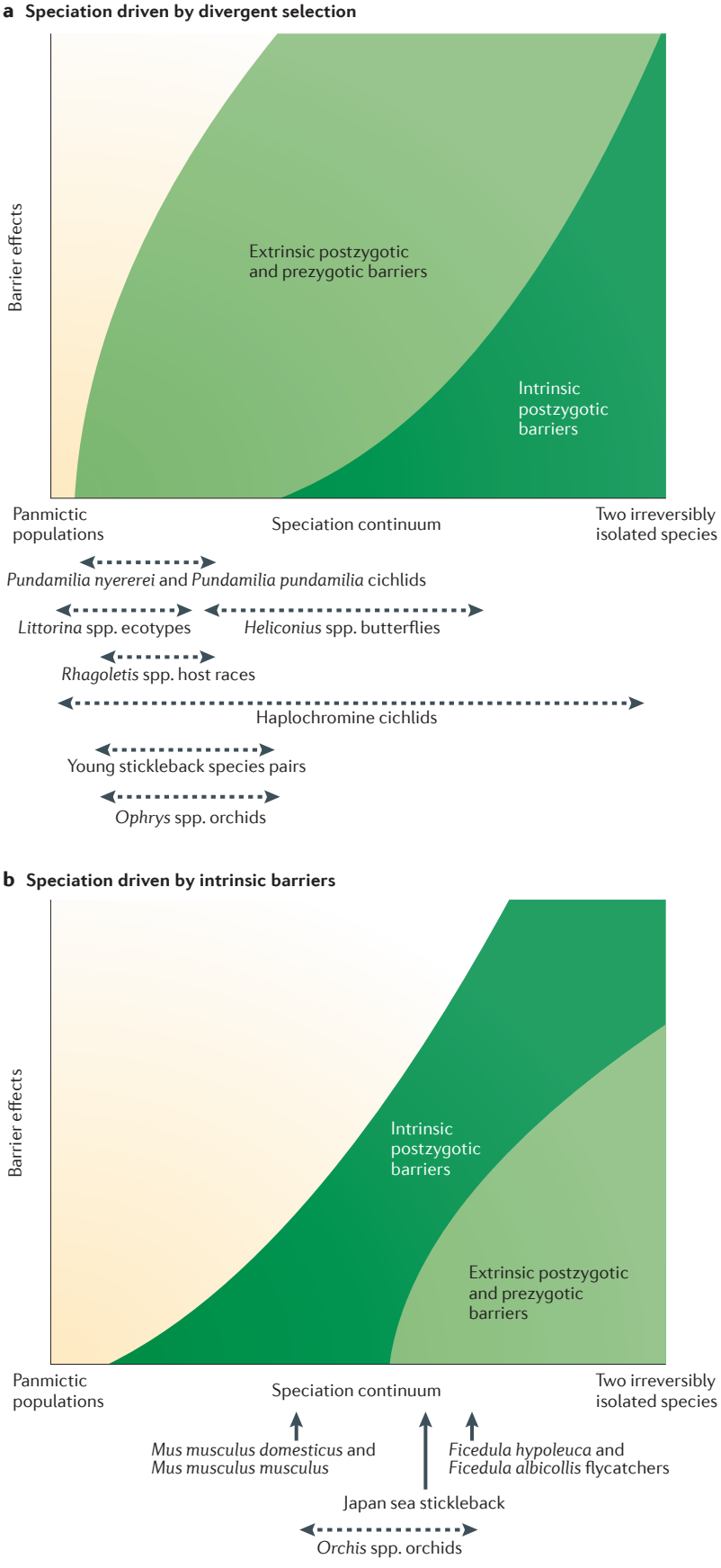
Box 2 | **Evolution of reproductive isolation**

Reproductive isolation can be usefully divided into three forms. First, extrinsic forms of postzygotic isolation result from divergent ecological or sexual selection and depend on interaction either with the environment or with other individuals; an example is reduced viability or fertility of migrants and hybrids due to ecological or behavioural factors. Second, intrinsic forms of postzygotic isolation are due to genetic incompatibilities that are independent of the environment (for example, Bateson–Dobzhansky–Muller incompatibilities). Third, prezygotic isolation includes phenological isolation, habitat isolation and sexual isolation that is due to assortative mating or fertilization.

In speciation driven by divergent ecological or sexual selection, extrinsic postzygotic and prezygotic barriers evolve first and often interact to produce reproductive isolation, and intrinsic postzygotic barriers will often only evolve later in the speciation process (see the figure, part **a**). By contrast, speciation driven by intrinsic barriers often results from epistatic incompatibilities, which may (although not necessarily[19]) accumulate in an accelerating 'snowball' manner[61,192] either as a by-product of selection or as a result of genetic drift (which only occurs slowly). Extrinsic postzygotic and prezygotic barriers may accumulate later, which facilitates both ecological coexistence between sibling species and reinforcement of reproductive isolation (see the figure, part **b**).

In each part of the figure, 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 between sister taxa. Shapes of the curves are hypothetical and reflect the idea that, in speciation driven by divergent selection, extrinsic postzygotic and prezygotic barriers rapidly arise early in the speciation process. Classes of barriers in each part of the figure are not necessarily additive or interactive, and the emergence of reproductive isolation through either of these barrier types should be viewed as independent trajectories. Movement along the speciation continuum — from weakly isolated species to irreversibly isolated ones — is not constant, and speciation can either go back and forth or be arrested at intermittent stages; the average timescales for speciation through the two processes contrasted in the figure may vary.

Arrows below the *x* axis in each panel (see the figure) indicate the positions of model systems along the speciation continuum that we have studied. These organisms vary in the strength and the types of barriers that isolate incipient and sister species. Studies of the genomics of speciation at different points in the speciation continuum are emerging in several systems, mainly where speciation is driven by divergent selection, as indicated by the dashed arrows that show time spans along the speciation continuum (see the figure). In many cases, particularly in ecological speciation[121], strong reproductive isolation may never evolve. Incomplete reproductive isolation may facilitate cases of 'speciation reversal' (REF. 193) and 'ephemeral' speciation[194].



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.

Postzygotic isolation
Effects of barriers that act after fertilization, such as hybrid sterility and hybrid inviability. It can be either extrinsic (that is, mediated by the environment) or intrinsic.

Disruptive selection
Selection within a single population that favours extreme phenotypes over intermediate phenotypes.

Sexual isolation
Reproductive isolation due to reduced mating between members of divergent populations, including behavioural assortative mate choice and assortative fertilization in animals, as well as pollinator-mediated assortative mating in plants. It is most often thought of as prezygotic but can also be postzygotic if there is disruptive sexual selection.

F_{ST}-outlier analyses
The comparison of the distribution of F_{ST} values across loci with the distribution expected in the absence of divergent selection for the same average differentiation. A locus with an F_{ST} value that exceeds expectation is likely to be influenced by divergent selection, either on the locus itself or on a linked locus.

Prezygotic isolation
Effect of barriers that act before or after mating but before fertilization, including the isolating effects of divergent mate choice, habitat preference, reproductive timing and gametic incompatibility.

Bateson–Dobzhansky–Muller incompatibilities
(BDMI). Intrinsic postmating barriers that are the result of epistatic interactions between alleles at two or more loci that reduce fitness in hybrids but not in the parental populations.

The available evidence suggests that negative epistatic interactions — Bateson–Dobzhansky–Muller incompatibilities (BDMIs; although hereafter we use the more common abbreviation, DMIs) — are the most frequent cause of intrinsic postzygotic isolation[1,19–21]. However, other mechanisms, including underdominance[22] and gene duplication, transposition and gene loss[23–25], may also cause intrinsic postzygotic isolation. The time course of the accumulation of DMIs is not well understood[19,26–28], and accumulation rates may vary among taxa and among mechanisms that underlie DMI evolution[19]. DMIs were long thought to arise as a consequence of genetic drift, as a result of stochastic deactivation of gene duplicates[29] or as a by-product of ecological selection[30]. However, theoretical considerations, such as the slow pace of neutral accumulation of barriers[31], and early empirical evidence for positive selection on loci that contribute to incompatibilities[32], suggested that genetic drift was unlikely to be a common source of incompatibilities. Instead, recent observations indicate that genomic conflict may be a common mechanism that drives DMI evolution[20,33–35] (BOX 3), as originally proposed in 1991 (REFS 34,35). Genomic conflict may arise from competing interests of males and females[36]; from meiotic drivers[37,38], mobile elements[39,40] or other 'selfish' genetic elements and their suppressors; and from competing interests between genomes of organelles and the nucleus[41,42]. Sexual conflict is thought to drive the evolution of new sex chromosomes[43,44], and empirical observations suggest that the turnover of sex chromosomes has a role in the evolution of reproductive isolation[45,46].

The different evolutionary mechanisms that underlie the build-up of both extrinsic and intrinsic postzygotic isolation and of prezygotic isolation suggest that their genomic signatures will also be distinct. The genomic architecture of extrinsic isolation is likely to resemble that of adaptive population divergence and is likely to be diverse and scattered across multiple regions in the genome (see below). However, there are theoretical arguments and empirical evidence that sites under selection in the genome will spatially cluster when adaptive evolution proceeds under prolonged bouts of divergent selection with either migration or recurrent hybridization[47]. For intrinsic isolation, incompatibility factors that are driven by genomic conflict are expected to accumulate in genomic regions of reduced recombination where linkage disequilibria between distorter loci and responder loci can become established[48,49]. Sex chromosomes are particularly susceptible to the accumulation of incompatibility factors that are derived from genomic conflict because these chromosomes are constantly in a 'battle' over segregation, whereas only small and tightly linked autosomal regions are in conflict with their homologues[34]. At the same time, there will be particularly strong selection for suppression of sex-linked distorter loci because they tend to bias sex ratios[50,51]. The genomic architecture of certain types of prezygotic isolation may also be influenced by regions of reduced recombination around sex-determining loci[52] or sex chromosomes[53], particularly when sex linkage resolves sexually antagonistic effects of sexual selection[54]. Alternatively, prezygotic isolation loci may accumulate near extrinsic ecological isolation loci (see the section below on genomic coupling of reproductive barriers). All of these signatures must be distinguished from background patterns of genetic diversity and divergence that depend on the populations' history of genetic drift, gene flow, background selection and episodes of positive selection that are unrelated to reproductive isolation.

The search for signatures in the genetic architecture of reproductive isolation has a long 'pre-genomic' history[55,56]. However, there has been a historical disconnection between research programmes that were focused on intrinsic isolation, which have typically concentrated on later stages of speciation[20,57], and those focused on extrinsic postzygotic isolation and prezygotic sexual isolation at early stages of speciation[2,30,15,16]. As a result of this disconnection, it is currently a challenge to compare the evolutionary rates of different components of reproductive isolation and their relevance to speciation. Such rates have been compared in the same taxon using pre-genomic methods[11,58–60], and the data suggest that prezygotic and extrinsic postzygotic isolation often evolve faster than intrinsic postzygotic isolation, which is consistent with expectations from classical theory[61]. The availability of genome-wide data will now permit testing of this pattern with a considerable increase in resolution.

Genomics and the 'speciation continuum'

After speciation is complete, populations accumulate differences as a result of mutation, genetic drift and ongoing selection. Therefore, reproductively isolated species often differ in traits that evolved under ecological selection and in others that evolved under sexual selection, and these species may also have intrinsic incompatibilities. A central task of speciation genetics is to reconstruct the sequence in which these different barriers originated in order to distinguish between causes and consequences of speciation. To achieve this, one would ideally take an unbiased view of the entire genome at all stages of the same speciation process. However, speciation can rarely be studied in real time in natural populations of sexually reproducing multicellular organisms. Estimates of variation among loci in the timing and the magnitude of gene flow could help to determine the order in which reproductive barriers emerged, but it is challenging to make such inferences, and current methods are not accurate enough for this purpose[62]. However, by integrating case studies of closely related taxa that vary in their extent of divergence (that is, the 'speciation continuum'), inferences can often be made about the chronology and the importance of different factors and processes involved.

Investigations of this speciation continuum have made important contributions to speciation research[63,64], and this approach is being adopted in NGS-based genome scan and transcriptome scan studies of speciation. The major questions being addressed are: to what extent is divergence at different stages of speciation localized in the genome (that is, the island view) and to what extent is it widespread? To what extent can heterogeneity in divergence be attributed to selective processes compared with genetic drift? What are the sources of

selection? Does genomic divergence tend to follow a common trajectory as it proceeds along the speciation continuum? And how are all of these affected by the extent of geographical isolation? A recently much cited scenario for speciation without strong geographical isolation, which is derived from earlier models[65,66], involves an early stage of divergence at which differentiation is limited to a small number of loci (that is, islands) that are under strong divergent selection. The size of these regions would gradually increase through the process of divergence hitchhiking, and the effective migration rate would eventually decrease globally across the genome, which gives rise to genome-wide divergence (that is, genomic hitchhiking)[67,68].

Genome scans of ecological speciation. Several NGS-based genome scans of the speciation continuum have found surprisingly variable patterns of genomic divergence. It seems that incipient species can quickly accumulate substantial divergence even in the presence of gene flow (FIG. 1). However, in some cases — such as those of *Heliconius* spp. butterflies[69], *Helianthus* spp. sunflowers[70] and poplar trees[71] — divergence between parapatric ecotype populations is limited to a few large genomic regions, whereas it is widespread across the genome in other cases[72–75]. NGS-based genome scans of sympatric sister species have generally reported genomically widespread and highly heterogeneous divergence that varies on a very local scale[75–81]. Few studies have looked for

Underdominance
Heterozygotic inferiority; that is, the phenotype expressed in heterozygotes has lower fitness than that of either homozygote. This can cause disruptive selection.

Meiotic drivers
Factors that distort Mendelian segregation. At a heterozygous site, the driving variant will be found in more than half of the gametes.

Sexual conflict
The evolution of phenotypic characteristics by sexual selection when the trait confers a fitness benefit on one sex but a fitness cost on the other.

Hybridization
Mating between individuals that belong to distinct species or populations. If postmating isolation is incomplete, hybridization leads to the introgression of genes from one population to another.

Linkage disequilibrium
The statistical association of the alleles at two loci within gametes in a population. Although linkage disequilibrium tends to be greater between linked loci, it can also arise between physically unlinked loci (for example, because of selection, nonrandom mating or gene flow).

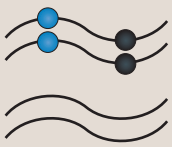
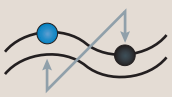
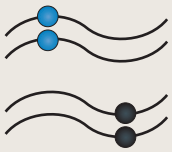
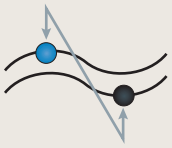
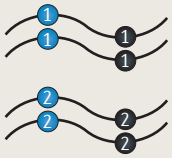
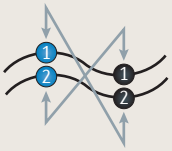
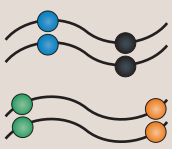
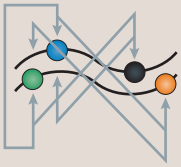
Distorter loci
Loci that underlie meiotic drive, which is the non-Mendelian segregation of alleles in meiosis. Distorter loci may act on other loci, so-called responder loci.

Responder loci
Loci that show deviations from Mendelian segregation (that is, meiotic drive) owing to the effect of distorter loci.

Speciation continuum
Variation of the strength of reproductive isolation between two incipient species either in different locations or in different species pairs that belong to the same evolutionary lineage and that diverge in similar ways.

Box 3 | **Models of hybrid incompatibility in a genomic conflict scenario**

In the classic model, Bateson–Dobzhansky–Muller incompatibilities (DMIs) are envisioned as two-locus, two-allele interactions, in which incompatibilities arise either between an ancestral allele and an allele that is derived in one lineage or between alleles that are derived in two separate lineages (see the table). Circles represent derived alleles. A special case of the model with separately derived alleles can refer to maternal-effect 'selfish' loci in which maternal 'poison' and zygotic 'antidote' are both due to divergence in developmental expression of the same locus. In co-evolutionary models, DMIs are continually fixed either at the same loci (that is, two-locus, two-allele) or at different loci (that is, four-locus, two-allele) (see the table). In all examples with two substitutions in a lineage, the selfish locus (left) drives the evolution of the restorer locus (right). Grey arrows indicate negative epistatic interactions between complementary loci. In all models, the ancestral state is wild type except for the two-locus two-allele co-evolutionary model. In this model, the ancestral state is a co-evolving selfish element-restorer system; numbers represent the lineages in which the derived alleles originated. Insights into the role of genomic conflict in speciation reveal the potential for further development of models of hybrid incompatibility. Models that incorporate the possibility for increased lag load due to ongoing co-evolution predict successively more severe incompatibilities as the lag load increases. Additional theoretical work is needed to investigate such co-evolutionary models.

Model type	Allelic substitutions	Divergence in allopatry	Hybridization in sympatry	Direction of DMI
Classic model				
Two-locus two-allele	Two substitutions in the same lineage			Ancestral–derived
Two-locus two-allele	One substitution in each lineage			Derived–derived
Co-evolutionary model				
Two-locus two-allele	Two substitutions in each lineage			Derived–derived
Four-locus two-allele	Two substitutions in each lineage			Derived–derived; ancestral–derived

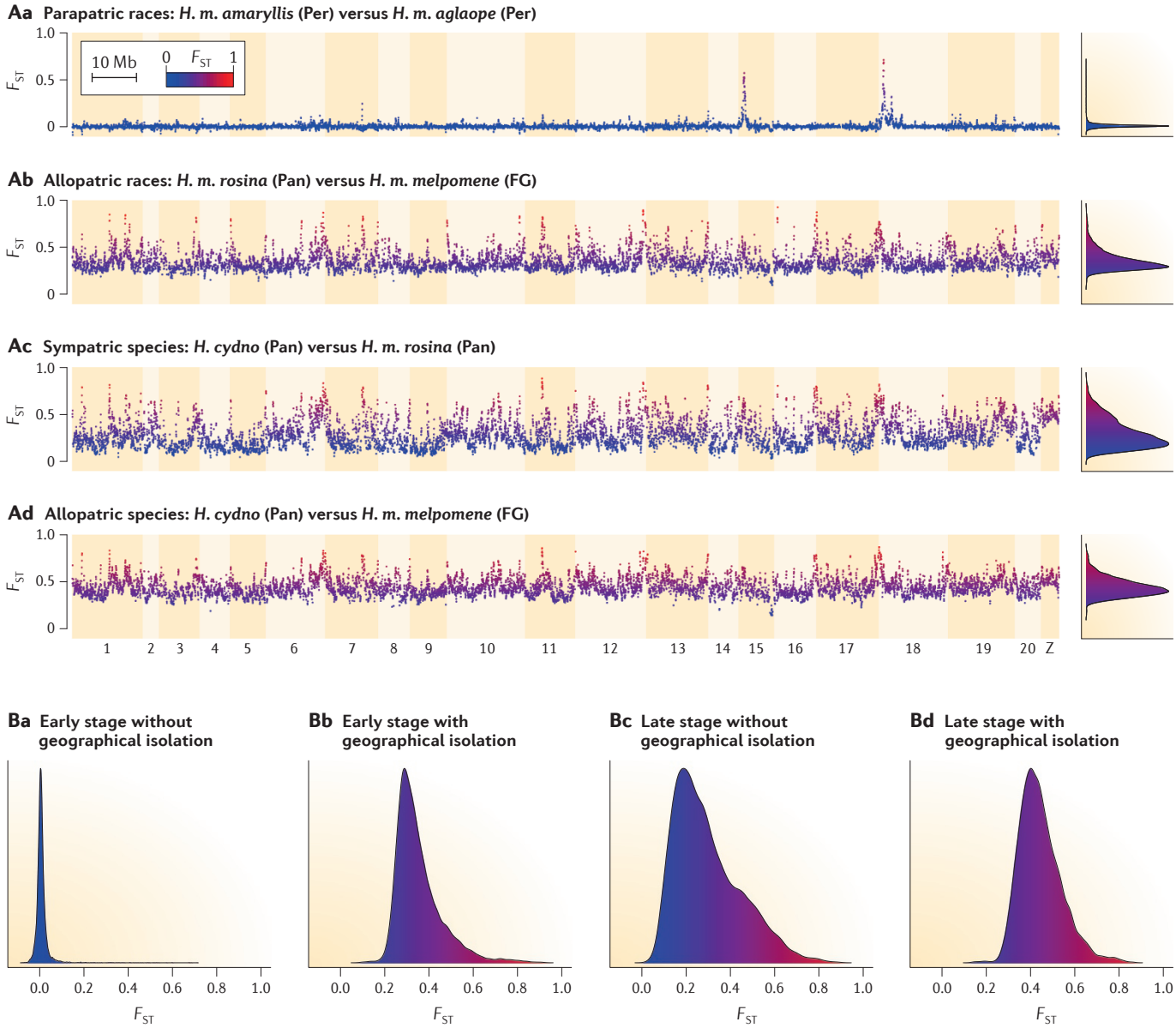


Figure 1 | **Genomic patterns of divergence along the 'speciation continuum' of *Heliconius* spp. butterflies.** The patterns of differentiation between hybridizing parapatric races (part **Aa**) and sympatric species (part **Ac**), as well as those between geographically isolated (that is, allopatric) races (part **Ab**) and species (part **Ad**), are shown along the genome; the *x* axes represent chromosome positions. Divergence is highly heterogeneous even between allopatric populations of the same species (part **Ab**). The shapes of the frequency distributions of locus-specific F_{ST} values (part **B**) clearly differ both between the different stages in the continuum and between geographical scenarios. For example, the greater variance is consistent with gene flow between species in sympatry (part **Bc**). However, the challenge is to distinguish between speciation with gene flow (parts **Ba**, **Bc**) and that without gene flow (parts **Bb**, **Bd**). All species shown are from the genus *Heliconius*, and all subspecies shown are from the species *Heliconius melpomene*. FG, French Guiana; Pan, Panama; Per, Peru. Figure is modified, with permission, from REF. 87 © (2013) Cold Spring Harbor Laboratory Press.

Genome scan
Comparison of genome-wide patterns of diversity within populations and/or divergence between populations at hundreds or thousands of markers. Until recently, most studies used amplified fragment length polymorphisms (AFLPs) but this has recently changed, and single-nucleotide polymorphisms (SNPs) generated by next-generation sequencing or SNP chips are being used.

evidence of divergence hitchhiking, and the available results are inconsistent[69,76,82]. Genome-wide average F_{ST} often increases as phenotypic divergence increases[80,83], but divergence seems to remain heterogeneous across the genome for a long time, which is potentially due to repeated episodes of interspecific gene flow even after reproductive isolation has become strong[84,85]. The first generation of NGS-based population genomic studies of ecological speciation has therefore shown that ecological selection can cause strong isolation of small genomic regions between diverging populations and that, when reproductive isolation is strong enough to permit persistence of incipient species in sympatry, many unlinked regions typically experience significant isolation.

So where does the heterogeneity in genomic divergence come from? It is commonly inferred to result

from locus-specific differences in the effects of divergent selection and gene flow. Indeed, genome scans have shown strong isolation at genomic loci that were known to be under divergent selection[64,69,70,72,74]. However, caution is warranted because different evolutionary processes can leave similar signatures in the genome. Heterogeneous genomic divergence is sometimes also observed between allopatric populations of the same species in the absence of any current gene flow[76,86,87] (FIG. 1). Indeed, many studies assume ongoing gene flow between species, even though stochastic variation due to recent coalescence times and incomplete lineage sorting can lead to low divergence and high heterogeneity in a similar way, particularly when they are combined with selection[88,89]. Statistical methods are available to distinguish divergence in isolation from that with gene flow, and these methods are increasingly being applied to genome-scale data sets (BOX 1; reviewed in REF. 90).

Even in the absence of selection, divergence is expected to vary owing to both the stochasticity of genetic drift and the complexities of population history, and this variation can be enhanced by confounding effects of genomic heterogeneity[91]. In particular, regions of low recombination and/or high gene density often show reduced intraspecific diversity, which inflates relative divergence as measured by either F_{ST} or D_a (that is, relative average divergence that is corrected for intraspecies diversity)[88]. This can result from background selection against deleterious mutations[92], intraspecific selective sweeps (in allopatry)[88] or even a direct influence of recombination on genetic diversity[93]. It is challenging to disentangle these processes[94]. Some have suggested correcting for recombination rate in the interpretation of F_{ST} patterns[83]. Others have suggested that absolute divergence measures such as D_{xy} are more robust to diversity artefacts[95], especially when they are corrected for local mutation rate[96]. It seems unlikely that any single parameter will reliably disentangle divergent selection and gene flow from neutral processes. Having a good knowledge of the geographical context of population divergence will help, but new parameter-rich modelling approaches[90] will be frequently required to distinguish between hypotheses of primary divergence with gene flow, secondary contact with hybridization and incomplete lineage sorting.

Adaptive divergence has been shown to accumulate preferentially in regions of low recombination[97], including the centres of chromosomes[83], the vicinity of centromeres[98], sites of inversions[74] and often (but not always[12,71]) sex chromosomes[98–100]. Heterogeneity in genomic divergence that is seen in allopatry might also result from gene-flow–selection balance that has occurred in the past[47,76]. Finally, the assumption that the baseline F_{ST} reflects neutral divergence may be violated in cases in which divergent selection is pervasive and multifarious, and this would bias against the detection of the selection signature[81].

There is evidence for repeated divergence of the same genes or the same genomic regions across replicate pairs of species or environmental contrasts, which strongly supports the idea that these regions are indeed involved in adaptation and/or reproductive isolation[72,74,85,97,101–103]. The detection of such parallel divergence may require dense sampling of genomes or transcriptomes because the highest levels of repeatability may be observed at the scale of genomic regions rather than that of individual genes or single-nucleotide polymorphisms (SNPs)[97]. In this case, the repeatability in the heterogeneity of genomic divergence may be due, at least partly, to shared genomic heterogeneity in both recombination and mutation rates rather than to parallel adaptive divergence, but the shared genomic structure may facilitate the repeated accumulation in the same genomic regions of adaptive differentiation[97]. Another approach involves combining classic cline theory with genome-wide analyses, which allows measurements of the strength of selection at specific loci[79] (BOX 1). In the future, parameter-rich coalescent models of divergence with gene flow fitted to genomic data may be able to account for the heterogeneity of demographic history across the genome when seeking to identify genomic regions that have reduced gene flow[84,104]. Finally, genome scans combined with manipulative selection[81], quantitative trait locus (QTL) mapping[82,105], candidate gene mapping[72,74] and admixture mapping[79,106–108] can be used to investigate whether divergent genomic regions contain loci that contribute to reproductive isolation.

Several recent studies have found a contribution of ancient alleles to recent divergence, as exemplified by sticklebacks[74,109], cichlids[77,110], *Rhagoletis* spp. flies[111] and *Heliconius* spp. butterflies[112]. Ancient alleles are identifiable owing to the accumulation of many substitutions or to the sharing across wide spatial or taxonomic ranges. The sources of such ancient allelic variation can be either standing genetic variation or hybridization[113]. It is difficult to distinguish between these hypotheses in practice because of the challenges of differentiating incomplete lineage sorting from hybridization[90] (BOX 1). The balance of evidence from NGS data implies introgressive hybridization rather than standing genetic variation as the source of ancient alleles in most of the above cases. Speciation in these cases might have been facilitated by hybridization that provides genetic material for both adaptation and reproductive isolation in the face of gene flow. Future research should test this hypothesis further by combining genomic and ecological approaches.

Genomic divergence and intrinsic isolation. Many studies have investigated DMI genes in strongly isolated species but, in many cases, it remained unclear whether the fixation of the underlying mutations was a cause or a consequence of speciation[20,57]. Regardless of whether identified DMI alleles are the first step in the origin of reproductive isolation, a striking pattern from recent work is that these alleles have evolved under strong positive selection rather than genetic drift and that genomic conflict is often implicated as the source of this selection. For example, one study identified *Overdrive* (*Ovd*), which is an X-linked gene that underlies both hybrid male sterility and sex ratio distortion in crosses between *Drosophila pseudoobscura pseudoobscura* and

Divergence hitchhiking
When divergent selection on a locus reduces the effective migration rate for physically linked regions, which increases the opportunity for divergence at loci under weaker selection in the surrounding regions. Regions of divergence hitchhiking may remain much larger than those of traditional hitchhiking after a selective sweep within populations because of the persistent reduction in the ability of flanking regions to recombine away from a divergently selected gene.

Parapatric
Pertaining to organisms, populations or species that inhabit either adjacent geographical regions or spatially distinct but adjacent habitats and that may exchange genes.

Sympatric
Pertaining to organisms, populations or species that share the same geographical region and that overlap in their use of space with no spatial barriers to gene exchange.

F_{ST}
(Also known as Wright's fixation index). A measure of population subdivision that compares the correlation between two gene copies that are randomly drawn from the same population to that between two gene copies that are drawn from two different populations. An F_{ST} of 1 indicates that two populations are fixed for alternative alleles.

Allopatric
Pertaining to organisms, populations or species that inhabit distinct geographical regions and are therefore not exchanging genes.

Coalescence
The merging of two genetic lineages in a common ancestor.

Incomplete lineage sorting
The situation in which some alleles share a more recent common ancestor with alleles in another species than with other alleles in the same species.

Drosophila pseudoobscura bogotana[51]. Another recent analysis found strong evidence for ongoing positive selection within *Drosophila mauritiana* in genes that have diverged between this species and its closest relatives and that are known to be involved in genomic conflict[12]. Two marked polymorphism troughs on the X chromosome were centred on a pair of genes that cause sex ratio distortion within *Drosophila simulans* and on *Odysseus* (*OdsH*), which is a rapidly evolving homeobox gene that was known to cause male sterility in *D. mauritiana*–*D. simulans* hybrids[32] and that may be involved in genomic conflict. These are two candidate cases of speciation by conflict-driven DMI evolution.

Genomic coupling of reproductive barriers. The build-up of associations between several traits or loci that are involved in reproductive isolation strengthens the total barrier to gene flow between diverging populations and is therefore important for the evolution of strong reproductive isolation[114,115]. Such genomic coupling can involve any prezygotic or postzygotic barriers[116]. Deviations from linkage equilibrium between barrier loci can be initially generated either by new mutations that arise on a particular genetic background or by genetic drift during divergence with limited gene flow. For example, coinciding barriers may arise through secondary contact between divergent populations, through the evolution of DMIs as an incidental by-product of divergent selection[117] or through hitchhiking of intrinsic incompatibility alleles with divergently selected alleles, as has been shown for populations of monkey flowers that have adapted to the presence of heavy metals[118]. However, for genomic coupling to be important in speciation, such coupling has to be maintained or even strengthened in the face of gene flow, which typically requires divergent selection[6].

Selection is expected to favour the coupling of barriers if this leads to an increase in mean fitness. In theory, this can involve multiple intrinsic barriers (such as DMIs)[119,120], both intrinsic and extrinsic postzygotic barriers, as well as sexual and other prezygotic isolation traits. Across an ecotone, multifarious extrinsic selection can assemble and maintain many coinciding clines at loci that are involved in adaptation[121], and these can become coupled with sexual isolation traits[122] and with DMIs[18,115,123]. Selection can also directly favour the evolution of increased prezygotic isolation, as in the case of reinforcement[124]. Finally, sexual conflict can couple intrinsic postzygotic isolation and prezygotic sexual isolation because DMIs that are driven by sexual conflict and genes underlying sexual traits or preferences which are expressed only in one sex may both accumulate on sex chromosomes[53,125]. Consistent with these expectations, loci for plumage colour, mating preferences and intrinsic postzygotic incompatibilities are coupled on the Z chromosome in flycatchers[52] and in Gouldian finches[126,127]. Similarly, loci for behavioural isolation and hybrid male sterility are coupled on the X chromosome in a species pair of Japanese sticklebacks[45].

As recombination tends to break up gene associations, genomic architectures that eliminate or decrease recombination are expected to facilitate coupling and hence speciation[128]. Most prominently, recombination will affect neither one-allele mechanisms nor associations among traits that are pleiotropically influenced by the same allele[129]. One-allele mechanisms do not leave a population-specific signature in the genome at the primary isolation locus, but such mechanisms should be detectable as sweeps that are shared by both diverging populations if they arise during speciation (for example, if an allele for imprinting on the phenotype of the father spreads across two incipient species that were connected by gene flow). Despite the theoretical expectation that one-allele mechanisms evolve more readily during speciation with gene flow than other types of barriers[6,16,129], we are not aware that the predicted genomic signature of shared sweeps at isolation loci has been detected in any case. The revelation of such a signature would be a strong contribution of speciation genomics to supporting a classical prediction of speciation theory.

Loci that underlie two-allele mechanisms are expected to be concentrated in regions of reduced recombination. Recent genomic studies have observed genomic architectures that either eliminate or reduce recombination between traits that are involved in reproductive isolation: there is evidence of synergistic pleiotropy in multiple-effect or 'magic' traits[16,130–132], which are traits that contribute both to adaptation and to reproductive isolation (for example, a trait that contributes to local adaptation and that is also used as a mating cue). Several genes that underlie isolating traits have been found together in inversions[133–135], on sex chromosomes[45,52,127] and also in otherwise tight physical linkage[118,136], including mating traits and mating preferences in cases of speciation with gene flow[137]. These data also provide some evidence that reinforcement of prezygotic isolation is facilitated either by linkage (as in the example of flycatchers[138]) or by pleiotropy (as in the example of phlox[131]). In other cases, reinforcement might be constrained[139] where loci are unlinked and where there is extensive gene flow. However, recent genomic studies have also provided empirical examples of coupling between unlinked loci in fully sympatric hybridizing species[77] and especially in hybrid zones, in which clines at many unlinked loci often coincide, although it is not always clear exactly how these loci are implicated in reproductive isolation[140]. Unbiased whole-genome resequencing data and genome scans from diverging populations, coupled with methods to reduce bias from NGS data[141] and with mapping of isolation traits, are needed to test the generality of these patterns.

Effect sizes and pleiotropy. A key question with a long history[55,142] is whether speciation is typically initiated by divergence at few loci of large and possibly pleiotropic effects or by divergence at many loci with small and additive effects[132,143]. This distinction is important because it will affect how speciation is constrained by the availability of suitable genetic variation and how likely it is that either selection or genetic drift may overcome

Sweeps
Increases in frequencies of alleles and closely linked chromosomal segments due to positive selection. Sweeps initially reduce variation and subsequently lead to a local excess of rare alleles as new unique mutations accumulate.

D_{xy}
The average number of nucleotide substitutions per site between two populations.

Secondary contact
The meeting of the distribution ranges of two distinct populations or species after a period of evolutionary divergence in geographical isolation (that is, allopatry).

Gene-flow–selection balance
A level of differentiation between subpopulations at which the homogenizing effect of gene flow and the differentiating effect of divergent selection are in equilibrium.

Multifarious
Pertaining to divergent selection that acts on multiple traits.

Cline
Directional variation in phenotype or genotype, or change in frequency (for example, of an allele) across a geographical region.

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

Quantitative trait locus
(QTL). A chromosomal region that has a significant effect on a phenotype.

Admixture mapping
The identification of genetic loci that contribute to phenotypic differences between ancestral populations by investigating genotype–phenotype correlations in a population of mixed ancestry.

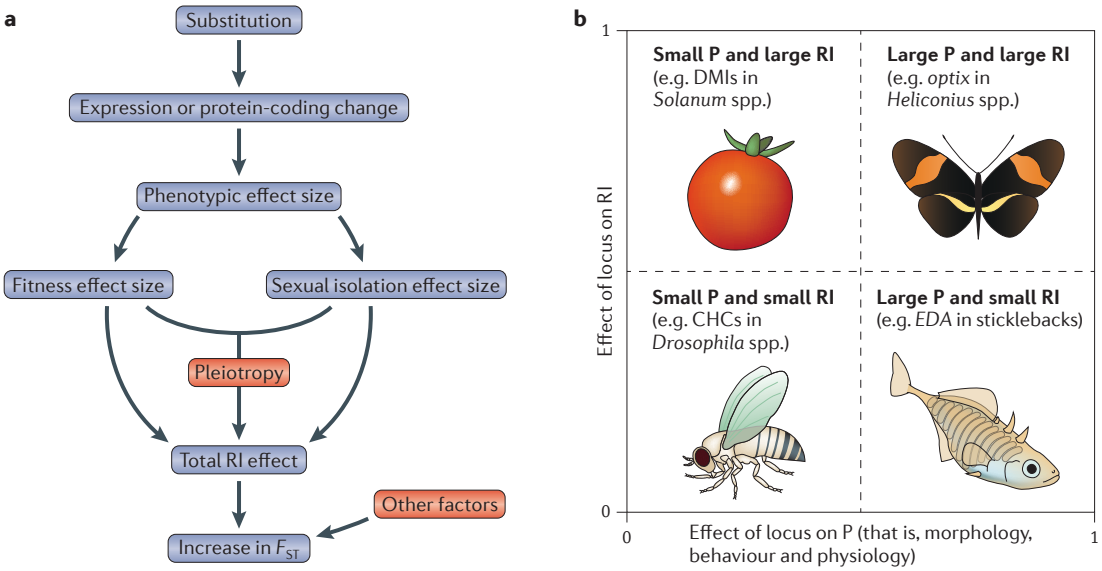


Figure 2 | **Effect sizes of substitutions on phenotype and on reproductive isolation.** **a** | Effects of variation at different levels and the connections between these levels are shown. The size of effect can vary at each step from zero or quite small to very large. A substitution can alter gene expression or protein coding, which in turn has some effect on a phenotype. This phenotype can have effects of varying sizes on environment-dependent fitness and hence possibly extrinsic postzygotic isolation; on environment-independent fitness and hence possibly intrinsic postzygotic isolation; and on prezygotic isolation. Alternatively, a phenotype may pleiotropically affect both fitness and prezygotic isolation. All of these effects combine to generate total reproductive isolation (RI), which will probably increase F_{ST}, although other factors also can affect F_{ST}. **b** | There is a lack of correlation between the effect of a locus on phenotype (P) and that on RI, such that phenotypic effect size does not necessarily predict RI effect size. Examples of different relationships between these effect sizes can be found in Bateson–Dobzhansky–Muller incompatibilities (DMIs) in *Solanum* spp.[27], the *optix* locus in *Heliconius* spp.[159], the cuticular hydrocarbon (CHC) loci in *Drosophila* spp.[200] and the ectodysplasin (*EDA*) locus in sticklebacks[190]. The relationships between phenotypic effect size, RI effect size and F_{ST} are currently unknown to a large extent.

Standing genetic variation
Allelic variation that is currently segregating within a population, as opposed to alleles that arise through new mutation events.

Introgressive hybridization
The introduction of genes from one population or species into another through hybridization.

Fixation
The situation in which a mutation or a variant has achieved a frequency of 100% in a population.

Ecotone
A zone in which there is a transition between two distinct biological communities, for example, between forest and grassland or between aquatic and terrestrial habitats. It is typically associated with changes in the physical environment.

Reinforcement
Selection for the strengthening of prezygotic barriers to avoid the production of unfit hybrids between taxa that have previously evolved some postzygotic isolation.

One-allele mechanisms
Mechanisms that produce reproductive barriers through the spread of the same allele in each of two diverging populations, such as alleles for behavioural imprinting or reduced migration.

Two-allele mechanisms
Mechanisms that produce reproductive barriers through the spread of different alleles at the same locus in two diverging populations, such as alleles for different habitat or mating preferences.

Hybrid zones
Spatially restricted regions where the distribution ranges of distinct populations or incipient species come into contact and where hybrids are formed.

gene flow. On their own, F_{ST} estimates from genome scans inform us little about the effect sizes of individual alleles on phenotypes, fitness or reproductive isolation[106] (FIG. 2). With regard to fitness, Fisher's geometric model predicts that the probability that a mutation is favourable decreases exponentially with mutational effect size, and we therefore expect few alleles of large positive fitness effects but many alleles of small effects[144–147] (but see REF. 147 for an argument that much evolutionary change relies only on small-effect alleles). However, this prediction does not take into account standing genetic variation, gene flow or changing environments. When these factors are considered, the predictions may change[47,146,148] and may even reverse[149].

Speciation with gene flow may require divergent or disruptive selection to be concentrated on a small number of regions in the genome that also have large effects on reproductive isolation[6]. Theoretically expected distributions of effect sizes in terms of reproductive isolation (rather than fitness) may be different for different classes of isolating barriers, but current data are equivocal (FIG. 2b). For example, the mapping of hybrid inferiority in natural environments for *Arabidopsis* spp. has shown that reproductive isolation is due to many genes with moderate effects[150]. By contrast, hybrid inviability in *Mimulus guttatus* is a consequence of two linked loci of major effect[118]. Predictions about the distribution of effect sizes expected for genes that underlie DMIs are also lacking in general, partly because effect sizes depend on mutation order and on the extent of background genomic divergence. Traits that govern prezygotic isolation and especially sexual isolation (BOX 2) are likely to have large effects on reproductive isolation because they directly influence mating or fertilization patterns[1,6,16,151–153]. To test this prediction with genomic data, existing quantitative genetic, mapping and candidate gene studies[45,108,110,127,137,154–158] should now be followed up by NGS-based genome scans to assess reproductive isolation around these loci[106].

Recently identified large-effect alleles that are involved in adaptation and speciation with gene flow are often highly pleiotropic; these alleles include *optix* in *Heliconius* spp.[159] and ectodysplasin (*EDA*) in sticklebacks[160], although we lack estimates of the effect of *EDA* on reproductive isolation or fitness. Such alleles may be rare among newly arising mutations, but alleles with synergistically pleiotropic effects may be more common in standing genetic variation. Recent theory suggests that large-effect or pleiotropic alleles may be favoured by selection during evolution in gene-flow–selection balance and hence eventually become enriched in taxa with divergence and gene flow[47].

Genomic constraint. The flipside of the coupling problem is that genetic correlation between traits that results from either pleiotropy or tight linkage may also constrain speciation. As new population genomic data are revealing divergence in many regions of the genome early in speciation, there is an opportunity to unite population genomics with a quantitative genetics perspective on the evolution of polygenic traits during speciation. In quantitative genetics terms, standing genetic variation is quantified by the G-matrix[161], which may indicate potential constraints on adaptive evolution that affect the response to directional selection[162,163], as well as potential constraints on genetic drift[164]. Tests to detect the effect of selection on the G-matrix are available[165]. Divergence among populations is biased along axes with greater genetic variation and covariation, and it is constrained along axes with little variation or covariation[163,166,167]. Importantly, however, genetic constraints are not only negative. Genetic covariation may align with correlational selection[168,169] and, as discussed above, pleiotropy can couple adaptation to reproductive isolation. It is not known how genes of major effect, compared with the traditional assumption of many genes of small effect, influence the structure of the G-matrix[170], and how higher moments of the distribution of genetic variation and covariation affect the response to divergent selection[171]. These questions can now be addressed with genomic methods, such as the direct estimation of the G-matrix in outbred populations using NGS data[172]. A different approach is to estimate the G-matrices for gene regulatory networks from gene expression data. The analysis of genomic data in a quantitative genetics framework in this way will illuminate how genomic constraint affects speciation[173].

Studying effects of hybridization is one promising application. Beyond being a source of allelic variation, hybridization may facilitate evolution and perhaps speciation by releasing populations from constraints that are caused by genetic correlations. Although there is accumulating empirical evidence that selection alters genomic architecture[168,174], the role of gene flow in aligning the G-matrix with the direction of divergent or disruptive selection has rarely been investigated[149]. The emerging consensus that hybridization frequently introduces adaptive variation[18] calls for empirical studies in this area. We predict that hybridization will influence speciation not only by generating novel and transgressive phenotypes but also by aligning the G-matrix with the axis of divergent selection (FIG. 3a). Even when early generation hybrids have reduced fitness, hybrid populations may benefit from increased evolvability over time[175]. Hybridization may alter patterns of genetic covariance at a rate that is much faster than is possible by selection alone and may lead to bursts of evolutionary diversification and speciation[113,176] (FIG. 3b–d). Genomic methods can now be used to assess these hypotheses in several ways, for example, by the direct estimation of the G-matrix in both parental and hybrid natural populations, and by association or admixture mapping of loci that contribute to novel adaptive phenotypes in hybrid populations[107].

G-matrix
The additive genetic variance–covariance matrix that summarizes the variance within and the covariance between multiple phenotypic traits.

Correlational selection
Selection for optimal combinations of characteristics.

Transgressive phenotypes
Phenotypes in hybrids that exceed the range of phenotypes that are observed in the parental taxa.

Speciation genomics: towards a synthesis

Speciation can proceed in many different ways, but these can be grouped in terms of 'drivers' (that is, genetic drift and different types of selection), causes (that is, extrinsic environment-dependent or intrinsic environment-independent isolation) and stage in the life cycle of reproductive isolation (that is, postzygotic or prezygotic isolation), which results in two major classes that are, at least in theory, quite distinct (BOX 2). Reproductive isolation is initiated by extrinsic selection in one class and by intrinsic incompatibility in the other class. Analyses of NGS data have begun to shed light on the signatures of these processes in the genome. Both classes of speciation can generate reproductively isolated species in allopatry, but parapatric and especially sympatric speciation are constrained to situations in which divergent natural and/or sexual selection has overcome the homogenizing effects of gene flow[1,6]. Whether speciation in such scenarios can proceed depends both on the strength of selection[2,6] and on the genetic architecture of adaptation and reproductive isolation[76,121]. Speciation that is driven by genomic conflict is much less likely to be initiated in the presence of gene flow because selfish genetic elements may then spread across populations and thereby prevent or slow down the accumulation of conflict-driven DMIs[177]. However, it remains possible that fairly brief periods of allopatry are sufficient for the origins of conflict-driven DMIs. Although DMIs may be removed by selection after secondary contact with gene flow, they may, in theory, facilitate speciation if they become coupled with other components of reproductive isolation before they are eliminated[115,178]. How often this happens is unknown.

These principles are not new[1], but they can and should now be examined with much greater resolution using genomic methods. Although speciation genomics is clearly still in its infancy, a few trends are emerging from the first generation of NGS-based genome scans, particularly in relation to non-allopatric speciation. The available evidence indicates that divergence can be genomically widespread early in speciation and may generally be so in species that coexist in full sympatry[74–77,80], whereas it can be restricted to few genomic islands in parapatric ecotypes[69,70]. Multifarious divergent selection or genomically widespread selection is perhaps important to generate sufficient reproductive isolation to permit maintenance and even accumulation of genetic differentiation in sympatry. More data are now needed to confirm this intriguing pattern.

Some genomic regions that are divergent between incipient and sibling species in geographical proximity contain genes with large effects on adaptation and pleiotropic effects on prezygotic isolation. The alleles at several of these loci have turned out to be ancient variants that were either present as standing genetic variation or brought together by hybridization in the ancestors of emerging species pairs[99,110,111]. Although it is premature to draw strong conclusions, this may turn out to be another emergent feature of speciation with gene flow. We expect larger effect sizes, less antagonistic pleiotropy and more synergistic pleiotropy in ancient alleles

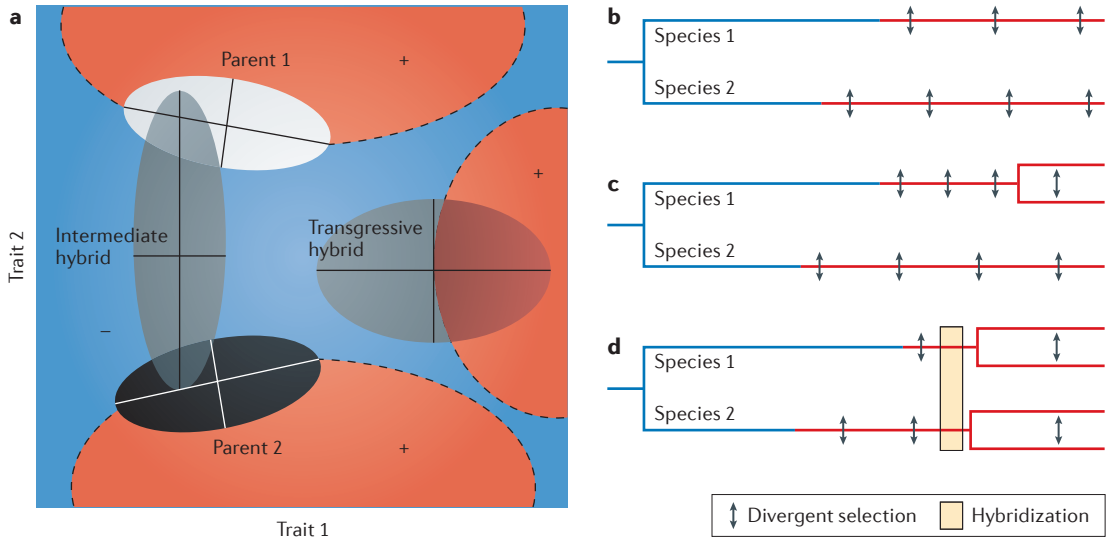


Figure 3 | **Influence of genetic constraints on speciation.** **a** | With the help of next-generation sequencing, it is now feasible to infer relatedness of individuals in any given natural population and thus to estimate a G-matrix without the use of pedigree data[172]. The G-matrix (represented as an ellipse in the space of two quantitative traits) can bias evolution in certain directions. Depending on the adaptive landscape (represented by regions of higher (+; red) and lower (–; blue) fitness than the parental populations (white and black)), the G-matrix might constrain adaptive divergence and speciation. Hybridization events may facilitate speciation either by aligning the G-matrix in the direction of divergence between parental species (that is, the intermediate hybrid) or by giving rise to novel phenotypes (that is, the transgressive hybrid) in new regions of positive fitness that cannot be reached through gradual evolution in either of the parental species. The influence of genetic constraints on speciation can be tested at the phylogenetic level. **b** | Constraints may persist over evolutionary time as a result of the inability of divergent selection to change genetic architecture, which prevents speciation from happening. **c** | Alternatively, other forms of selection (for example, correlational selection) can alter the structure and the orientation of the G-matrix and can potentially facilitate divergence and speciation over moderate timescales. **d** | Hybridization and gene flow can markedly alter the G-matrix in only a few generations, which 'fuels' adaptive divergence and results 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.

that have been refined by selection over time than in alleles that are newly arising through mutation. We hypothesize that the substitution of such ancient alleles at major-effect loci has the potential to quickly reduce gene flow, to the point at which alleles with smaller effect at other loci can also spread in one of the two divergent incipient species without affecting the other one. Genome scans of divergence early in the speciation continuum at low overall reproductive isolation (BOX 2) should allow explicit tests of these hypotheses.

Alternative mechanisms and geographical modes of speciation make different predictions for patterns in genomic data. Specifically, we predict that speciation due to conflict-driven DMIs involves greater divergence at centromeres and at sex chromosomes, and these regions should therefore bear signatures of selective sweeps. Divergence under ecological selection may be more widely distributed across the genome, and sweeps at individual loci may be less marked. The available data are consistent with these expectations, although theory predicts an accumulation of genes for ecological divergence in regions of low recombination when selection is antagonized by gene flow[128]. Divergence by sexual selection may be concentrated on sex chromosomes[52], but support for this prediction is not always found and predictions vary with the sex determination system[54]. Further testing of these predictions requires many more population genomic studies of divergence in a wider range of taxa and across a greater range of points in the speciation continuum. More generally, future work should seek to determine the extent to which different evolutionary mechanisms and geographical modes of speciation can be distinguished on the basis of genomic data and, in turn, the extent to which genomic features can predict the modes and the mechanisms of speciation that apply to a given evolutionary lineage.

Taxonomic variation is prevalent in the propensity for speciation without geographical isolation[179], and it will be interesting to learn whether variation in genomic architecture explains some of this. Apart from the total strength of selection, whether selection can overcome gene flow depends on the number of genomic regions that are targeted by selection, on the rate of recombination between these regions and on the extent of pleiotropy. The analyses of genomic data together with ecological data therefore hold promise to help to explain why non-allopatric speciation occurs readily in

Box 4 | **New data for new theory: speciation genomics and patterns in biodiversity**

As speciation produces the 'raw material' for patterns in biodiversity, an important goal is to connect speciation processes to these patterns[195]. We envisage that speciation genomics can make vital and unique contributions to elucidating these connections. Studies of the distribution of species richness among clades provide evidence for non-uniform diversification rates among taxa, which can arise from differences in speciation and/or extinction rates[196]. Speciation rates that are estimated from the fossil record are much slower than those predicted from mathematical models and those observed in studies of recent diversification. 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[194]. This has been documented in cases of 'speciation reversal' (REFS 193,197,198), which is possible when speciation does not reach completion[121,199].

A better understanding of the genomic basis of speciation might help us to understand the influence of speciation on both species persistence and patterns of species diversity. For example, ecological speciation readily and rapidly produces divergent, partially isolated ecotypes, as well as species that may immediately be able to coexist without competitive exclusion. Ecological speciation might thereby contribute disproportionately to the build-up of biodiversity compared with non-ecological mechanisms[195]. However, isolation between young ecologically differentiated species is often extrinsically based and is contingent on the persistence of divergent selection (BOX 2). The species that arise most rapidly may therefore be those that are the most vulnerable to extinction early in their histories[199]. By contrast, speciation through 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 in which ecological speciation is common, and ecological differences must evolve after speciation 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 bases influence rates of speciation as well as the persistence and coexistence of young species. If mechanisms of speciation leave distinctive genomic signatures, then correlation between genomic patterns and disparity in species richness among clades could be quantitatively tested using comparative phylogenetic approaches.

cichlid fish, whitefish, sticklebacks, *Rhagoletis* spp. flies, *Heliconius* spp. butterflies, *Coprosma* spp. shrubs[180] and some other taxa but is not reported in the majority of others. This combination of approaches may also help more generally to explain why some taxa undergo speciation and accumulate species diversity a lot more readily than others. Answers to such questions will also facilitate an understanding of larger-scale patterns in species diversity (BOX 4).

There is still a lack of population genomic studies that explicitly compare the rates of evolution and the genomic distribution of prezygotic, extrinsic postzygotic and intrinsic postzygotic barriers to gene flow. We believe that such studies hold considerable promise to overcome old dichotomies in speciation genetics. As the discovery of DMIs used to be laborious, we cannot yet answer the question of how often DMIs are caused by genomic conflict, ecological selection or genetic drift. Hopefully, this will also change soon, as genomic data now allow the identification of DMI loci at an increasing pace[12,26] (BOX 1).

A part of a synthesis in speciation genomics that is still missing is the integration of evolutionary developmental biology. Insights from this field make several relevant suggestions for speciation genomics[181]. First, mutations in protein-coding sequences may have pleiotropic effects more often than those in *cis*-regulatory regions. Second, pleiotropy will be more common when selection targets genes that have central roles in gene regulatory networks, and many morphological developmental genes are in such positions. Third, because of the first two predictions, morphological evolution may often be constrained to take place through changes in *cis*-regulatory mutations, whereas physiology may be freer to evolve through protein-coding mutations. These predictions make interesting yet little explored connections between some of the above questions in speciation research and the debate about the prevalence of protein-coding mutations compared with *cis*-regulatory mutations in evolution[181,182]. Notwithstanding possible ascertainment biases, empirical data suggest that divergence between sibling species and between conspecific populations is predominantly due to evolution of protein-coding genes and is independent of their positions in gene regulatory networks, but morphological differences between species that diverged a longer time ago are predominantly a result of *cis*-regulatory evolution[181]. To explain this, selection that acts early during population divergence has been proposed to partly overcome the negative fitness effects of antagonistic pleiotropy that are expected from protein-coding mutations, but such selection may not be strong enough to fix these mutations[181]. Over time, as more mutations become available, *cis*-regulatory mutations that have more specific effects and less antagonistic pleiotropy would replace the protein-coding variants. An interesting implication is that the mutations that are responsible for phenotypic differences between older species may be distinct from those that are causally important in the process of population divergence and speciation, even when the mechanisms of speciation and the diverging phenotypes are the same. This hypothesis needs to be tested by studies of the genomic basis (that is, whether mutations are in protein-coding or *cis*-regulatory regions) of species divergence in incipient species compared with older species in the same taxon. We are not aware that such data exist.

These are exciting times for speciation research, and major progress in the field is likely to come from integrating the analyses of genomic data with studies of ecology, behaviour, developmental biology and theory. We propose three major building blocks as a 'roadmap' for such continued integration.

First, there is a need for more comparative genome scans both at different stages of the speciation continuum in closely related taxa and in replicate species pairs in the same taxon. These data need to be combined with annotation of the effects of alleles on phenotypes and on reproductive isolation, which can be done through QTL mapping or functional analyses in the context of annotated reference genomes. This would allow divergent genomic regions to be associated with mechanisms of reproductive isolation. Such studies need to be repeated in the following scenarios: in taxa in which speciation is driven by ecological selection, sexual selection and intrinsic incompatibilities (BOX 2); in different spatial contexts; and in taxa that have not speciated but that occupy similar environments to those that have undergone speciation. Sampling design should explicitly aim to explore variation both in different stages of the speciation continuum and for different degrees of geographical isolation (FIG. 1), and the history of geographical isolation should ideally be known. With replication and clever experimental and comparative study designs, it will eventually become possible to understand whether different mechanisms and modes of speciation can be distinguished on the basis of patterns observed in genome-wide data.

Second, experimental population genomic studies of speciation are needed to measure the strength and the multifarious nature of selection, and more generally to test hypotheses about processes that underlie differentiation and isolation, including genomic conflict, coupling and heterogeneity in recombination rates.

Third, there is a need for theoretical modelling that includes the influences of variable demography, recombination rates and time, and that explicitly considers standing genetic variation and different sources of incompatibilities. Such models will be helpful in generating predictions that can be tailored to individual empirical study systems to make them testable. These predictions could include genomic signatures of alternative speciation modes and mechanisms, and how such modes and mechanisms can be inferred from genomic patterns at different stages of the speciation continuum. Improved methods for estimating the timing of long-term gene flow would also be valuable[90]. Given the increasingly widespread evidence for recruitment of ancient genetic variation into recent speciation events, we also need analytical methods for rigorous hypothesis testing regarding the source of such variation (that is, the contributions of hybridization and standing genetic variation). Such methods could include comparisons of the phylogenetic histories of genomic regions that confer adaptation and reproductive isolation with those of other genomic regions of young sister species[74,77,99,111].

Conclusions

New approaches for gathering large amounts of genomic data in non-model organisms have begun to produce intriguing and unexpected insights into the genetics of speciation. Sympatrically coexisting species are characterized by heterogeneous differentiation that is widely scattered across the genome even when these species are still young, but adaptive differentiation between parapatric populations can be restricted to a few genomic islands. Ancient alleles with large and pleiotropic effects characterize both types of divergence and were often acquired by interspecific hybridization. Genomic conflict may be a frequent source of intrinsic postzygotic isolation and may be recognized in genome scans either as strong sweep signatures on sex chromosomes or in isolated islands of divergence on autosomes. We now need more strongly integrated studies that cover multiple components of reproductive isolation both at multiple stages of the speciation continuum and in geographical settings that range from complete allopatry to full sympatry, and we need to pay additional attention to the history of population contact (that is, primary or secondary). With the rapid development of approaches to generate and analyse genomic data, it will then soon become possible to construct an integrated picture of speciation that starts from the evolution of reproductive barriers and to understand how this is influenced by ecological and genomic constraints, through the way speciation creates signatures of genomic divergence, to how genomic properties of organisms bear witness to history and interact with ecology in shaping patterns in biodiversity. There is no doubt that a new phase of discovery has begun, which will lead to a greatly increased understanding of the origin of species.

1. Coyne, J. & Orr, H. *Speciation* (Sinauer Associates, 2004).
This comprehensive book is a must-read for every student of speciation. It provides excellent background information and reviews of all facets of research on the speciation process.
2. Nosil, P. *Ecological Speciation* (Oxford Univ. Press, 2012).
This is an in-depth treatment of speciation by divergent natural selection.
3. Price, T. *Speciation in Birds* (Roberts & Company, 2008).
4. Dobzhansky, T. Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics* **21**, 113–135 (1936).
5. Muller, H. J. & Pontecorvo, G. Recessive genes causing interspecific sterility and other disharmonies between *Drosophila melanogaster* and *simulans*. *Genetics* **27**, 157 (1942).
6. Gavrilets, S. *Fitness Landscapes and the Origin of Species* (Princeton Univ. Press, 2004).
7. van Doorn, G. S., Edelaar, P. & Weissing, F. J. On the origin of species by natural and sexual selection. *Science* **326**, 1704–1707 (2009).
8. M'Gonigle, L. K., Mazzucco, R., Otto, S. P. & Dieckmann, U. Sexual selection enables long-term coexistence despite ecological equivalence. *Nature* **484**, 506–509 (2012).
9. Gavrilets, S. & Losos, J. B. Adaptive radiation: contrasting theory with data. *Science* **323**, 732–737 (2009).
10. Dieckmann, U., Doebeli, M., Metz, J. A. J. & Tautz, D. (eds) *Adaptive Speciation* (Cambridge Univ. Press, 2004).
11. Coyne, J. A. & Orr, H. A. "Patterns of speciation in *Drosophila*" revisited. *Evolution* **51**, 295–303 (1997).
This is a key comparative study on the rate of evolution of reproductive isolation and was a model for similar studies in other taxa.
12. 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 Res.* **23**, 99–110 (2013).
In this paper, genome-sequence data show how widespread strong selection due to genomic conflict can be, which suggests that such selection may be a potent source of incompatibilities between previously isolated populations.
13. van der Sluijs, I. *et al.* Female mating preference functions predict sexual selection against hybrids between sibling species of cichlid fish. *Phil. Trans. R. Soc. B* **363**, 2871–2877 (2008).
14. Panhuis, T. M., Butlin, R., Zuk, M. & Tregenza, T. Sexual selection and speciation. *Trends Ecol. Evol.* **16**, 364–371 (2001).
15. Boughman, J. W. How sensory drive can promote speciation. *Trends Ecol. Evol.* **17**, 571–577 (2002).

16. Maan, M. E. & Seehausen, O. Ecology, sexual selection and speciation. *Ecol. Lett.* **14**, 591–602 (2011).
17. Hendry, A. P., Nosil, P. & Rieseberg, L. H. The speed of ecological speciation. *Funct. Ecol.* **21**, 455–464 (2007).
18. Abbott, R. *et al.* Hybridization and speciation. *J. Evol. Biol.* **26**, 229–246 (2013).
19. Presgraves, D. C. Speciation genetics: search for the missing snowball. *Curr. Biol.* **20**, R1073–R1074 (2010).
20. Presgraves, D. C. The molecular evolutionary basis of species formation. *Nature Rev. Genet.* **11**, 175–180 (2010).
This review summarizes empirical evidence on the genetic basis of hybrid dysfunction and focuses on work in genetic model organisms in which key genes have been identified.
21. Rieseberg, L. H. & Blackman, B. K. Speciation genes in plants. *Ann. Bot.* **106**, 439–455 (2010).
22. Rieseberg, L. H. & Burke, J. M. A genic view of species integration — commentary. *J. Evol. Biol.* **14**, 883–886 (2001).
23. Bikard, D. *et al.* Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. *Science* **323**, 623–626 (2009).
24. 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).
25. Ting, C. T. *et al.* Gene duplication and speciation in *Drosophila*: evidence from the *Odysseus* locus. *Proc. Natl Acad. Sci. USA* **101**, 12232–12235 (2004).
26. 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 (2010).
27. Moyle, L. C. & Nakazato, T. Hybrid incompatibility "snowballs" between *Solanum* species. *Science* **329**, 1521–1523 (2010).
28. Stadler, T., Florez-Rueda, A. M. & Paris, M. Testing for "snowballing" hybrid incompatibilities in *Solanum*: impact of ancestral polymorphism and divergence estimates. *Mol. Biol. Evol.* **29**, 31–34 (2012).
29. Lynch, M. & Force, A. G. The origin of interspecific genomic incompatibility via gene duplication. *Am. Naturalist* **156**, 590–605 (2000).
30. Schluter, D. Evidence for ecological speciation and its alternative. *Science* **323**, 737–741 (2009).
31. Nei, M., Maruyama, T. & Wu, C. I. Models of evolution of reproductive isolation. *Genetics* **103**, 557–579 (1983).
32. Ting, C. T., Tsaur, S. C., Wu, M. L. & Wu, C. I. A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* **282**, 1501–1504 (1998).
33. Crespi, B. & Nosil, P. Conflictual speciation: species formation via genomic conflict. *Trends Ecol. Evol.* **28**, 48–57 (2013).
34. Frank, S. A. Divergence of meiotic drive-suppression systems as an explanation for sex-biased hybrid sterility and inviability. *Evolution* **45**, 262–267 (1991).
35. 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).
36. Cocquet, J. *et al.* A genetic basis for a postmeiotic X versus Y chromosome intragenomic conflict in the mouse. *PLoS Genet.* **8**, e1002900 (2012).
37. Malik, H. S. in *Progress in Molecular and Subcellular Biology* (ed. Ugarovic, D.) 33–52 (Springer, 2009).
38. 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).
39. Rebollo, R., Horard, B., Hubert, B. & Vieira, C. Jumping genes and epigenetics: towards new species. *Gene* **454**, 1–7 (2010).
40. Watson, E. T. & Demuth, J. P. in *Speciation: Natural Processes, Genetics and Biodiversity* (ed. Michalak, P.) (Nova Science Publishers, 2013).
41. Burton, R. S. & Barreto, F. S. A disproportionate role for mtDNA in Dobzhansky–Muller incompatibilities? *Mol. Ecol.* **21**, 4942–4957 (2012).
42. Sambatti, J. B. M., Ortiz-Barrientos, D., Baack, E. J. & Rieseberg, L. H. Ecological selection maintains cytonuclear incompatibilities in hybridizing sunflowers. *Ecol. Lett.* **11**, 1082–1091 (2008).
43. van Doorn, G. S. & Kirkpatrick, M. Turnover of sex chromosomes induced by sexual conflict. *Nature* **449**, 909–912 (2007).
44. Charlesworth, D. & Charlesworth, B. Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genet. Res.* **35**, 205–214 (1980).
45. Kitano, J. *et al.* A role for a neo-sex chromosome in stickleback speciation. *Nature* **461**, 1079–1083 (2009).
46. Parnell, N. F. & Streelman, J. T. Genetic interactions controlling sex and color establish the potential for sexual conflict in Lake Malawi cichlid fishes. *Heredity* **110**, 239–246 (2013).
47. Yeaman, S. & Whitlock, M. C. The genetic architecture of adaptation under migration–selection balance. *Evolution* **65**, 1897–1911 (2011).
48. Silver, L. The peculiar journey of a selfish chromosome: mouse t haplotypes and meiotic drive. *Trends Genet.* **9**, 250–254 (1993).
49. Larracuente, A. M. & Presgraves, D. C. The selfish segregation distorter gene complex of *Drosophila melanogaster*. *Genetics* **192**, 33–53 (2012).
50. Bull, J. J. *Evolution of Sex-determining Mechanisms*. (Benjamin/Cumings, 1983).
51. Phadnis, N. & Orr, H. A. A single gene causes both male sterility and segregation distortion in *Drosophila* hybrids. *Science* **323**, 376–379 (2009).
52. Saether, S. A. *et al.* Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science* **318**, 95–97 (2007).
53. Lindholm, A. & Breden, F. Sex chromosomes and sexual selection in poeciliid fishes. *Am. Naturalist* **160**, S214–S224 (2002).
54. Reeve, H. K. & Pfennig, D. W. Genetic biases for showy males: are some genetic systems especially conducive to sexual selection? *Proc. Natl Acad. Sci. USA* **100**, 1089–1094 (2003).
55. Templeton, A. R. Mechanisms of speciation — a population genetic approach. *Annu. Rev. Ecol. Systemat.* **12**, 23–48 (1981).
56. Ayala, F. J. & Tracey, M. L. Genetic differentiation within and between species of *Drosophila willistoni* group. *Proc. Natl Acad. Sci. USA* **71**, 999–1003 (1974).
57. Coyne, J. A. Genetics and speciation. *Nature* **355**, 511–515 (1992).
58. Price, T. D. & Bouvier, M. M. The evolution of F_1 postzygotic incompatibilities in birds. *Evolution* **56**, 2083–2089 (2002).
59. Stelkens, R. B., Young, K. A. & Seehausen, O. The accumulation of reproductive incompatibilities in African cichlid fish. *Evolution* **64**, 617–632 (2010).
60. Scopece, G., Musacchio, A., Widmer, A. & Cozzolino, S. Patterns of reproductive isolation in Mediterranean deceptive orchids. *Evolution* **61**, 2623–2642 (2007).
61. Orr, H. A. & Turelli, M. The evolution of postzygotic isolation: accumulating Dobzhansky–Muller incompatibilities. *Evolution* **55**, 1085–1094 (2001).
This paper provides background on the evolution of reproductive isolation through postzygotic intrinsic barriers. It considers the dynamics of accumulation of Dobzhansky–Muller incompatibilities in diverging lineages.
62. Strasburg, J. L. & Rieseberg, L. H. Interpreting the estimated timing of migration events between hybridizing species. *Mol. Ecol.* **20**, 2353–2366 (2011).
63. Nosil, P., Funk, D. J. & Ortiz-Barrientos, D. Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* **18**, 375–402 (2009).
64. Strasburg, J. L. *et al.* What can patterns of differentiation across plant genomes tell us about adaptation and speciation? *Phil. Trans. R. Soc. B* **367**, 364–373 (2012).
65. Barton, N. Gene flow and speciation. *Heredity* **50**, 213–213 (1983).
66. Wu, C. I. The genic view of the process of speciation. *J. Evol. Biol.* **14**, 851–865 (2001).
This review discusses the process of speciation from a genetic perspective by highlighting the difference between thinking of the evolution of reproductive isolation as a whole-genome process and understanding the influence of specific loci on reproductive isolation and gene exchange. The idea that genes, but not whole genomes, are the unit of species differentiation is a seminal perspective, which is crucial to much of the current work in speciation genetics.
67. Via, S. & West, J. The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Mol. Ecol.* **17**, 4334–4345 (2008).
68. Feder, J. L., Egan, S. P. & Nosil, P. The genomics of speciation-with-gene-flow. *Trends Genet.* **28**, 342–350 (2012).
This review brings together empirical studies with theory on the effect of divergent selection on gene flow elsewhere in the genome to examine how reproductive isolation might spread through the genome as speciation proceeds.
69. Nadeau, N. J. *et al.* Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. *Phil. Trans. R. Soc. B* **367**, 343–353 (2012).
70. Andrew, R. L. & Rieseberg, L. H. Divergence is focused on few genomic regions early in speciation: incipient speciation of sunflower ecotypes. *Evolution* **67**, 2468–2482 (2013).
Applying genomic analysis at different points of the speciation continuum is important for understanding how reproductive isolation develops. This study shows how differentiation can be focused on a small proportion of the genome early in speciation.
71. Stolting, K. N. *et al.* Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. *Mol. Ecol.* **22**, 842–855 (2013).
72. Hohenlohe, P. A. *et al.* Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* **6**, e100086 (2010).
73. Deagle, B. E. *et al.* Population genomics of parallel phenotypic evolution in stickleback across stream-lake ecological transitions. *Proc. R. Soc. B* **279**, 1277–1286 (2012).
74. Jones, F. C. *et al.* The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55–61 (2012).
This is an example of how genomic data can inform about the genetic basis for repeated adaptation to similar environments. This study shows that stickleback fish populations throughout the northern hemisphere have repeatedly used the same genetic elements to adapt to freshwater environments.
75. Lawniczak, M. K. N. *et al.* Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. *Science* **330**, 512–514 (2010).
76. Renaut, S. *et al.* Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nature Commun.* **4**, 1827 (2013).
77. 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.* **22**, 2848–2863 (2012).
78. Jones, F. C. *et al.* A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Curr. Biol.* **22**, 83–90 (2012).
79. Parchman, T. L. *et al.* The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. *Mol. Ecol.* **22**, 3304–3317 (2013).
80. 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* **67**, 2483–2497 (2013).
81. Michel, A. P. *et al.* Widespread genomic divergence during sympatric speciation. *Proc. Natl Acad. Sci. USA* **107**, 9724–9729 (2010).
82. Via, S., Conte, G., Mason-Foley, C. & Mills, K. Localizing F_{ST} outliers on a QTL map reveals evidence for large genomic regions of reduced gene exchange during speciation-with-gene-flow. *Mol. Ecol.* **21**, 5546–5560 (2012).
83. 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 (2012).
84. Garrigan, D. *et al.* Genome sequencing reveals complex speciation in the *Drosophila simulans* clade. *Genome Res.* **22**, 1499–1511 (2012).
85. Neafsey, D. E. *et al.* SNP genotyping defines complex gene-flow boundaries among African malaria vector mosquitoes. *Science* **330**, 514–517 (2010).
86. Nosil, P. *et al.* Genomic consequences of multiple speciation processes in a stick insect. *Proc. R. Soc. B* **279**, 5058–5065 (2012).
87. Martin, S. H. *et al.* Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* **23**, 1817–1828 (2013).

88. 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).
89. Hahn, M. W., White, B. J., Muir, C. D. & Besansky, N. J. No evidence for biased co-transmission of speciation islands in *Anopheles gambiae*. *Phil. Trans. R. Soc. B* **367**, 374–384 (2012).
90. Sousa, V. & Hey, J. Understanding the origin of species with genome-scale data: modelling gene flow. *Nature Rev. Genet.* **14**, 404–414 (2013).
91. Scally, A. *et al.* Insights into hominid evolution from the gorilla genome sequence. *Nature* **483**, 169–175 (2012).
92. Charlesworth, B., Morgan, M. T. & Charlesworth, D. The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303 (1993).
93. Spencer, C. C. A. *et al.* The influence of recombination on human genetic diversity. *PLoS Genet.* **2**, 1375–1385 (2006).
94. Nachman, M. W. & Payseur, B. A. Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Phil. Trans. R. Soc. B.* **367**, 409–421 (2012).
95. Nei, M. *Molecular Evolutionary Genetics* (Columbia Univ. Press, 1987).
96. Charlesworth, B. Measures of divergence between populations and the effect of forces that reduce variability. *Mol. Biol. Evol.* **15**, 538–543 (1998).
97. Renaut, S., Owens, G. L. & Rieseberg, L. H. Shared selective pressure and local genomic landscape lead to repeatable patterns of genomic divergence in sunflowers. *Mol. Ecol.* **23**, 311–324 (2014).
98. Ellegren, H. *et al.* The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* **491**, 756–760 (2012).
This paper presents a truly genome-wide view of differentiation in a system with remarkable ecological and behavioural information available. Strikingly heterogeneous levels of divergence were observed, including major divergence peaks at telomeres that are so far unexplained.
99. Heliconius Genome Consortium. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94–98 (2012).
100. 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 (2012).
101. Elmer, K. R. & Meyer, A. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol. Evol.* **26**, 298–306 (2011).
102. 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).
103. 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. *Mol. Biol. Evol.* **21**, 945–956 (2004).
104. Excoffier, L., Hofer, T. & Foll, M. Detecting loci under selection in a hierarchically structured population. *Heredity* **103**, 285–298 (2009).
105. Gagnaire, P.-A., Normandeau, E., Pavey, S. A. & Bernatchez, L. Mapping phenotypic, expression and transmission ratio distortion QTL using RAD markers in the Lake Whitefish (*Coregonus clupeaformis*). *Mol. Ecol.* **22**, 3036–3048 (2013).
106. 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**, 164 (2012).
107. Lindtke, D., González-Martínez, S. C., Macaya-Sanz, D. & Lexer, C. Admixture mapping of quantitative traits in Populus hybrid zones: power and limitations. *Heredity* **111**, 474–485 (2013).
This is one of the first studies to use admixture mapping in a natural hybrid population to examine the genetic basis of traits that contribute to reproductive isolation. Application of this powerful approach is likely to make important contributions to speciation research.
108. 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 (2012).
109. Schluter, D. & Conte, G. L. Genetics and ecological speciation. *Proc. Natl Acad. Sci. USA* **106**, 9955–9962 (2009).
110. Seehausen, O. *et al.* Speciation through sensory drive in cichlid fish. *Nature* **455**, 620–626 (2008).
This study shows how divergent female preferences that initially have an ecological basis can impose selection on male signal traits and generate reproductive isolation in the face of gene flow given the right balance between intensity of selection and distribution of habitats.
111. Feder, J. L. *et al.* Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. *Proc. Natl Acad. Sci. USA* **100**, 10314–10319 (2003).
112. Nadeau, N. J. *et al.* Genome-wide patterns of divergence and gene flow across a butterfly radiation. *Mol. Ecol.* **22**, 814–826 (2013).
113. Seehausen, O. Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**, 198–207 (2004).
114. Smadja, C. M. & Butlin, R. K. A framework for comparing processes of speciation in the presence of gene flow. *Mol. Ecol.* **20**, 5123–5140 (2011).
115. 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 (2011).
116. Hermann, K. *et al.* Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in *Petunia*. *Curr. Biol.* **23**, 873–877 (2013).
117. Bank, C., Burger, R. & Hermisson, J. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent-island model. *Genetics* **191**, 845–863 (2012).
118. 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 Biol.* **11**, e1001497 (2013).
119. Barton, N. H. Multilocus clines. *Evolution* **37**, 454–471 (1983).
120. Barton, N. H. & de Cara, M. A. The evolution of strong reproductive isolation. *Evolution* **63**, 1171–1190 (2009).
This theoretical paper considers the conditions under which selection can overcome recombination to bring together multiple reproductive barriers and therefore generate strong reproductive isolation.
121. Nosil, P., Harmon, L. J. & Seehausen, O. Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol.* **24**, 145–156 (2009).
122. Doebeli, M. & Dieckmann, U. Speciation along environmental gradients. *Nature* **421**, 259–264 (2003).
123. Butlin, R. K. & Ritchie, M. G. Pulling together or pulling apart: hybridization in theory and practice. *J. Evol. Biol.* **26**, 294–298 (2013).
124. Servedio, M. R. & Noor, M. A. F. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Evol. Systemat.* **34**, 339–364 (2003).
125. Qvarnström, A. & Bailey, R. I. Speciation through evolution of sex-linked genes. *Heredity* **102**, 4–15 (2009).
126. Pryke, S. R. & Griffith, S. C. Postzygotic genetic incompatibility between sympatric color morphs. *Evolution* **63**, 793–798 (2009).
127. Pryke, S. R. Sex chromosome linkage of mate preference and color signal maintains assortative mating between interbreeding finch morphs. *Evolution* **64**, 1301–1310 (2010).
128. Kirkpatrick, M. & Barton, N. Chromosome inversions, local adaptation and speciation. *Genetics* **173**, 419–434 (2006).
This paper gives a thorough theoretical background and new insights into the role of chromosome inversions in adaptation and speciation.
129. Felsenstein, J. Skepticism towards Santa Rosalia, or why are there so few kinds of animals. *Evolution* **35**, 124–138 (1981).
This key paper introduces the antagonism between recombination and the build-up of linkage disequilibrium that lies at the heart of the speciation process.
130. Hopkins, R. & Rausher, M. D. Pollinator-mediated selection on flower color allele drives reinforcement. *Science* **335**, 1090–1092 (2012).
131. Hopkins, R. & Rausher, M. D. Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*. *Nature* **469**, 411–414 (2011).
132. 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 (2011).
133. Lowry, D. B. & Willis, J. H. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* **8**, e1000500 (2010).
134. Joron, M. *et al.* Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **477**, 203–206 (2011).
135. Noor, M. A. F., Grams, K. L., Bertucci, L. A. & Reiland, J. Chromosomal inversions and the reproductive isolation of species *Proc. Natl Acad. Sci. USA* **98**, 12084–12088 (2001).
136. Shaw, K. L. & Lesnick, S. C. Genomic linkage of male song and female acoustic preference QTL underlying a rapid species radiation. *Proc. Natl Acad. Sci. USA* **106**, 9737–9742 (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* **278**, 511–518 (2011).
138. Saetre, G. P. & Saether, S. A. Ecology and genetics of speciation in *Ficedula* flycatchers. *Mol. Ecol.* **19**, 1091–1106 (2010).
139. Bimova, B. V. *et al.* Reinforcement selection acting on the European house mouse hybrid zone. *Mol. Ecol.* **20**, 2403–2424 (2011).
Using powerful analyses of the exceptionally well-studied mouse hybrid zone, this paper provides clear evidence for the operation of reinforcement and also for the limits on its effectiveness in reducing gene flow.
140. Teeter, K. C. *et al.* The variable genomic architecture of isolation between hybridizing species of house mice. *Evolution* **64**, 472–485 (2010).
141. Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y. & Wang, J. SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data. *PLoS ONE* **7**, e37558 (2012).
142. Orr, H. A. The genetics of species differences. *Trends Ecol. Evol.* **16**, 343–350 (2001).
143. 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 (2012).
144. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon Press, 1930).
145. Orr, H. A. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution* **52**, 935–949 (1998).
146. Orr, H. A. The genetic theory of adaptation: a brief history. *Nature Rev. Genet.* **6**, 119–127 (2005).
147. Rockman, M. V. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66**, 1–17 (2012).
148. Gordo, I. & Campos, P. R. Evolution of clonal populations approaching a fitness peak. *Biol. Lett.* **9**, 20120239 (2013).
149. Guillaume, F. & Whitlock, M. C. Effects of migration on the genetic covariance matrix. *Evolution* **61**, 2398–2409 (2007).
150. Bomblies, K. & Weigel, D. *Arabidopsis* and relatives as models for the study of genetic and genomic incompatibilities. *Phil. Trans. R. Soc. B* **365**, 1815–1823 (2010).
151. Leary, G. P. *et al.* Single mutation to a sx pheromone receptor provides adaptive specificity between closely related moth species. *Proc. Natl Acad. Sci. USA* **109**, 14081–14086 (2012).
152. 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).
153. Feder, J. L. *et al.* Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. *Proc. Natl Acad. Sci. USA* **91**, 7990–7994 (1994).
154. Klahre, U. *et al.* Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. *Curr. Biol.* **21**, 730–739 (2011).
155. Dambroski, H. R. *et al.* The genetic basis for fruit odor discrimination in *Rhagoletis* flies and its significance for sympatric host shifts. *Evolution* **59**, 1953–1964 (2005).
156. Haesler, M. P. & Seehausen, O. Inheritance of female mating preference in a sympatric sibling species pair of Lake Victoria cichlids: implications for speciation. *Proc. R. Soc. B* **272**, 237–245 (2005).
157. Fan, P. *et al.* Genetic and neural mechanisms that inhibit *Drosophila* from mating with other species. *Cell* **154**, 89–102 (2013).

158. Ballerini, E. S. *et al.* QTL mapping reveals the genetic architecture of loci affecting pre- and post-zygotic isolating barriers in Louisiana Iris. *BMC Plant Biol.* **12**, 91 (2012).
159. Reed, R. D. *et al. optix* drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* **333**, 1137–1141 (2011).
160. 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).
161. 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).
162. Lande, R. Quantitative genetic-analysis of multivariate evolution, applied to brain — body size allometry. *Evolution* **33**, 402–416 (1979).
163. Schluter, D. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766–1774 (1996).
This key conceptual paper shows how the structure of genetic variances and covariances among quantitative traits can influence the direction of evolution and thus the progress of adaptive radiation.
164. Roff, D. The evolution of the G matrix: selection or drift? *Heredity* **84**, 135–142 (2000).
165. Martin, G., Chapuis, E. & Goudet, J. Multivariate Q_{st}–F_{st} comparisons: a neutrality test for the evolution of the G matrix in structured populations. *Genetics* **180**, 2135–2149 (2008).
166. Hansen, T. F. & Houle, D. Measuring and comparing evolvability and constraint in multivariate characters. *J. Evol. Biol.* **21**, 1201–1219 (2008).
167. Chenoweth, S. F., Rundle, H. D. & Blows, M. W. The contribution of selection and genetic constraints to phenotypic divergence. *Am. Naturalist* **175**, 186–196 (2010).
168. Roff, D. A. & Fairbairn, D. J. A test of the hypothesis that correlational selection generates genetic correlations. *Evolution* **66**, 2953–2960 (2012).
169. 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).
170. 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).
171. 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. *J. Evol. Biol.* **25**, 2210–2231 (2012).
172. Yang, J. *et al.* Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genet.* **43**, 519–525 (2011).
173. Kopp, A. Metamodels and phylogenetic replication: a systematic approach to the evolution of developmental pathways. *Evolution* **63**, 2771–2789 (2009).
174. Eroukhmanoff, F. & Svensson, E. I. Evolution and stability of the G-matrix during the colonization of a novel environment. *J. Evol. Biol.* **24**, 1363–1373 (2011).
175. Nolte, A. W. & Tautz, D. Understanding the onset of hybrid speciation. *Trends Genet.* **26**, 54–58 (2010).
176. Selz, O. M., Lucek, K., Young, K. A. & Seehausen, O. Relaxed trait covariance in interspecific cichlid hybrids predicts morphological diversity in adaptive radiations. *J. Evol. Biol.* **27**, 11–24 (2014).
177. Macaya-Sanz, D. *et al.* Genetic analysis of post-mating reproductive barriers in hybridizing European *Populus* species. *Heredity* **107**, 478–486 (2011).
178. Seehausen, O. Conditions when hybridization might predispose populations for adaptive radiation. *J. Evol. Biol.* **26**, 279–281 (2013).
179. Bolnick, D. I. & Fitzpatrick, B. M. Sympatric speciation: models and empirical evidence. *Annu. Rev. Ecol. Evol. Systemat.* **38**, 459–487 (2007).
180. Papadopulos, A. S. T. *et al.* A comparative analysis of the mechanisms underlying speciation on Lord Howe Island. *J. Evol. Biol.* **26**, 733–745 (2013).
181. Stern, D. L. & Orgogozo, V. The loci of evolution: how predictable is genetic evolution? *Evolution* **62**, 2155–2177 (2008).
182. Hoekstra, H. E. & Coyne, J. A. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* **61**, 995–1016 (2007).
183. Eriksson, A. & Manica, A. Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominins. *Proc. Natl Acad. Sci. USA* **109**, 13956–13960 (2012).
184. 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 (2010).
185. Lawson, D. J., Hellenthal, G., Myers, S. & Falush, D. Inference of population structure using dense haplotype data. *PLoS Genet.* **8**, e1002453 (2012).
186. Slate, J. Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Mol. Ecol.* **14**, 363–379 (2005).
187. Buerkle, C. A. & Lexer, C. Admixture as the basis for genetic mapping. *Trends Ecol. Evol.* **23**, 686–694 (2008).
188. Gompert, Z. & Buerkle, C. A. Bayesian estimation of genomic clines. *Mol. Ecol.* **20**, 2111–2127 (2011).
189. Trier, C. N., Hermansen, J. S., Sætre, G. P. & Bailey, R. I. Evidence for mito-nuclear and sex-linked incompatibilities between the hybrid Italian sparrow and its parent species. *PLoS Genet.* **10**, e1004075 (2014).
190. Barrett, R. D. H., Rogers, S. M. & Schluter, D. Natural selection on a major armor gene in threespine stickleback. *Science* **322**, 255–257 (2008).
This study showed unexpected complexity in the response to selection for alleles of the *EDA* locus, probably because of pleiotropic effects on other fitness-related traits. Such pleiotropic effects may be widespread and have major impacts on the progress of adaptation and speciation.
191. Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. Mapping complex disease traits with global gene expression. *Nature Rev. Genet.* **10**, 184–194 (2009).
192. Orr, H. A. The population-genetics of speciation — the evolution of hybrid incompatibilities. *Genetics* **139**, 1805–1813 (1995).
193. Vonlanthen, P. *et al.* Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature* **482**, 357–362 (2012).
194. Rosenblum, E. B. *et al.* Goldilocks meets Santa Rosalia: an ephemeral speciation model explains patterns of diversification across time scales. *Evol. Biol.* **39**, 255–261 (2012).
195. Butlin, R. K., Bridle, J. R. & Schluter, D. *Speciation and Patterns of Diversity* (Cambridge Univ. Press, 2009).
196. Sanderson, M. J. & Donoghue, M. J. Reconstructing shifts in diversification rates on phylogenetic trees. *Trends Ecol. Evol.* **11**, 15–20 (1996).
197. Seehausen, O., van Alphen, J. J. M. & Witte, F. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277**, 1808–1811 (1997).
198. 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).
199. Seehausen, O., Takimoto, G., Roy, D. & Jokela, J. Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Mol. Ecol.* **17**, 30–44 (2008).
200. 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 (2010).

Acknowledgements
The authors thank the European Science Foundation networking programme Frontiers in Speciation Research (FroSpects) for funding a workshop on Genetics and Genomics of Speciation and for contributing towards travel expenses for graduate students, and N. Pepe and L. Greuter for help with the organization of the workshop. They thank C. Lexer and two anonymous reviewers for suggestions that improved this Review.

Competing interests statement
The authors declare no competing interests.