



Genomics and the origin of species

**Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E.,
Boughman, J.W., Hohenlohe, P.A., Peichel, C.A.,
Saetre, G.-P., Bank, C. and Brannstrom, A.**

**IIASA Interim Report
2013**



Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., Peichel, C.A., Saetre, G.-P., Bank, C. and Brannstrom, A. (2013) Genomics and the origin of species. IIASA Interim Report. IIASA, Laxenburg, Austria, IR-13-066 Copyright © 2013 by the author(s). <http://pure.iiasa.ac.at/10701/>

Interim Reports on work of the International Institute for Applied Systems Analysis receive only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work. All rights reserved. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage. All copies must bear this notice and the full citation on the first page. For other purposes, to republish, to post on servers or to redistribute to lists, permission must be sought by contacting repository@iiasa.ac.at



Interim Report

IR-13-066

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, Glenn-Peter Saetre
Claudia Bank, Åke Brännström (brnstrom@iiasa.ac.at)
Alan Brelsford, Chris S. Clarkson
Fabrice Eroukhmanoff, Jeffrey L. Feder
Martin C. Fischer, Andrew D. Foote
Paolo Franchini, Chris D. Jiggins
Felicity C. Jones, Anna K. Lindholm
Kay Lucek, Martine E. Maan
David A. Marques, Simon H. Martin
Blake Matthews, Joana I. Meier
Markus Möst, Michael W. Nachman
Etsuko Nonaka, Diana J. Rennison
Julia Schwarzer, Eric T. Watson
Anja M. Westram, Alex Widmer

Approved by

Ulf Dieckmann
Director, Evolution and Ecology Program

June 2015

1 Genomics and the origin of species

2
3 Ole Seehausen^{1,2#*}, Roger K. Butlin^{3,4#}, Irene Keller^{1,2,5#}, Catherine E. Wagner^{1,2#}, Janette W.
4 Boughman^{1,6}, Paul A. Hohenlohe⁷, Catherine L. Peichel⁸, Glenn-Peter Saetre⁹

5
6 Claudia Bank¹⁰, Åke Brännström¹¹, Alan Brelsford¹², Chris S. Clarkson¹³, Fabrice Eroukhmanoff⁹,
7 Jeffrey L. Feder¹⁴, Martin C. Fischer⁵, Andrew D. Foote^{15, 28}, Paolo Franchini¹⁶, Chris D. Jiggins¹⁷,
8 Felicity C. Jones¹⁸, Anna K. Lindholm¹⁹, Kay Lucek^{1,2}, Martine E. Maan²⁰, David A. Marques^{1,2,27}, Simon
9 H. Martin¹⁷, Blake Matthews²¹, Joana I. Meier^{1,2,27}, Markus Möst^{17,21}, Michael W. Nachman²², Etsuko
10 Nonaka²³, Diana J. Rennison²⁴, Julia Schwarzer^{1,2,25}, Eric T. Watson²⁶, Anja M. Westram³, Alex Widmer⁵
11 (authors in this section ordered alphabetically)

12
13 *correspondence ole.seehausen@eawag.ch
14

15 ¹ Department of Fish Ecology and Evolution, EAWAG Swiss Federal Institute of Aquatic Science and Technology, Center for Ecology, Evolution and
16 Biogeochemistry, Seestrasse 79, 6047 Kastanienbaum, Switzerland

17 ² Division of Aquatic Ecology and Macroevolution, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland

18 ³ Department of Animal and Plant Sciences, The University of Sheffield, Sheffield, UK

19 ⁴ Sven Lovén Centre – Tjärnö, University of Gothenburg, S-452 96 Strömstad, Sweden

20 ⁵ Institute of Integrative Biology, ETH Zürich, Universitätsstrasse 16, ETH Zentrum CHN, 8092 Zürich, Switzerland

21 ⁶ Department of Zoology; Ecology, Evolutionary Biology & Behavior Program; BEACON, Michigan State University, 203 Natural Sciences, East Lansing, Michigan
22 48824, USA

23 ⁷ Department of Biological Sciences, Institute of Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844-3051, USA

24 ⁸ Division of Human Biology, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, Washington 98109, USA

25 ⁹ Department of Biosciences, Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, P. O. Box 1066, Blindern, N-0316 Oslo, Norway

26 ¹⁰ School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

27 ¹¹ Integrated Science Lab & Department of Mathematics and Mathematical Statistics, Umeå University, 90187 Umeå, Sweden

28 ¹² Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland

29 ¹³ Liverpool School of Tropical Medicine, Liverpool, UK

30 ¹⁴ Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, 46556-0369 USA

31 ¹⁵ Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, ØsterVoldgade 5-7, DK-1350 Copenhagen, Denmark

32 ¹⁶ Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz, Universitätstrasse 10, 78457 Konstanz, Germany

33 ¹⁷ Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

34 ¹⁸ Friedrich Miescher Laboratory of the Max Planck Society, Tübingen, Germany

35 ¹⁹ Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

36 ²⁰ Behavioural Biology Group, Centre for Behaviour and Neurosciences, University of Groningen, P.O. Box 11103, 9700 CC Groningen, The Netherlands

37 ²¹ Department of Aquatic Ecology, Centre of Ecology, Evolution and Biogeochemistry, Eawag Swiss Federal Institute of Aquatic Science and Technology,
38 Kastanienbaum, Switzerland

39 ²² Museum of Vertebrate Zoology and Department of Integrative Biology, 3101 Valley Life Sciences Building, University of California, Berkeley, Berkeley, CA
40 94720-3160

41 ²³ Integrated Science Lab & Department of Ecology and Environmental Science, Umeå University, 90187 Umeå, Sweden

42 ²⁴ Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada

43 ²⁵ Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee 160, 53113 Bonn, Germany

44 ²⁶ Department of Biology, The University of Texas at Arlington, TX, USA

45 ²⁷ Computational and Molecular Population Genetics Lab, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland

46 ²⁸ Current address: Dept of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden

47 #core writing team
48

49 **Preface**

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

57

58 **Introduction**

59 Major insights into the genetics of speciation have come from a number of approaches (**Box 1**),
 60 ranging from the mapping of individual genes causing **reproductive isolation** (RI) to the
 61 characterization of genome-wide patterns of differentiation, and from quantitative genetic
 62 approaches to admixture analyses associating phenotypes with reduced **gene flow** between
 63 populations¹⁻³. These empirical approaches have a long history, starting with the work of
 64 Dobzhansky⁴ and Muller⁵. Theoretical understanding of the genetics of speciation has advanced
 65 markedly⁶⁻¹⁰. However, the deluge of empirical data coming from **next generation sequencing** (NGS),
 66 along with the emergence of new analytical approaches, necessitate the integration of this
 67 theoretical work to strengthen the conceptual foundations of the nascent field of **speciation**
 68 **genomics**. Such integration will help elucidate the relationships between evolutionary processes and
 69 genomic divergence patterns on the one hand, and between genomic properties and speciation
 70 processes on the other, and it will help unify research on the ecological and non-ecological causes of
 71 speciation.

72 In this review, we first discuss areas in which genomic approaches have begun to make important
 73 contributions to speciation research (**Box 1**), for example by elucidating patterns and rates of
 74 genome-wide divergence, improving our understanding of the genomic basis and evolution of
 75 intrinsic and extrinsic reproductive barriers, and identifying mechanisms by which different barriers
 76 become genomically coupled. We also highlight areas that would benefit from further attention;
 77 these areas include the distributions of **locus effect sizes**, **pleiotropy** and genomic constraint. We
 78 conclude by discussing how NGS data and innovative population genomic analyses could contribute
 79 to further progress in integrating these study areas into a more comprehensive and coherent
 80 understanding of the genomics of speciation.

81

82

83 **The evolution of reproductive barriers: Theory and classical evidence**

84 In line with others^{1, 3}, we define speciation as the origin of reproductive barriers among populations
 85 that permit maintenance of genetic and phenotypic distinctiveness of these populations in
 86 geographical proximity. The origin of reproductive barriers can either be initiated by **divergent**
 87 **selection** (that is, “ecological” or sexual selection creating **extrinsic reproductive isolation**), or by the
 88 evolution - through genetic drift, as an indirect consequence of selection or through genomic conflict
 89 - of genetic incompatibilities that cause **intrinsic reproductive isolation** (**Box 2**). Studying the
 90 accumulation of intrinsic isolation has a strong tradition in evolutionary biology^{1, 11}. Yet, most recent
 91 population genomic studies of divergence across the genomes of incipient and sister species have

92 investigated cases of putative **ecological speciation** and have focused on divergent adaptation and
 93 extrinsic isolation (but see¹² discussed below).

94
 95 Extrinsic postzygotic isolation arises as a consequence of divergent or **disruptive natural selection**
 96 when the viability or fertility of migrants or of individuals with intermediate genotypes is reduced².
 97 Prezygotic **sexual isolation** and also extrinsic postzygotic isolation, when hybrids have reduced
 98 mating success¹³, may evolve as a consequence of divergent sexual selection^{3, 14} which is often, but
 99 not always, mediated by differences in environments^{15, 16}. Prezygotic sexual isolation and extrinsic
 100 **postzygotic isolation** are, hence, dependent on genotype-environment interactions in the wider
 101 sense (where mating partners are part of the external environment). In contrast, intrinsic postzygotic
 102 isolation is independent of the external environment. Consequently, different types of genes and
 103 gene networks and different evolutionary processes may be involved in generating these classes of
 104 isolation. Extrinsic postzygotic isolation and sexual isolation can evolve rapidly¹⁷, and they often
 105 interact with each other¹⁶ and with the evolution of intrinsic postzygotic isolating barriers¹⁸ (Box 2).
 106 Selection can initiate speciation in situations with and without gene flow between populations, while
 107 intrinsic incompatibilities are less likely to accumulate when gene flow is present⁶. This being said,
 108 adaptive divergence and ecological speciation are not the same. Divergent adaptation alone rarely
 109 causes sufficient reproductive isolation to allow the accumulation or persistence of species
 110 differences in geographical proximity: this typically requires the evolution of **prezygotic isolation**^{1, 3}
 111 (Box 2), although it is possible that this varies between major taxonomic groups such as insects
 112 versus vertebrates or plants.

113
 114 The available evidence suggests that negative epistatic interactions, so called **Bateson-Dobzhansky-**
 115 **Muller incompatibilities** (BDMIs, or often just referred to as DMIs), are the most frequent cause of
 116 intrinsic postzygotic isolation^{1, 19-21}. However, other mechanisms, including **underdominance**²² and
 117 gene duplication, transposition and gene loss²³⁻²⁵ may also cause intrinsic postzygotic isolation. The
 118 time course of the accumulation of DMIs is not well understood^{19, 26-28}, and rates may vary among
 119 taxa and among mechanisms underlying DMI evolution¹⁹. DMIs were long thought to arise either as a
 120 consequence of genetic drift, as a result of stochastic deactivation of gene duplicates²⁹ or as a by-
 121 product of ecological selection³⁰. However, theoretical considerations, such as the slow pace of
 122 neutral accumulation of barriers³¹, and early empirical evidence for positive selection on loci
 123 contributing to incompatibilities³², suggested that drift was unlikely to be a common source of
 124 incompatibilities. Recent observations suggest instead that **intragenomic conflict** may be a common
 125 mechanism driving their evolution^{20, 33-35} (Fig. 1), as originally proposed in 1991^{34, 35}. Genomic conflict
 126 may arise from competing interests of males and females³⁶, from **meiotic drivers**^{37, 38}, mobile
 127 elements^{39, 40}, or other selfish genetic elements and their suppressors, and from competing interests
 128 between organellar and nuclear genomes^{41, 42}. **Sexual conflict** is thought to drive the evolution of
 129 new sex chromosomes^{43, 44}, and empirical observations suggest sex chromosome turnover has a role
 130 in the evolution of reproductive isolation^{45, 46}.

131
 132 The different evolutionary mechanisms underlying the build-up of extrinsic and intrinsic postzygotic
 133 and of prezygotic isolation suggest that genomic signatures will also be distinct. The genomic
 134 architecture of extrinsic isolation is likely to resemble that of adaptive population divergence, and be
 135 diverse and scattered across multiple regions in the genome (see below). However, there are
 136 theoretical arguments and empirical evidence for spatial clustering of sites under selection in the
 137 genome when adaptive evolution proceeds under prolonged bouts of divergent selection with
 138 migration or recurrent hybridization⁴⁷. For intrinsic isolation, incompatibility factors driven by
 139 genomic conflict are expected to accumulate in genomic regions of reduced recombination where
 140 linkage disequilibria between **distorter loci** and **responder loci** can become established^{48, 49}. Sex
 141 chromosomes are particularly susceptible to the accumulation of incompatibility factors derived from
 142 genomic conflict because sex chromosomes are constantly in a battle over segregation, whereas only
 143 small and tightly linked autosomal regions are in conflict with their homologs³⁴. At the same time,

144 there will be particularly strong selection for suppression of sex-linked distorter loci because they
145 tend to bias sex ratios^{50,51}. The genomic architecture of certain types of prezygotic isolation may also
146 be influenced by regions of reduced recombination around sex determining loci⁵² or sex
147 chromosomes⁵³, particularly when sex-linkage resolves sexually antagonistic effects of sexual
148 selection⁵⁴. Alternatively, prezygotic isolation loci may accumulate near extrinsic ecological isolation
149 loci (see section below, “**Genomic coupling of reproductive barriers**”). All of these signatures must
150 be distinguished from background patterns of genetic diversity and divergence that depend on the
151 populations’ history of genetic drift, gene flow, background selection and episodes of positive
152 selection unrelated to reproductive isolation.

153
154 Looking for signatures in the genetic architecture of reproductive isolation has a long “pre-genomic”
155 history (^{55,56}). However, there has been a historical disconnect between research programs focused
156 on intrinsic isolation, which have typically concentrated on later stages of speciation^{20,57}, versus
157 extrinsic postzygotic isolation and prezygotic sexual isolation at early stages of speciation^{2,30,15,16}.
158 Because of this disconnect, comparing the rates of evolution of components of reproductive
159 isolation, and their relevance to speciation, is currently a challenge. Where rates have been
160 compared in the same taxon using “pre-genomic” methods^{11,58-60}, the data suggest that prezygotic
161 and extrinsic postzygotic isolation often evolve faster than intrinsic postzygotic isolation, consistent
162 with expectations from classical theory⁶¹. Genome-wide data will now permit testing of this pattern
163 with a tremendous increase in resolution.

164
165

166 **Genomics and the speciation continuum**

167 Once speciation is complete, populations accumulate differences due to mutation and genetic drift
168 as well as ongoing selection. Reproductively isolated species, therefore, often differ in traits that
169 evolved under ecological selection and others that evolved under sexual selection, and may also have
170 intrinsic incompatibilities. A central task of speciation genetics is to reconstruct the sequence in
171 which these different barriers originated so as to distinguish between causes and consequences of
172 speciation. To achieve this, one would ideally take an unbiased view of the entire genome at all
173 stages of the same speciation process. However, speciation can rarely be studied in real time in
174 natural populations of sexually reproducing multicellular organisms. Estimates of variation among
175 loci in the timing and magnitude of gene flow could help determine the order in which reproductive
176 barriers emerged, but such inferences are challenging and current methods are not accurate enough
177 for this purpose⁶². However, by integrating case studies of closely related taxa that vary in their
178 extent of divergence (the “speciation continuum”), inferences can often be made about the
179 chronology and significance of different factors and processes at play.

180
181 Investigations of this “**speciation continuum**” have made important contributions to speciation
182 research^{63,64} and this approach is being adopted in NGS-based genome and transcriptome scan
183 studies of speciation. The major questions being addressed are the extent to which divergence at
184 different stages in speciation is either localized in the genome (the “island view”) or widespread, the
185 extent to which heterogeneity in divergence can be attributed to selective processes versus genetic
186 drift, the sources of selection, whether genomic divergence tends to follow a common trajectory as it
187 proceeds along the speciation continuum, and how all this is affected by the extent of geographical
188 isolation. A recently much cited scenario for speciation without strong geographical isolation, derived
189 from earlier models^{65,66}, involves an early stage of divergence where differentiation is limited to a
190 small number of loci (islands) under strong divergent selection. Gradually, these regions would grow
191 through the process of **divergence hitchhiking**, and eventually the effective migration rate would
192 become reduced globally across the genome fostering genome-wide divergence (‘genome
193 hitchhiking’)^{67,68}.

194

195 **Genome scans of ecological speciation**

196 Several NGS-based **genome scans** of the speciation continuum have found surprisingly variable
 197 patterns of genomic divergence. It appears that incipient species can quickly accumulate substantial
 198 divergence, even in the presence of gene flow (Fig. 2). However, whereas in some examples - such as
 199 *Heliconius* butterflies⁶⁹, *Helianthus* sunflowers⁷⁰, and poplar trees⁷¹ - divergence between **parapatric**
 200 ecotype populations is limited to a few large genomic regions, in others it is widespread across the
 201 genome⁷²⁻⁷⁵. NGS-based genome scans of **sympatric** sister species have generally reported
 202 genomically widespread and highly heterogeneous divergence that varies on a very local scale⁷⁵⁻⁸¹.
 203 Few studies have looked for evidence of divergence hitchhiking and the available results are
 204 inconsistent^{69, 76, 82}. Genome-wide average F_{ST} often increases as phenotypic divergence increases^{80, 83}
 205 but divergence seems to remain heterogeneous across the genome for a very long time, potentially
 206 due to repeated episodes of interspecific gene flow even after RI has become strong^{84, 85}. The first
 207 generation of NGS-based population-genomic studies of ecological speciation has therefore shown
 208 that ecological selection can cause strong isolation of small genomic regions between diverging
 209 populations, and that when RI is strong enough to permit persistence of incipient species in
 210 sympatry, many unlinked regions typically experience significant isolation.

211

212 So where does the heterogeneity in genomic divergence come from? It is commonly inferred to
 213 result from locus-specific differences in the effects of divergent selection and gene flow. Indeed,
 214 genome scans have shown strong isolation at genomic loci that were known to be under divergent
 215 selection^{64, 69, 70, 72, 74}. However, caution is warranted as different evolutionary processes can leave
 216 similar signatures in the genome. Heterogeneous genomic divergence is sometimes also observed
 217 between allopatric populations of the same species in the absence of any current gene flow^{76, 86, 87}
 218 (Fig. 2). Indeed, many studies assume ongoing gene flow between species, even though stochastic
 219 variation due to recent **coalescence** times and **incomplete lineage sorting** can similarly lead to low
 220 divergence and high heterogeneity, particularly when in combination with selection^{88, 89}. Statistical
 221 methods are available to distinguish divergence in isolation from divergence with gene flow, and
 222 these methods are increasingly being applied to genome scale datasets (reviewed in⁹⁰; Box 1).

223

224 Even in the absence of selection, divergence is expected to vary due to the stochasticity of genetic
 225 drift and the complexities of population history, and this variation can be enhanced by confounding
 226 effects of genomic heterogeneity⁹¹. In particular, regions of low recombination and/or high gene
 227 density often show reduced intra-specific diversity, which inflates relative divergence as measured by
 228 F_{ST} or D_a ⁸⁸. This can result from background selection against deleterious mutations⁹², intraspecific
 229 selective sweeps (in allopatry)⁸⁸ or even a direct influence of recombination on genetic diversity⁹³.
 230 Disentangling these processes is challenging⁹⁴. Some have suggested correcting for recombination
 231 rate in interpreting F_{ST} patterns⁸³. Others have suggested that absolute divergence measures such as
 232 D_{xy} are more robust to diversity artefacts⁹⁵, especially when corrected for local mutation rate⁹⁶. It
 233 seems unlikely that any single parameter will reliably disentangle divergent selection and gene flow
 234 from neutral processes. Good knowledge of the geographical context of population divergence will
 235 help, but distinguishing between hypotheses of speciation with gene flow, secondary contact and
 236 incomplete lineage sorting will often require new, parameter-rich modeling approaches⁹⁰.

237

238 Adaptive divergence has been shown to accumulate preferentially in regions of low recombination⁹⁷,
 239 including the centers of chromosomes⁸³, the vicinity of centromeres⁹⁸, inversions⁷⁴ or often (but not
 240 always^{12, 71}) on sex chromosomes⁹⁸⁻¹⁰⁰. Heterogeneity in genomic divergence seen in allopatry might
 241 also result from **gene-flow-selection balance** that has occurred in the past^{47, 76}. Finally, the
 242 assumption that the baseline F_{ST} reflects neutral divergence may be violated in cases where divergent
 243 selection is pervasive and multifarious, and this would bias against the detection of the signature of
 244 selection⁸¹.

245

246 Evidence for repeated divergence of the same genes or genomic regions across replicate pairs of
 247 species or environmental contrasts already provides strong evidence that these regions are indeed
 248 involved in adaptation and/or RI^{72, 74, 85, 97, 101-103}. Detecting such parallel divergence may require
 249 dense sampling of genomes or transcriptomes because the highest levels of repeatability may be
 250 observed at the scale of genomic regions rather than individual genes or SNPs⁹⁷. In this case, the
 251 repeatability in the heterogeneity of genomic divergence may be due at least in part to shared
 252 genomic heterogeneity in recombination and mutation rates rather than parallel adaptive
 253 divergence, but the shared genomic structure may facilitate the repeated accumulation in the same
 254 genomic regions of adaptive differentiation⁹⁷. Another approach involves combining classic **cline**
 255 theory with genome-wide analyses, allowing measurements of the strength of selection at specific
 256 loci⁷⁹ (Box 1). In the future, parameter-rich **coalescent** models of divergence with gene flow fitted to
 257 genomic data may be able to account for the heterogeneity of demographic history across the
 258 genome when seeking to identify genomic regions with reduced gene flow^{104, 105}. Finally, genome
 259 scans combined with manipulative selection⁸¹, QTL mapping^{82, 106}, candidate gene mapping^{72, 74} and
 260 **admixture mapping**^{79, 107-109} can be used to investigate whether divergent genomic regions contain
 261 loci contributing to RI.

262
 263 Several recent studies have found a contribution of ancient alleles to recent divergence, as
 264 exemplified by stickleback^{74, 110}, cichlids^{77, 111}, *Rhagoletis* flies¹¹² and *Heliconius* butterflies¹¹³. Ancient
 265 alleles are identifiable due to the accumulation of many substitutions or sharing over wide spatial or
 266 taxonomic ranges. The sources of such ancient allelic variation can either be **standing genetic**
 267 **variation**, or **hybridization**¹¹⁴. Distinguishing between these hypotheses is difficult in practice due to
 268 the challenges of distinguishing **incomplete lineage sorting** from hybridization⁹⁰ (Box 1). The balance
 269 of evidence from NGS data implies **introgressive hybridization** rather than standing variation as the
 270 source of ancient alleles in most of the above cases. Speciation in these cases might have been
 271 facilitated by hybridization providing genetic material for adaptation and reproductive isolation in
 272 the face of gene flow, which puts a new twist on an old idea¹. Future research combining genomic
 273 and ecological approaches should test this hypothesis further.

274 **Genomic divergence and intrinsic isolation**

275 Many studies have investigated DMI genes in strongly isolated species, but in many cases it remained
 276 unclear if the **fixation** of the underlying mutations was a cause or a consequence of speciation^{20, 57}.
 277 Regardless of whether identified DMI alleles are the first step in the origin of reproductive isolation, a
 278 striking pattern to emerge from recent work is that they have evolved under strong positive selection
 279 rather than genetic drift and that **genomic conflict** is often implicated as the source of this selection.
 280 For example, one study identified *Ovd*, an X-linked gene that underlies both hybrid male sterility and
 281 sex-ratio distortion in crosses between *Drosophila pseudoobscura pseudoobscura* and *D. p.*
 282 *bogotana*⁵¹. Another example is a recent analysis that found strong evidence for ongoing positive
 283 selection within *Drosophila mauritiana* in genes that have diverged between this species and its
 284 closest relatives and that are known to be involved in genomic conflict¹². Two pronounced
 285 polymorphism troughs on the X chromosome were centered on a pair of genes that cause sex-ratio
 286 distortion within *D. simulans*, and on *Odysseus*, a rapidly evolving homeobox gene that was known to
 287 cause male sterility in *D. mauritiana* x *D. simulans* hybrids³² and may be involved in genomic conflict.
 288 These are two candidate cases of speciation by conflict-driven DMI evolution.

289 **Genomic coupling of reproductive barriers**

290
 291 The build-up of associations between several traits or loci involved in RI strengthens the total barrier
 292 to gene flow between diverging populations, and is therefore important for the evolution of strong
 293 reproductive isolation^{115, 116}. Such **genomic coupling** can involve any pre- or post-zygotic barriers¹¹⁷.
 294 Deviations from linkage equilibrium between barrier loci can initially be generated by new mutations
 295 arising on a particular genetic background, or by genetic drift during divergence with limited gene
 296 flow. Coinciding barriers may, for example, arise through secondary contact between divergent
 297

298 populations, through the evolution of DMIs as an incidental by-product of divergent selection¹¹⁸, or
 299 via hitchhiking of intrinsic incompatibility alleles with divergently selected alleles, as has been shown
 300 for heavy-metal adapted populations of monkey flowers¹¹⁹. However, for barrier coupling to be
 301 important in speciation, coupling has to be maintained or even strengthened in the face of gene
 302 flow, and this typically requires divergent selection⁶.

303 Selection is expected to favour the coupling of barriers if this leads to an increase in mean fitness. In
 304 theory this can involve multiple intrinsic barriers (like DMIs)^{120, 121} or intrinsic and extrinsic
 305 postzygotic barriers as well as sexual and other prezygotic isolation traits. Across an **ecotone**,
 306 multifarious extrinsic selection can assemble and maintain many coinciding clines at loci involved in
 307 adaptation¹²², and these can become coupled with sexual isolation traits¹²³ and with DMIs^{18,116, 124}.
 308 Selection can also directly favour the evolution of increased prezygotic isolation, as in the case of
 309 **reinforcement**¹²⁵. Finally, sexual conflict can couple intrinsic postzygotic and prezygotic sexual
 310 isolation because DMIs driven by sexual conflict and genes underlying sexual traits or preferences
 311 expressed only in one sex may both accumulate on sex chromosomes^{53, 126}. Consistent with these
 312 expectations, loci for plumage colour, mating preferences and intrinsic postzygotic incompatibilities
 313 are coupled on the Z chromosome in flycatchers⁵² and Gouldian finches^{127, 128}. Similarly, loci for
 314 behavioural isolation and hybrid male sterility are coupled on the X chromosome in a species pair of
 315 Japanese stickleback⁴⁵.

316
 317 Because recombination tends to break up gene associations, genomic architectures that eliminate or
 318 decrease recombination are expected to facilitate coupling, and hence speciation¹²⁹. Most
 319 prominently, recombination will affect neither associations among traits that are pleiotropically
 320 influenced by the same allele, nor **'one-allele' mechanisms**, where the presence of the same allele in
 321 different genetic backgrounds confers RI¹³⁰. One-allele mechanisms do not leave a population-
 322 specific signature in the genome at the primary isolation locus but they should be detectable as
 323 **sweeps** shared by both diverging populations if they arise during speciation (as for instance if an
 324 allele for imprinting on the phenotype of the father spreads across two incipient species that were
 325 connected by gene flow). Despite the theoretical expectation that 'one-allele' mechanisms evolve
 326 more readily during speciation with gene flow than other types of barriers^{6, 16, 130}, we are not aware
 327 that the predicted genomic signature of shared sweeps at isolation loci has yet been detected in any
 328 case. Revealing such a signature would be a strong contribution of speciation genomics to
 329 demonstrating a classical prediction of speciation theory.

330
 331 Loci underlying **'two-allele' mechanisms** are expected to be concentrated in regions of reduced
 332 recombination. Recent genomic studies have observed genomic architectures that eliminate or
 333 reduce recombination between traits involved in RI: There is evidence of synergistic pleiotropy in
 334 **multiple-effect or "magic" traits**^{16, 131-133}, and multiple genes underlying isolating traits have been
 335 found together in inversions¹³⁴⁻¹³⁶, on sex-chromosomes^{45, 52, 128} and also in otherwise tight physical
 336 linkage^{119, 137}, including mating traits and mating preferences in cases of speciation with gene flow¹³⁸.
 337 These data also provide some evidence that reinforcement of prezygotic isolation is facilitated by
 338 linkage, as in flycatchers¹³⁹, or by pleiotropy, as in phlox¹³². In other cases reinforcement might be
 339 constrained¹⁴⁰ where loci are not linked and where there is extensive gene flow. However, recent
 340 genomic studies have also provided empirical examples of coupling between unlinked loci in fully
 341 **sympatric** hybridizing species⁷⁷ and especially in **hybrid zones**, where clines at many unlinked loci
 342 often coincide, although it is not always clear exactly how these loci are implicated in RI¹⁴¹. Unbiased
 343 whole-genome re-sequencing data and genome scans from diverging populations, coupled with
 344 methods to reduce bias from NGS data¹⁴² and with mapping of isolation traits, are needed to test the
 345 generality of these patterns.

346 **Effect sizes and pleiotropy**

347
 348 A key question, with a long history^{55, 143}, is whether speciation is typically initiated by divergence at
 349 few loci of large and possibly pleiotropic effect or by divergence at many loci with small and additive

350 effects^{133, 144}. The distinction is important because it will affect how speciation is constrained by the
 351 availability of suitable genetic variation, and will also affect how likely it is that selection or genetic
 352 drift may overcome gene flow. On their own, F_{ST} estimates from genome scans tell us little about the
 353 effect sizes of individual alleles on phenotypes, fitness or RI ¹⁰⁷ (Fig 3). With regard to fitness, Fisher's
 354 geometric model predicts that the probability that a mutation is favourable decreases exponentially
 355 with mutational effect size, hence we expect few alleles of large positive fitness effect but many of
 356 small effect¹⁴⁵⁻¹⁴⁷ (but see¹⁴⁸). However, this prediction does not take into account standing genetic
 357 variation, gene flow or changing environments. When those factors are considered, the predictions
 358 change^{47, 147, 149} and may even reverse¹⁵⁰.

359
 360 Speciation with gene flow may require divergent or disruptive selection to be concentrated on a
 361 small number of regions in the genome that also have large effects on RI ⁶. Theoretically expected
 362 distributions of effect sizes in terms of RI (rather than fitness) may be different for different classes of
 363 isolating barriers, but current data are equivocal (Fig. 3b). For example, mapping hybrid inferiority in
 364 natural environments for *Arabidopsis* has shown RI to be due to many genes with moderate
 365 effects¹⁵¹. In contrast, hybrid inviability in *Mimulus guttatus* is a consequence of two linked loci of
 366 major effect¹¹⁹. Predictions about the distribution of effect sizes expected for genes that underlie
 367 DMIs are also generally lacking, partly because effect sizes depend on mutation order and the extent
 368 of background genomic divergence. Traits governing prezygotic isolation, and especially sexual
 369 isolation (Box 2), are likely to have large effects on RI because they directly influence mating or
 370 fertilization patterns^{1, 6, 16, 152-154}. To test this prediction with genomic scale data, existing quantitative
 371 genetic, mapping and candidate gene studies^{45, 109, 111, 128, 138, 155-157, 158, 159} should now be followed up
 372 by NGS-based genome scans assessing RI around these loci¹⁰⁷.

373
 374 Recently identified large-effect alleles involved in adaptation and speciation with gene flow, are
 375 often highly **pleiotropic** (e.g., *Optix* in *Heliconius*¹⁶⁰ and Ectodysplasin [*Eda*] in sticklebacks¹⁶¹,
 376 although we lack estimates of the effect *Eda* has on RI or fitness). Such alleles may be rare among
 377 newly arising mutations but alleles with synergistically pleiotropic effects may be more common in
 378 standing genetic variation. Recent theory suggests that large-effect or pleiotropic alleles may be
 379 favoured by selection during evolution in gene-flow-selection balance, and hence eventually become
 380 enriched in taxa with divergence and gene flow⁴⁷.

381
 382 **Genomic constraint**

383 The flipside of the coupling problem is that genetic correlation between traits as a result of
 384 pleiotropy or tight linkage may also constrain speciation. With new population genomic data
 385 revealing divergence in many regions of the genome early in speciation, there is an opportunity to
 386 unite population genomics with a quantitative genetics perspective on the evolution of polygenic
 387 traits during speciation. In quantitative genetics terms, standing genetic variation is quantified by the
 388 **G-matrix** of additive genetic variance and covariance¹⁶². **G** may indicate potential constraints on
 389 adaptive evolution that affect the response to directional selection^{163, 164}, as well as constraints on
 390 genetic drift¹⁶⁵. Tests to detect the impact of selection on **G** are available (e.g.¹⁶⁶). Divergence among
 391 populations is biased along axes with greater genetic variation and covariation and constrained along
 392 axes with little variation or covariation^{164, 167, 168}. Importantly, however, genetic constraints are not
 393 only negative. Genetic covariation may align with **correlational selection**^{169, 170} and, as discussed
 394 above, pleiotropy can couple adaptation to RI . It is not known how genes of major effect, versus the
 395 traditional assumption of many genes of small effect, influence the structure of **G**¹⁷¹, and how higher
 396 moments of the distribution of genetic variation and covariation affect the response to divergent
 397 selection¹⁷². These questions can now be addressed with genomic methods, such as directly
 398 estimating **G** in outbred populations using NGS data¹⁷³. A different approach is to estimate the
 399 genetic variance-covariance matrices for gene regulatory networks from gene expression data.
 400 Analyzing genomic data in a quantitative genetics framework in this way will illuminate how genomic
 401 constraint affects speciation¹⁷⁴.

402
 403 Studying effects of hybridization is one promising application. Beyond being a source of allelic
 404 variation, hybridization may facilitate evolution and perhaps speciation by releasing populations from
 405 constraints caused by genetic correlations. While empirical evidence has accumulated that suggests
 406 that selection alters genomic architecture^{169, 175}, the role of gene flow in aligning **G** with the direction
 407 of divergent or disruptive selection has rarely been investigated¹⁵⁰. The emerging consensus that
 408 hybridization frequently introduces adaptive variation¹⁸ calls for empirical studies in this area. We
 409 predict that hybridization will influence speciation not only by generating novel and **transgressive**
 410 **phenotypes** but also by aligning **G** with the axis of divergent selection (Fig. 4a). Even when early
 411 generation hybrids are maladapted, hybrid populations may over time benefit from increased
 412 evolvability¹⁷⁶. Hybridization may alter patterns of genetic covariance much faster than is possible by
 413 selection alone, and may lead to bursts of evolutionary diversification and speciation^{114, 177} (Fig. 4b-d).
 414 Genomic methods can now be used in assessing these hypotheses in several ways, such as direct
 415 estimation of **G** in both parental and hybrid natural populations and through association or
 416 **admixture mapping** of loci contributing to novel adaptive phenotypes in hybrid populations¹⁰⁸.

419 **Speciation genomics: towards a synthesis**

420 Speciation can proceed in many different ways, but these can be grouped in terms of drivers (drift
 421 and different types of selection), causes (extrinsic environment-dependent versus intrinsic
 422 environment-independent) and stage in the life cycle (postzygotic or prezygotic) of reproductive
 423 isolation, resulting in two major classes that are at least in theory quite distinct (Box 2). In one, RI is
 424 initiated by extrinsic selection, in the other by intrinsic incompatibility. Analysis of NGS data has
 425 begun to shed light on the signatures of these processes in the genome. Both of these classes of
 426 processes can generate reproductively isolated species in allopatry, but parapatric and especially
 427 sympatric speciation are constrained to situations where divergent natural and/or sexual selection
 428 overcome the homogenizing effects of gene flow^{1, 6}. Whether speciation in such scenarios can
 429 proceed depends on the strength of selection^{2, 6} and the genetic architecture of adaptation and
 430 reproductive isolation^{76, 122}. Speciation driven by genomic conflict is much less likely to be initiated in
 431 the presence of gene flow because selfish genetic elements may then spread across populations and
 432 thereby prevent or slow down the accumulation of conflict-driven DMIs¹⁷⁸. However, it remains
 433 possible that relatively brief periods of **allopatry** are sufficient for the origins of conflict-driven DMIs.
 434 Although DMIs may be removed by selection after **secondary contact**, they may, in theory, facilitate
 435 speciation if they become coupled with other components of RI before they are purged^{116, 179}. How
 436 often this happens is unknown.

437
 438 These principles are not new¹, but they can and should now be examined with much greater
 439 resolution using genomic methods. Although speciation genomics is clearly still in its infancy, a few
 440 trends are emerging from the first generation of NGS-based genome scans, particularly in relation to
 441 non-allopatric speciation: The available evidence suggests that divergence can be genomically
 442 widespread very early in speciation, and may generally be so in species that coexist in full sympatry<sup>74-
 443 77, 80</sup>, whereas it can be restricted to very few islands of divergence in parapatric ecotypes^{69, 70}.
 444 Perhaps **multifarious divergent selection** or genomically widespread selection is important to
 445 generate sufficient RI to permit maintenance and perhaps buildup of genetic differentiation in
 446 sympatry. More data are now needed to confirm this intriguing pattern.

447
 448 Some genomic regions that are divergent between incipient and sibling species in geographical
 449 proximity contain genes with large effects on adaptation and pleiotropic effects on prezygotic
 450 isolation. The alleles at several such loci have turned out to be ancient variants that were present as
 451 standing variation or were brought together by hybridization in the ancestors of emerging species

452 pairs^{99, 111, 112}. Although it is premature to draw strong conclusions, this may turn out to be another
453 emergent feature of speciation with gene flow. We expect effect sizes to be larger, antagonistic
454 pleiotropy to be less frequent and synergistic pleiotropy to be more frequent in ancient alleles that
455 have been honed by selection over time than in alleles arising newly through mutation. We
456 hypothesize that substitution of such ancient alleles at major effect loci has the potential to reduce
457 gene flow quickly, to the point where substitutions with smaller effects at other loci can also spread.
458 Genome scans of divergence very early in the speciation continuum (at low overall RI, Box 2) should
459 allow explicit tests of these hypotheses.

460
461 Alternative mechanisms and geographical modes of speciation make different predictions for
462 patterns in genomic data. Specifically, we predict that speciation due to conflict-driven DMIs involves
463 greater divergence at centromeres and sex chromosomes, and so these regions should bear
464 signatures of selective sweeps. Divergence under ecological selection may be more widely
465 distributed across the genome, and sweeps at individual loci less pronounced. The available data are
466 consistent with these expectations, although theory predicts accumulation of genes for ecological
467 divergence in regions of low recombination when selection is antagonized by gene flow¹²⁹.
468 Divergence by sexual selection may be concentrated on sex chromosomes⁵², but support for this
469 prediction is not always found and predictions vary with the sex determination system⁵⁴. Many more
470 population genomic studies of divergence in a wider range of taxa and across a greater range of
471 points along the speciation continuum are needed to test these predictions further. Speaking more
472 broadly, future work should seek to determine to what extent different evolutionary mechanisms
473 and geographical modes of speciation can be distinguished based on genomic data and, in turn, the
474 extent to which genomic features can predict the modes and mechanisms of speciation that apply to
475 a given evolutionary lineage.

476 Taxonomic variation in the propensity for speciation without geographical isolation is prevalent¹⁸⁰
477 and it will be interesting to learn if variation in genomic architecture explains some of this. Whether
478 selection can overcome gene flow depends, besides the total strength of selection, on the number of
479 genome regions targeted by selection, on the rate of recombination between them, and on the
480 extent of pleiotropy. When analyzed in conjunction with ecological data, genomic data therefore
481 hold promise to help explain why non-allopatric speciation occurs readily in cichlid fish, whitefish,
482 stickleback, *Rhagoletis* flies, *Heliconius* butterflies, *Coprosma* shrubs¹⁸¹ and some other taxa, but is
483 not reported in the majority of others. This combination of approaches may also help more generally
484 to explain why some taxa undergo speciation and accumulate species diversity a lot more readily
485 than others. Answering such questions will also facilitate an understanding of larger-scale patterns in
486 species diversity (Box 3).

487
488 Population-genomic studies that explicitly compare rates of evolution and the genomic distribution
489 of prezygotic, extrinsic postzygotic and intrinsic postzygotic barriers to gene flow have yet to
490 materialize. We believe that such studies hold considerable promise to overcome old dichotomies in
491 speciation genetics. Because the discovery of DMIs used to be laborious, we cannot yet answer the
492 question how often DMIs are caused by conflict, ecological selection or genetic drift. This too will
493 hopefully soon change as genomic data allow the identification of DMI loci at an increasing pace^{12, 26}
494 (Box 1).

495
496 A still missing part of a synthesis in speciation genomics is the integration of evolutionary
497 developmental biology. Insights from this field make several relevant suggestions for speciation
498 genomics¹⁸²: First, mutations in coding sequences may more often have pleiotropic effects than
499 those in *cis*-regulatory regions. Second, pleiotropy will be more common when selection targets
500 genes with central roles in gene regulatory networks, and many morphological developmental genes
501 are in such positions. Third, because of the first two predictions, morphological evolution may often
502 be constrained to take place through changes in *cis*-regulatory mutations, whereas physiology may

503 be more free to evolve through coding mutations. These predictions make for interesting yet little
504 explored connections between some of the above discussed questions in speciation research and the
505 debate about the prevalence of coding versus *cis*-regulatory mutations in evolution^{182, 183}. Possible
506 ascertainment bias notwithstanding, empirical data suggest that divergence between sibling species
507 and conspecific populations is predominantly due to evolution of coding genes, independent of their
508 positions in gene regulatory networks, but morphological differences between species that diverged
509 longer ago are predominantly due to *cis*-regulatory evolution¹⁸². The following explanation has been
510 offered: Selection acting early during population divergence may partly overcome the negative
511 fitness effects of antagonistic pleiotropy that are expected for coding mutations, but may not be
512 strong enough to fix these mutations¹⁸². Over time, as more mutations become available, *cis*-
513 regulatory mutations with more specific effects and less antagonistic pleiotropy would replace the
514 coding variants. An interesting implication is that the mutations responsible for phenotypic
515 differences between older species may be distinct from those that are causally important in the
516 process of population divergence and speciation, even when the mechanism of speciation and the
517 diverging phenotypes are the same. Studies of the genomic basis (coding versus regulatory) of
518 species divergence in incipient versus older species in the same taxon are needed to test this
519 hypothesis. We are not aware that such data exist.

520

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

524

525 First, there is a need for more comparative genome scans at different stages in the speciation
526 continuum in closely related taxa or in replicate species pairs in the same taxon. These data need
527 to be combined with annotation of the effects of alleles on phenotypes and on RI, which can be
528 done through QTL mapping or functional analyses in the context of annotated reference
529 genomes. This would allow the association of divergent genomic regions with mechanisms of RI.
530 Such studies need to be repeated in the following scenarios: in taxa in which speciation is driven
531 by ecology, sexual selection and intrinsic incompatibilities (Box 2); in different spatial contexts;
532 and in taxa that have not speciated, but that occupy similar environments to those taxa that have
533 undergone speciation. Sampling design should explicitly aim to explore variation, both in
534 different stages on the speciation continuum and for different degrees of geographical isolation
535 (Fig. 2), and the history of geographical isolation should ideally be known. Eventually, with
536 replication and clever experimental and comparative study designs, it will become possible to
537 understand whether different mechanisms and modes of speciation can be distinguished based
538 on patterns observed in genome-wide data.

539

540 Second, experimental **population genomics** studies of speciation are needed to measure the
541 strength and multifarious nature of selection, and more generally to test hypotheses about
542 processes underlying differentiation and isolation, including intragenomic conflict, heterogeneity
543 in recombination rates, and coupling.

544

545 Third, theoretical modeling is needed that includes the influences of variable demography,
546 recombination rates and time, and explicitly considers standing genetic variation and different
547 sources of incompatibilities. Such models will be helpful in generating predictions that can be
548 tailored to individual empirical study systems to make them testable. Such predictions could
549 include genomic signatures of alternative speciation modes and mechanisms, and how modes
550 and mechanisms can be inferred from patterns found in genomes at different stages of the
551 speciation continuum. Improved methods for estimating the timing of long-term gene flow
552 would also be very valuable⁹⁰. Given the increasingly widespread evidence for recruitment of
553 ancient genetic variation into recent speciation events, analytical methods for rigorous
554 hypothesis-testing regarding the source of such variation – that is, the contributions of

555 hybridization and standing genetic variation – are also needed. Such methods could include
556 comparisons of the phylogenetic histories of genomic regions that confer adaptation and
557 reproductive isolation with those of other segments of the genomes of young sister species^{74, 77,}
558 ^{99, 112}.
559
560

561 **Conclusions**

562 New approaches for gathering large amounts of genomic data in non-model organisms have begun
563 to produce intriguing and unexpected insights into the genetics of speciation. Sympatrically
564 coexisting species are characterized by heterogeneous differentiation that is widely scattered across
565 the genome even when these species are still very young, but adaptive differentiation between
566 parapatric populations can be restricted to a few genomic islands. Ancient alleles with large and
567 pleiotropic effects characterize both types of divergence, and were often acquired by interspecific
568 hybridization. Genomic conflict may be a frequent source of intrinsic postzygotic isolation. It may be
569 recognized in genome scans as strong sweep signatures on sex chromosomes or in isolated islands of
570 divergence on autosomes. More strongly integrated studies are now needed that cover multiple
571 components of RI at multiple stages of the speciation continuum, and in geographical settings
572 ranging from complete allopatry to full sympatry, paying additional attention to the history of
573 population contact (primary or secondary). With the rapid growth of genomic data generation and
574 analysis approaches, it will then soon become possible to construct an integrated picture of
575 speciation starting from the evolution of reproductive barriers and how this is influenced by
576 ecological and genomic constraints, through the way speciation creates signatures of genomic
577 divergence, to how genomic properties of organisms interact with history and ecology in shaping
578 patterns in biodiversity. There is no doubt that a new phase of discovery has begun that will usher in
579 a greatly increased understanding of the origin of species.

580
581 Author contributions. This paper was initiated during a workshop funded by the ESF networking
582 program Frontiers in Speciation Research (FroSpects). OS led workshop organization and
583 coordination, and manuscript preparation with assistance from CEW, IK and RKB. JWB, PAH, CLP, G-
584 PS, CEW and IK led discussion groups and initial drafting of sections of the paper, ETW, CDJ, CSC,
585 SHM, JWB, JS and CEW prepared figures, OS and RKB drafted general sections, other authors
586 contributed during the workshop and commented on drafts.

587
588 Acknowledgements. We thank the European Science Foundation (networking program Frontiers in
589 Speciation Research, FroSpects) for funding a workshop on “Genetics and Genomics of Speciation”
590 and for contributing towards publication costs, and Nadja Pepe and Lucie Greuter for help with the
591 organization of the workshop. We thank Christian Lexer and two anonymous reviewers for
592 constructive suggestions that improved our paper.
593

594	Glossary
595	Items included in the glossary are bolded in their first appearance in the main text.
596	
597	Admixture mapping
598	Identification of genetic loci that contribute to phenotypic differences between ancestral
599	populations, by exploring genotype-phenotype correlations in a population of mixed ancestry.
600	
601	Allopatric
602	Organisms, populations or species inhabiting distinct geographical regions and therefore not
603	exchanging genes.
604	
605	Allopatry
606	Occurrence in geographically isolated regions.
607	
608	Cline
609	Directional variation in phenotype or genotype, or change in frequency (e.g. of an allele), across a
610	geographic region.
611	
612	Coalescence
613	The merging of two genetic lineages in a common ancestor.
614	
615	Coalescent
616	A statistical framework for the analysis of genetic data where the genotypes shared by populations
617	or species are traced back in time to their most recent common ancestor.
618	
619	Correlational selection
620	Selection for optimal character combinations.
621	
622	Disruptive selection
623	Selection within a single population that favours extreme phenotypes over intermediate phenotypes.
624	
625	Distorter loci
626	Loci underlying meiotic drive , the non-Mendelian segregation of alleles in meiosis. Distorter loci may
627	act on other loci, so-called responder loci.
628	
629	Divergence hitchhiking (DH)
630	Occurs when divergent selection on a locus reduces the effective migration rate for physically linked
631	regions, which increases the opportunity for divergence at loci under weaker selection in these
632	surrounding regions. DH regions may remain much larger than traditional hitchhiking regions after a
633	selective sweep within populations because of the persistent reduction in the ability of flanking
634	regions to recombine away from a divergently selected gene.
635	
636	Divergent selection
637	Selection favouring different phenotypes in different populations.
638	
639	D_{xy}
640	The average number of nucleotide substitutions per site between two populations.
641	
642	Bateson-Dobzhansky-Muller Incompatibility (BDMI or mostly just referred to as DMI)
643	An intrinsic postmating barrier that is the result of epistatic interactions between alleles at two or
644	more loci that cause reduced fitness in hybrids but not in the parental populations.
645	

646 **Ecological speciation**

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

649

650 **Ecotone**

651 A zone where there is a transition between two distinct biological communities, e.g. between forest
652 and grassland or aquatic and terrestrial habitats. Ecotones are typically associated with changes in
653 the physical environment.

654

655 **Extrinsic reproductive isolation**

656 Fitness reduction in hybrids that is dependent on the environment, i.e. mediated by genotype-
657 environment interactions.

658

659 **Fixation**

660 Describes the situation in which a mutation or variant has achieved a frequency of 100% in a
661 population.

662

663 **F_{ST}**

664 A measure of population subdivision that compares the correlation between two gene copies that
665 are randomly drawn from the same population to that between two gene copies drawn from two
666 different populations. An F_{ST} of 1 indicates that two populations are fixed (**fixation**) for alternative
667 alleles.

668

669 **F_{ST} -outlier analysis**

670 Comparison of the distribution of F_{ST} values across loci with the distribution expected in the absence
671 of divergent selection for the same average differentiation. Loci whose F_{ST} values exceed expectation
672 are likely to be influenced by divergent selection, either on the locus itself or on a linked locus.

673

674 **Gene flow**

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

678

679 **Gene-flow-selection balance**

680 A level of differentiation between sub-populations at which the homogenizing effect of gene flow
681 and the differentiating effect of divergent selection are in equilibrium.

682

683 **Genome scan**

684 Comparison of genome-wide patterns of diversity within populations and/or divergence between
685 populations at hundreds or thousands of markers. Most studies until recently used Amplified
686 Fragment Length Polymorphisms (AFLPs) but this has recently changed, and SNPs generated by NGS
687 or SNP chips are being used.

688

689 **Genomic conflict**

690 Genomic conflict arises between genes or genetic elements within the same genome when these are
691 not transmitted by the same rules (e.g. biparental vs uniparental inheritance), or when a gene causes
692 its own transmission to the detriment of the rest of the genome. The presence of elements (distorter
693 loci) that bias transmission is expected to lead to the evolution of loci that restore Mendelian
694 segregation (restorer loci).

695

696 **Genomic coupling**

697 The statistical association between different traits and loci involved in RI.

- 698
699 **G-matrix**
700 The additive genetic variance-covariance matrix that summarizes the variances within and
701 covariances between multiple phenotypic traits.
702
- 703 **Hybridization**
704 Mating between individuals that belong to distinct species or populations. If postmating isolation is
705 incomplete, hybridization leads to the introgression of genes from one population to the other.
706
- 707 **Hybrid zones**
708 Spatially restricted regions where the distribution ranges of distinct populations or incipient species
709 come into contact and hybrids are formed.
710
- 711 **Incomplete lineage sorting**
712 Situation in which some alleles share a more recent common ancestor with alleles in another species
713 than with other alleles in the same species.
714
- 715 **Intragenomic conflict**
716 Antagonistic selection among genomic elements with different fitness interests in an individual.
717
- 718 **Intrinsic reproductive isolation**
719 Fitness reduction in hybrids that is independent of the environment.
720
- 721 **Introgressive hybridization**
722 The introduction of genes from one population or species into another through hybridization.
723
- 724 **Linkage disequilibrium**
725 The statistical association of the alleles at two loci within gametes in a population. Although linkage
726 disequilibrium tends to be greater between linked loci, it can also arise between physically unlinked
727 loci — for example, because of selection, non-random mating or gene flow.
728
- 729 **Locus or allele effect size**
730 The magnitude of the influence of a locus, or a specific allele, on a phenotypic trait. This can be
731 expressed, for example, as the proportion of phenotypic variation attributable to a specific locus or
732 the phenotypic difference between genotypes with and without a specific allele.
733
- 734 **Multifarious divergent selection**
735 Divergent selection acting on multiple traits.
736
- 737 **Multiple-effect traits or “magic” traits**
738 Traits that contribute to more than one component of reproductive isolation, such as a trait
739 contributing to local adaptation that is also used as a mating cue.
740
- 741 **Meiotic drivers**
742 Factors distorting Mendelian segregation. At a heterozygous site, the driving variant will be found in
743 more than half of the gametes.
744
- 745 **Next Generation Sequencing**
746 A class of high-throughput sequencing methods that rely on technologies that parallelize the
747 sequencing process, producing thousands or millions of sequences concurrently. Next Generation
748 Sequencing technologies increase throughput and lower the cost of DNA sequencing by orders of
749 magnitude compared to standard dye-terminator methods.

- 750
- 751 **One-allele mechanism**
- 752 Reproductive barriers arise through spreading of the same allele in each of two diverging
- 753 populations, such as an allele for behavioural imprinting or reduced migration.
- 754
- 755 **Parapatric**
- 756 Organisms, populations or species that inhabit adjacent geographical regions or spatially distinct but
- 757 adjacent habitats and may exchange genes.
- 758
- 759 **Pleiotropy**
- 760 Effect of an allele on more than one trait.
- 761
- 762 **Prezygotic isolation**
- 763 Effect of barriers acting before or after mating but before fertilisation, including the isolating effects
- 764 of divergent mate choice, habitat preference, reproductive timing and gametic incompatibility.
- 765
- 766 **Population genomics**
- 767 Use of genome-wide data (typically based on next-generation sequencing methods) to make
- 768 inferences about evolutionary processes in natural populations.
- 769
- 770 **Postzygotic isolation**
- 771 Effects of barriers acting after fertilisation, such as hybrid sterility and hybrid inviability. Can be
- 772 extrinsic (mediated by the environment) or intrinsic.
- 773
- 774 **Quantitative trait locus (QTL)**
- 775 Chromosomal region with a statistically significant effect on a phenotype.
- 776
- 777 **Reinforcement**
- 778 Selection for the strengthening of prezygotic barriers to avoid the production of unfit hybrids
- 779 between taxa that have previously evolved some postzygotic isolation.
- 780
- 781 **Reproductive isolation**
- 782 Absence or restriction of gene flow between populations over and above that due to spatial
- 783 separation alone.
- 784
- 785 **Responder loci**
- 786 Loci showing deviations from Mendelian segregation (meiotic drive) due to the effect of a distorter
- 787 locus.
- 788
- 789 **Secondary contact**
- 790 The meeting of the distribution ranges of two distinct populations or species after a period of
- 791 evolutionary divergence in geographical isolation (allopatry).
- 792
- 793 **Sexual conflict**
- 794 The evolution of phenotypic characteristics by sexual selection, when the trait confers a fitness
- 795 benefit to one sex but a fitness cost to the other.
- 796
- 797 **Sexual isolation**
- 798 Reproductive isolation as a consequence of reduced mating between members of divergent
- 799 populations, including behavioural assortative mate choice and assortative fertilization in animals, as
- 800 well as pollinator-mediated assortative mating in plants. Most often thought of as prezygotic, but can
- 801 be postzygotic if there is disruptive sexual selection.

802

803 **Speciation continuum**

804 Pattern where the strength of reproductive isolation between two incipient species varies in
805 different locations or in different species pairs that belong to the same evolutionary lineage and
806 diverge in similar ways.

807

808 **Speciation genomics**

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

813

814 **Standing genetic variation**

815 Allelic variation that is currently segregating within a population; as opposed to alleles that arise
816 through new mutation events.

817

818 **Sweep**

819 Increase in frequency of an allele and closely linked chromosomal segments due to positive selection.
820 Sweeps initially reduce variation and subsequently lead to a local excess of rare alleles as new unique
821 mutations accumulate.

822

823 **Sympatric**

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

826

827 **Transgressive phenotypes**

828 Expression of phenotypic variation in hybrids that exceeds the range of phenotypes observed in the
829 parental taxa.

830

831 **Two-allele mechanism**

832 Reproductive barriers arise through spreading of different alleles at the same locus in two diverging
833 populations, such as alleles for different habitat or mating preferences.

834

835 **Underdominance**

836 Heterozygote inferiority. The phenotype expressed in heterozygotes has lower fitness than that of
837 either homozygote. Underdominance can be a cause of **disruptive selection**.

- 838 1. Coyne, J. & Orr, H. Speciation (Sinauer Associates, Sunderland, M.A., 2004).
- 839 2. Nosil, P. Ecological Speciation (Oxford University Press, Oxford and New York, 2012).
- 840 3. Price, T. Speciation in Birds (Roberts & Company, Greenwood Village, Colorado, 2008).
- 841 4. Dobzhansky, T. Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila*
- 842 *pseudoobscura* hybrids. *Genetics* **21**, 113–135 (1936).
- 843 5. Muller, H.J. & Pontecorvo, G. Recessive genes causing interspecific sterility and other
- 844 disharmonies between *Drosophila melanogaster* and *simulans*. *Genetics* **27**, 157 (1942).
- 845 6. Gavrillets, S. Fitness landscapes and the origin of species. *Princeton University Press* (2004).
- 846 7. van Doorn, G.S., Edelaar, P. & Weissing, F.J. On the Origin of Species by Natural and Sexual
- 847 Selection. *Science* **326**, 1704-1707 (2009).
- 848 8. M'Gonigle, L.K., Mazzucco, R., Otto, S.P. & Dieckmann, U. Sexual selection enables long-term
- 849 coexistence despite ecological equivalence. *Nature* **484**, 506-509 (2012).
- 850 9. Gavrillets, S. & Losos, J.B. Adaptive Radiation: Contrasting Theory with Data. *Science* **323**, 732-
- 851 737 (2009).
- 852 10. Dieckmann, U., Doebeli, M., Metz, J.A.J. & Tautz, D. (eds.) Adaptive Speciation (Cambridge
- 853 University Press, 2004).
- 854 11. Coyne, J.A. & Orr, H.A. "Patterns of speciation in *Drosophila*" revisited. *Evolution* **51**, 295-303
- 855 (1997).
- 856 12. Nolte, V., Pandey, R.V., Kofler, R. & Schlötterer, C. Genome-wide patterns of natural variation
- 857 reveal strong selective sweeps and ongoing genomic conflict in *Drosophila mauritiana*.
- 858 *Genome Research* **23**, 99-110 (2013).
- 859 13. van der Sluijs, I. et al. Female mating preference functions predict sexual selection against
- 860 hybrids between sibling species of cichlid fish. *Philosophical Transactions of the Royal Society*
- 861 *B-Biological Sciences* **363**, 2871-2877 (2008).
- 862 14. Panhuis, T.M., Butlin, R., Zuk, M. & Tregenza, T. Sexual selection and speciation. *Trends in*
- 863 *Ecology & Evolution* **16**, 364-371 (2001).
- 864 15. Boughman, J.W. How sensory drive can promote speciation. *Trends in Ecology & Evolution*
- 865 **17**, 571-577 (2002).
- 866 16. Maan, M.E. & Seehausen, O. Ecology, sexual selection and speciation. *Ecology Letters* **14**,
- 867 591-602 (2011).
- 868 17. Hendry, A.P., Nosil, P. & Rieseberg, L.H. The speed of ecological speciation. *Functional*
- 869 *Ecology* **21**, 455-464 (2007).
- 870 18. Abbott, R. et al. Hybridization and speciation. *Journal of Evolutionary Biology* **26**, 229-246
- 871 (2013).
- 872 19. Presgraves, D.C. Speciation Genetics: Search for the Missing Snowball. *Current Biology* **20**,
- 873 R1073-R1074 (2010).
- 874 20. Presgraves, D.C. The molecular evolutionary basis of species formation. *Nature Reviews*
- 875 *Genetics* **11**, 175-180 (2010).
- 876 21. Rieseberg, L.H. & Blackman, B.K. Speciation genes in plants. *Annals of Botany* **106**, 439-455
- 877 (2010).
- 878 22. Rieseberg, L.H. & Burke, J.M. A genic view of species integration - Commentary. *Journal of*
- 879 *Evolutionary Biology* **14**, 883-886 (2001).
- 880 23. Bikard, D. et al. Divergent evolution of duplicate genes leads to genetic incompatibilities
- 881 within *A. thaliana*. *Science* **323**, 623-626 (2009).
- 882 24. Masly, J.P., Jones, C.D., Noor, M.A.F., Locke, J. & Orr, H.A. Gene transposition as a cause of
- 883 hybrid sterility in *Drosophila*. *Science* **313**, 1448-1450 (2006).
- 884 25. Ting, C.T. et al. Gene duplication and speciation in *Drosophila*: Evidence from the Odysseus
- 885 locus. *Proceedings of the National Academy of Sciences of the United States of America* **101**,
- 886 12232-12235 (2004).
- 887 26. Matute, D.R., Butler, I.A., Turissini, D.A. & Coyne, J.A. A Test of the Snowball Theory for the
- 888 Rate of Evolution of Hybrid Incompatibilities. *Science* **329**, 1518-1521 (2010).

- 889 27. Moyle, L.C. & Nakazato, T. Hybrid incompatibility “snowballs” between *Solanum* species.
890 *Science* **329**, 1521-1523 (2010).
- 891 28. Stadler, T., Florez-Rueda, A.M. & Paris, M. Testing for "Snowballing" Hybrid Incompatibilities
892 in *Solanum*: Impact of Ancestral Polymorphism and Divergence Estimates. *Molecular Biology*
893 *and Evolution* **29**, 31-34 (2012).
- 894 29. Lynch, M. & Force, A.G. The origin of interspecific genomic incompatibility via gene
895 duplication. *American Naturalist* **156**, 590-605 (2000).
- 896 30. Schluter, D. Evidence for Ecological Speciation and Its Alternative. *Science* **323**, 737-741
897 (2009).
- 898 31. Nei, M., Maruyama, T. & Wu, C.I. Models of evolution of reproductive isolation. *Genetics* **103**,
899 557-579 (1983).
- 900 32. Ting, C.T., Tsaur, S.C., Wu, M.L. & Wu, C.I. A rapidly evolving homeobox at the site of a hybrid
901 sterility gene. *Science* **282**, 1501-1504 (1998).
- 902 33. Crespi, B. & Nosil, P. Conflictual speciation: species formation via genomic conflict. *Trends in*
903 *Ecology & Evolution* **28**, 48-57 (2013).
- 904 34. Frank, S.A. Divergence of meiotic drive-suppression systems as an explanation for sex-biased
905 hybrid sterility and inviability. *Evolution* **45**, 262-267 (1991).
- 906 35. Hurst, L.D. & Pomiankowski, A. Causes of sex ratio bias may account for unisexual sterility in
907 hybrids: a new explanation of Haldane's rule and related phenomena. *Genetics* **128**, 841-858
908 (1991).
- 909 36. Cocquet, J. et al. A genetic basis for a postmeiotic X Versus Y chromosome intragenomic
910 conflict in the mouse. *PLOS Genetics* **8**, e1002900 (2012).
- 911 37. Malik, H.S. in *Progress in Molecular and Subcellular Biology* (ed. Ugarovic, D.) 33-52
912 (Springer, 2009).
- 913 38. Campbell, P., Good, J.M. & Nachman, M.W. Meiotic sex chromosome inactivation is
914 disrupted in sterile hybrid male house mice. *Genetics* **193**, 819-828 (2013).
- 915 39. Rebollo, R., Horard, B., Hubert, B. & Vieira, C. Jumping genes and epigenetics: Towards new
916 species. *Gene* **454**, 1-7 (2010).
- 917 40. Watson, E.T. & Demuth, J.P. in *Speciation: Natural Processes, Genetics and Biodiversity* (ed.
918 Michalak, P.) (Nova Science Publishers, 2013).
- 919 41. Burton, R.S. & Barreto, F.S. A disproportionate role for mtDNA in Dobzhansky–Muller
920 incompatibilities? *Molecular Ecology* **21**, 4942-4957 (2012).
- 921 42. Sambatti, J.B.M., Ortiz-Barrientos, D., Baack, E.J. & Rieseberg, L.H. Ecological selection
922 maintains cytonuclear incompatibilities in hybridizing sunflowers. *Ecology Letters* **11**, 1082-
923 1091 (2008).
- 924 43. van Doorn, G.S. & Kirkpatrick, M. Turnover of sex chromosomes induced by sexual conflict.
925 *Nature* **449**, 909-912 (2007).
- 926 44. Charlesworth, D. & Charlesworth, B. Sex differences in fitness and selection for centric
927 fusions between sex-chromosomes and autosomes. *Genet. Res.* **35**, 205–214 (1980).
- 928 45. Kitano, J. et al. A role for a neo-sex chromosome in stickleback speciation. *Nature* **461**, 1079-
929 1083 (2009).
- 930 46. Parnell, N.F. & Streelman, J.T. Genetic interactions controlling sex and color establish the
931 potential for sexual conflict in Lake Malawi cichlid fishes. *Heredity* **110**, 239-246 (2013).
- 932 47. Yeaman, S. & Whitlock, M.C. The genetic architecture of adaptation under migration-
933 selection balance. *Evolution* **65**, 1897-1911 (2011).
- 934 48. Silver, L. The peculiar journey of a selfish chromosome: mouse t-haplotypes and meiotic
935 drive. *Trends in Genetics* **9**, 250-254 (1993).
- 936 49. Larracuente, A.M. & Presgraves, D.C. The selfish segregation distorter gene complex of
937 *Drosophila melanogaster*. *Genetics* **192**, 33-53 (2012).
- 938 50. Bull, J.J. *Evolution of Sex-determining Mechanisms*. (Benjamin/Cummings, London, 1983).
- 939 51. Phadnis, N. & Orr, H.A. A Single Gene Causes Both Male Sterility and Segregation Distortion
940 in *Drosophila* Hybrids. *Science* **323**, 376-379 (2009).

- 941 52. Saether, S.A. et al. Sex chromosome-linked species recognition and evolution of reproductive
942 isolation in flycatchers. *Science* **318**, 95-97 (2007).
- 943 53. Lindholm, A. & Breden, F. Sex chromosomes and sexual selection in poeciliid fishes. *American*
944 *Naturalist* **160**, S214-S224 (2002).
- 945 54. Reeve, H.K. & Pfennig, D.W. Genetic biases for showy males: Are some genetic systems
946 especially conducive to sexual selection? *Proceedings of the National Academy of Sciences of*
947 *the United States of America* **100**, 1089-1094 (2003).
- 948 55. Templeton, A.R. Mechanisms of Speciation - a Population Genetic Approach. *Annual Review*
949 *of Ecology and Systematics* **12**, 23-48 (1981).
- 950 56. Ayala, F.J. & Tracey, M.L. Genetic differentiation within and between species of *Drosophila*
951 *willistoni*-group. *Proceedings of the National Academy of Sciences of the United States of*
952 *America* **71**, 999-1003 (1974).
- 953 57. Coyne, J.A. Genetics and speciation. *Nature* **355**, 511-515 (1992).
- 954 58. Price, T.D. & Bouvier, M.M. The evolution of F-1 postzygotic incompatibilities in birds.
955 *Evolution* **56**, 2083-2089 (2002).
- 956 59. Stelkens, R.B., Young, K.A. & Seehausen, O. The Accumulation of Reproductive
957 Incompatibilities in African Cichlid Fish. *Evolution* **64**, 617-632 (2010).
- 958 60. Scopece, G., Musacchio, A., Widmer, A. & Cozzolino, S. Patterns of reproductive isolation in
959 Mediterranean deceptive orchids. *Evolution* **61**, 2623-2642 (2007).
- 960 61. Orr, H.A. & Turelli, M. The evolution of postzygotic isolation: Accumulating Dobzhansky-
961 Muller incompatibilities. *Evolution* **55**, 1085-1094 (2001).
- 962 62. Strasburg, J.L. & Rieseberg, L.H. Interpreting the estimated timing of migration events
963 between hybridizing species. *Molecular Ecology* **20**, 2353-2366 (2011).
- 964 63. Nosil, P., Funk, D.J. & Ortiz-Barrientos, D. Divergent selection and heterogeneous genomic
965 divergence. *Molecular Ecology* **18**, 375-402 (2009).
- 966 64. Strasburg, J.L. et al. What can patterns of differentiation across plant genomes tell us about
967 adaptation and speciation? *Philosophical Transactions of the Royal Society B-Biological*
968 *Sciences* **367**, 364-373 (2012).
- 969 65. Barton, N. Gene flow and speciation. *Heredity* **50**, 213-213 (1983).
- 970 66. Wu, C.I. The genic view of the process of speciation. *Journal of Evolutionary Biology* **14**, 851-
971 865 (2001).
- 972 67. Via, S. & West, J. The genetic mosaic suggests a new role for hitchhiking in ecological
973 speciation. *Molecular Ecology* **17**, 4334-4345 (2008).
- 974 68. Feder, J.L., Egan, S.P. & Nosil, P. The genomics of speciation-with-gene-flow. *Trends in*
975 *Genetics* **28**, 342-350 (2012).
- 976 69. Nadeau, N.J. et al. Genomic islands of divergence in hybridizing *Heliconius* butterflies
977 identified by large-scale targeted sequencing. *Philosophical Transactions of the Royal Society*
978 *B-Biological Sciences* **367**, 343-353 (2012).
- 979 70. Andrew, R.L. & Rieseberg, L.H. Divergence is focused on few genomic regions early in
980 speciation: incipient speciation of sunflower ecotypes *Evolution* **67**, 2468-2482 (2013).
- 981 71. Stolting, K.N. et al. Genomic scan for single nucleotide polymorphisms reveals patterns of
982 divergence and gene flow between ecologically divergent species. *Molecular Ecology* **22**, 842-
983 855 (2013).
- 984 72. Hohenlohe, P.A. et al. Population Genomics of Parallel Adaptation in Threespine Stickleback
985 using Sequenced RAD Tags. *Plos Genetics* **6** (2010).
- 986 73. Deagle, B.E. et al. Population genomics of parallel phenotypic evolution in stickleback across
987 stream-lake ecological transitions. *Proceedings of the Royal Society B-Biological Sciences* **279**,
988 1277-1286 (2012).
- 989 74. Jones, F.C. et al. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*
990 **484**, 55-61 (2012).
- 991 75. Lawnczak, M.K.N. et al. Widespread Divergence Between Incipient *Anopheles gambiae*
992 Species Revealed by Whole Genome Sequences. *Science* **330**, 512-514 (2010).

- 993 76. Renaut, S. et al. Genomic islands of divergence are not affected by geography of speciation in
994 sunflowers. *Nature Communications* (2013).
- 995 77. Keller, I. et al. Population genomic signatures of divergent adaptation, gene flow, and hybrid
996 speciation in the rapid radiation of Lake Victoria cichlid fishes. *Molecular Ecology*, doi:
997 10.1111/mec.12083 (2012).
- 998 78. Jones, F.C. et al. A Genome-wide SNP Genotyping Array Reveals Patterns of Global and
999 Repeated Species-Pair Divergence in Sticklebacks. *Current Biology* **22**, 83-90 (2012).
- 1000 79. Parchman, T.L. et al. The genomic consequences of adaptive divergence and reproductive
1001 isolation between species of manakins. *Molecular Ecology* **22**, 3304-3317 (2013).
- 1002 80. Gagnaire, P.-A., Pavey, S.A., Normandeau, E. & Bernatchez, L. The genetic architecture of
1003 reproductive isolation during speciation-with-gene-flow in lake whitefish pairs assessed by
1004 RAD-sequencing. *Evolution* **67**, 2483-2497 (2013).
- 1005 81. Michel, A.P. et al. Widespread genomic divergence during sympatric speciation. *Proceedings*
1006 *of the National Academy of Sciences of the United States of America* **107**, 9724-9729 (2010).
- 1007 82. Via, S., Conte, G., Mason-Foley, C. & Mills, K. Localizing F-ST outliers on a QTL map reveals
1008 evidence for large genomic regions of reduced gene exchange during speciation-with-gene-
1009 flow. *Molecular Ecology* **21**, 5546-5560 (2012).
- 1010 83. Roesti, M., Hendry, A.P., Salzburger, W. & Berner, D. Genome divergence during evolutionary
1011 diversification as revealed in replicate lake-stream stickleback population pairs. *Molecular*
1012 *Ecology* **21**, 2852-2862 (2012).
- 1013 84. Garrigan, D. et al. Genome sequencing reveals complex speciation in the *Drosophila simulans*
1014 clade. *Genome Research* **22**, 1499-1511 (2012).
- 1015 85. Neafsey, D.E. et al. SNP Genotyping Defines Complex Gene-Flow Boundaries Among African
1016 Malaria Vector Mosquitoes. *Science* **330**, 514-517 (2010).
- 1017 86. Nosil, P. et al. Genomic consequences of multiple speciation processes in a stick insect.
1018 *Proceedings of the Royal Society B-Biological Sciences* **279**, 5058-5065 (2012).
- 1019 87. Martin, S.H. et al. Genome-wide evidence for speciation with gene flow in *Heliconius*
1020 butterflies. *Genome Research* (2013).
- 1021 88. Noor, M.A.F. & Bennett, S.M. Islands of speciation or mirages in the desert? Examining the
1022 role of restricted recombination in maintaining species. *Heredity* **103**, 439-444 (2009).
- 1023 89. Hahn, M.W., White, B.J., Muir, C.D. & Besansky, N.J. No evidence for biased co-transmission
1024 of speciation islands in *Anopheles gambiae*. *Philosophical Transactions of the Royal Society B-*
1025 *Biological Sciences* **367**, 374-384 (2012).
- 1026 90. Sousa, V. & Hey, J. Understanding the origin of species with genome-scale data: modelling
1027 gene flow. *Nature Reviews Genetics* **14**, 404-414 (2013).
- 1028 91. Scally, A. et al. Insights into hominid evolution from the gorilla genome sequence. *Nature*
1029 **483**, 169-175 (2012).
- 1030 92. Charlesworth, B., Morgan, M.T. & Charlesworth, D. The effect of deleterious mutations on
1031 neutral molecular variation. *Genetics* **134**, 1289-1303 (1993).
- 1032 93. Spencer, C.C.A. et al. The influence of recombination on human genetic diversity. *Plos*
1033 *Genetics* **2**, 1375-1385 (2006).
- 1034 94. Nachman, M.W. & Payseur, B.A. Recombination rate variation and speciation: theoretical
1035 predictions and empirical results from rabbits and mice. *Philosophical Transactions of the*
1036 *Royal Society B-Biological Sciences* **367**, 409-421 (2012).
- 1037 95. Nei, M. *Molecular evolutionary genetics* (Columbia University Press, New York, NY, 1987).
- 1038 96. Charlesworth, B. Measures of divergence between populations and the effect of forces that
1039 reduce variability. *Molecular Biology and Evolution* **15**, 538-543 (1998).
- 1040 97. Renaut, S., Owens, G.L. & Rieseberg, L.H. Shared selective pressure and local genomic
1041 landscape lead to repeatable patterns of genomic divergence in sunflowers. *Molecular*
1042 *Ecology* (accepted).
- 1043 98. Ellegren, H. et al. The genomic landscape of species divergence in *Ficedula* flycatchers.
1044 *Nature* **491**, 756-760 (2012).

- 1045 99. Consortium, H.G. Butterfly genome reveals promiscuous exchange of mimicry adaptations
 1046 among species. *Nature* **487**, 94-98 (2012).
- 1047 100. Janousek, V. et al. Genome-wide architecture of reproductive isolation in a naturally
 1048 occurring hybrid zone between *Mus musculus musculus* and *M. m. domesticus*. *Molecular*
 1049 *Ecology* **21**, 3032-3047 (2012).
- 1050 101. Elmer, K.R. & Meyer, A. Adaptation in the age of ecological genomics: insights from
 1051 parallelism and convergence. *Trends in Ecology & Evolution* **26**, 298-306 (2011).
- 1052 102. Nosil, P., Egan, S.P. & Funk, D.J. Heterogeneous genomic differentiation between walking-
 1053 stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection. *Evolution*
 1054 **62**, 316-336 (2008).
- 1055 103. Campbell, D. & Bernatchez, L. Generic scan using AFLP markers as a means to assess the role
 1056 of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology*
 1057 *and Evolution* **21**, 945-956 (2004).
- 1058 104. Garrigan, D. et al. Genome sequencing reveals complex speciation in the *Drosophila simulans*
 1059 clade. *Genome Research* **22**, 1499-1511 (2012).
- 1060 105. Excoffier, L., Hofer, T. & Foll, M. Detecting loci under selection in a hierarchically structured
 1061 population. *Heredity* **103**, 285-298 (2009).
- 1062 106. Gagnaire, P.-A., Normandeau, E., Pavey, S.A. & Bernatchez, L. Mapping phenotypic,
 1063 expression and transmission ratio distortion QTL using RAD markers in the Lake Whitefish
 1064 (*Coregonus clupeaformis*). *Molecular Ecology* **22**, 3036-3048 (2013).
- 1065 107. Nosil, P., Parchman, T.L., Feder, J.L. & Gompert, Z. Do highly divergent loci reside in genomic
 1066 regions affecting reproductive isolation? A test using next-generation sequence data in
 1067 *Timema* stick insects. *Bmc Evolutionary Biology* **12** (2012).
- 1068 108. Lindtke, D., González-Martínez, S.C., Macaya-Sanz, D. & Lexer, C. Admixture mapping of
 1069 quantitative traits in *Populus* hybrid zones: power and limitations. *Heredity* (2013).
- 1070 109. Malek, T.B., Boughman, J.W., Dworkin, I. & Peichel, C.L. Admixture mapping of male nuptial
 1071 colour and body shape in a recently formed hybrid population of threespine stickleback.
 1072 *Molecular Ecology* **21**, 5265-5279 (2012).
- 1073 110. Schluter, D. & Conte, G.L. Genetics and ecological speciation. *Proceedings of the National*
 1074 *Academy of Sciences of the United States of America* **106**, 9955-9962 (2009).
- 1075 111. Seehausen, O. et al. Speciation through sensory drive in cichlid fish. *Nature* **455**, 620-626
 1076 (2008).
- 1077 112. Feder, J.L. et al. Allopatric genetic origins for sympatric host-plant shifts and race formation
 1078 in *Rhagoletis*. *Proceedings of the National Academy of Sciences of the United States of*
 1079 *America* **100**, 10314-10319 (2003).
- 1080 113. Nadeau, N.J. et al. Genome-wide patterns of divergence and gene flow across a butterfly
 1081 radiation. *Molecular Ecology* **22**, 814-826 (2013).
- 1082 114. Seehausen, O. Hybridization and adaptive radiation. *Trends in Ecology & Evolution* **19**, 198-
 1083 207 (2004).
- 1084 115. Smadja, C.M. & Butlin, R.K. A framework for comparing processes of speciation in the
 1085 presence of gene flow. *Molecular Ecology* **20**, 5123-5140 (2011).
- 1086 116. Bierne, N., Welch, J., Loire, E., Bonhomme, F. & David, P. The coupling hypothesis: why
 1087 genome scans may fail to map local adaptation genes. *Molecular Ecology* **20**, 2044-2072
 1088 (2011).
- 1089 117. Hermann, K. et al. Tight Genetic Linkage of Prezygotic Barrier Loci Creates a Multifunctional
 1090 Speciation Island in *Petunia*. *Current Biology* **23**, 873-877 (2013).
- 1091 118. Bank, C., Burger, R. & Hermisson, J. The Limits to Parapatric Speciation: Dobzhansky-Muller
 1092 Incompatibilities in a Continent-Island Model. *Genetics* **191**, 845-U345 (2012).
- 1093 119. Wright, K.M., Lloyd, D., Lowry, D.B., Macnair, M.R. & Willis, J.H. Indirect Evolution of Hybrid
 1094 Lethality Due to Linkage with Selected Locus in *Mimulus guttatus*. *Plos Biology* **11** (2013).
- 1095 120. Barton, N.H. Multilocus clines. *Evolution* **37**, 454-471 (1983).

- 1096 121. Barton, N.H. & de Cara, M.A.R. The evolution of strong reproductive isolation. *Evolution* **63**,
1097 1171-1190 (2009).
- 1098 122. Nosil, P., Harmon, L.J. & Seehausen, O. Ecological explanations for (incomplete) speciation.
1099 *Trends in Ecology & Evolution* **24**, 145-156 (2009).
- 1100 123. Doebeli, M. & Dieckmann, U. Speciation along environmental gradients. *Nature* **421**, 259-264
1101 (2003).
- 1102 124. Butlin, R.K. & Ritchie, M.G. Pulling together or pulling apart: hybridization in theory and
1103 practice. *Journal of Evolutionary Biology* **26**, 294-298 (2013).
- 1104 125. Servedio, M.R. & Noor, M.A.F. The role of reinforcement in speciation: Theory and data.
1105 *Annual Review of Ecology Evolution and Systematics* **34**, 339-364 (2003).
- 1106 126. Qvarnström, A. & Bailey, R.I. Speciation through evolution of sex-linked genes. *Heredity* **102**,
1107 4-15 (2009).
- 1108 127. Pryke, S.R. & Griffith, S.C. Postzygotic Genetic Incompatibility between Sympatric Color
1109 Morphs. *Evolution* **63**, 793-798 (2009).
- 1110 128. Pryke, S.R. Sex chromosome linkage of mate preference and color signal maintains
1111 assortative mating between interbreeding finch morphs. *Evolution* **64**, 1301-1310 (2010).
- 1112 129. Kirkpatrick, M. & Barton, N. Chromosome inversions, local adaptation and speciation.
1113 *Genetics* **173**, 419-434 (2006).
- 1114 130. Felsenstein, J. Skepticism Towards Santa Rosalia, or Why Are There So Few Kinds of Animals.
1115 *Evolution* **35**, 124-138 (1981).
- 1116 131. Hopkins, R. & Rausher, M.D. Pollinator-Mediated Selection on Flower Color Allele Drives
1117 Reinforcement. *Science* **335**, 1090-1092 (2012).
- 1118 132. Hopkins, R. & Rausher, M.D. Identification of two genes causing reinforcement in the Texas
1119 wildflower *Phlox drummondii*. *Nature* **469**, 411-+ (2011).
- 1120 133. Servedio, M.R., Van Doorn, G.S., Kopp, M., Frame, A.M. & Nosil, P. Magic traits in speciation:
1121 'magic' but not rare? *Trends in Ecology & Evolution* **26**, 389-397 (2011).
- 1122 134. Lowry, D.B. & Willis, J.H. A Widespread Chromosomal Inversion Polymorphism Contributes to
1123 a Major Life-History Transition, Local Adaptation, and Reproductive Isolation. *Plos Biology* **8**
1124 (2010).
- 1125 135. Joron, M. et al. Chromosomal rearrangements maintain a polymorphic supergene controlling
1126 butterfly mimicry. *Nature* **477**, 203-U102 (2011).
- 1127 136. Noor, M.A.F., Grams, K.L., Bertucci, L.A. & Reiland, J. Chromosomal inversions and the
1128 reproductive isolation of species *Proceedings of the National Academy of Sciences of the*
1129 *United States of America* **98**, 12084-12088 (2001).
- 1130 137. Shaw, K.L. & Lesnick, S.C. Genomic linkage of male song and female acoustic preference QTL
1131 underlying a rapid species radiation. *Proceedings of the National Academy of Sciences of the*
1132 *United States of America* **106**, 9737-9742 (2009).
- 1133 138. Merrill, R.M., Van Schooten, B., Scott, J.A. & Jiggins, C.D. Pervasive genetic associations
1134 between traits causing reproductive isolation in *Heliconius* butterflies. *Proceedings of the*
1135 *Royal Society B-Biological Sciences* **278**, 511-518 (2011).
- 1136 139. Saetre, G.P. & Saether, S.A. Ecology and genetics of speciation in *Ficedula* flycatchers.
1137 *Molecular Ecology* **19**, 1091-1106 (2010).
- 1138 140. Bimova, B.V. et al. Reinforcement selection acting on the European house mouse hybrid
1139 zone. *Molecular Ecology* **20**, 2403-2424 (2011).
- 1140 141. Teeter, K.C. et al. The variable genomic architecture of isolation between hybridizing species
1141 of house mice. *Evolution* **64**, 472-485 (2010).
- 1142 142. Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y. & Wang, J. SNP Calling, Genotype Calling,
1143 and Sample Allele Frequency Estimation from New-Generation Sequencing Data. *Plos One* **7**
1144 (2012).
- 1145 143. Orr, H.A. The genetics of species differences. *Trends in Ecology & Evolution* **16**, 343-350
1146 (2001).

- 1147 144. Haller, B.C., De Leon, L.F., Rolshausen, G., Gotanda, K.M. & Hendry, A.P. Magic traits:
 1148 distinguishing the important from the trivial. *Trends in Ecology & Evolution* **27**, 4-5 (2012).
- 1149 145. Fisher, R.A. The genetical theory of natural selection (Clarendon Press, Oxford, 1930).
- 1150 146. Orr, H.A. The population genetics of adaptation: The distribution of factors fixed during
 1151 adaptive evolution. *Evolution* **52**, 935-949 (1998).
- 1152 147. Orr, H.A. The genetic theory of adaptation: a brief history. *Nature Reviews Genetics* **6**, 119-
 1153 127 (2005).
- 1154 148. Rockman, M.V. The QTN program and the alleles that matter for evolution: all that's gold
 1155 does not glitter. *Evolution* **66**, 1-17 (2012).
- 1156 149. Gordo, I. & Campos, P.R.A. Evolution of clonal populations approaching a fitness peak.
 1157 *Biology Letters* **9**, 20120239 (2013).
- 1158 150. Guillaume, F. & Whitlock, M.C. Effects of migration on the genetic covariance matrix.
 1159 *Evolution* **61**, 2398-2409 (2007).
- 1160 151. Bomblies, K. & Weigel, D. Arabidopsis and relatives as models for the study of genetic and
 1161 genomic incompatibilities. *Philosophical Transactions of the Royal Society B: Biological*
 1162 *Sciences* **365**, 1815-1823 (2010).
- 1163 152. Leary, G.P. et al. Single mutation to a sex pheromone receptor provides adaptive specificity
 1164 between closely related moth species. *Proceedings of the National Academy of Sciences of*
 1165 *the United States of America* **109**, 14081-14086 (2012).
- 1166 153. Bradshaw, H.D. & Schemske, D.W. Allele substitution at a flower colour locus produces a
 1167 pollinator shift in monkey flowers. *Nature* **426**, 176-178 (2003).
- 1168 154. Feder, J.L. et al. Host fidelity is an effective premating barrier between sympatric races of the
 1169 apple maggot fly. *Proceedings of the National Academy of Sciences of the United States of*
 1170 *America* **91**, 7990-7994 (1994).
- 1171 155. Klahre, U. et al. Pollinator Choice in Petunia Depends on Two Major Genetic Loci for Floral
 1172 Scent Production. *Current Biology* **21**, 730-739 (2011).
- 1173 156. Dambroski, H.R. et al. The genetic basis for fruit odor discrimination in Rhagoletis flies and its
 1174 significance for sympatric host shifts. *Evolution* **59**, 1953-1964 (2005).
- 1175 157. Haesler, M.P. & Seehausen, O. Inheritance of female mating preference in a sympatric sibling
 1176 species pair of Lake Victoria cichlids: implications for speciation. *Proceedings of the Royal*
 1177 *Society B-Biological Sciences* **272**, 237-245 (2005).
- 1178 158. Fan, P. et al. Genetic and Neural Mechanisms that Inhibit Drosophila from Mating with Other
 1179 Species. *Cell* **154**, 89-102 (2013).
- 1180 159. Ballerini, E.S. et al. QTL mapping reveals the genetic architecture of loci affecting pre- and
 1181 post-zygotic isolating barriers in Louisiana Iris. *Bmc Plant Biology* **12** (2012).
- 1182 160. Reed, R.D. et al. optix Drives the Repeated Convergent Evolution of Butterfly Wing Pattern
 1183 Mimicry. *Science* **333**, 1137-1141 (2011).
- 1184 161. Barrett, R.D.H., Rogers, S.M. & Schluter, D. Environment specific pleiotropy facilitates
 1185 divergence at the Ectodysplasin locus in threespine stickleback. *Evolution* **63**, 2831-2837
 1186 (2009).
- 1187 162. Arnold, S.J., Burger, R., Hohenlohe, P.A., Ajie, B.C. & Jones, A.G. Understanding the Evolution
 1188 and Stability of the G-Matrix. *Evolution* **62**, 2451-2461 (2008).
- 1189 163. Lande, R. Quantitative genetic-analysis of multivariate evolution, applied to brain - body size
 1190 allometry. *Evolution* **33**, 402-416 (1979).
- 1191 164. Schluter, D. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766-
 1192 1774 (1996).
- 1193 165. Roff, D. The evolution of the G matrix: selection or drift? *Heredity* **84**, 135-142 (2000).
- 1194 166. Martin, G., Chapuis, E. & Goudet, J. Multivariate Q(st)-F-st Comparisons: A Neutrality Test for
 1195 the Evolution of the G Matrix in Structured Populations. *Genetics* **180**, 2135-2149 (2008).
- 1196 167. Hansen, T.F. & Houle, D. Measuring and comparing evolvability and constraint in multivariate
 1197 characters. *Journal of Evolutionary Biology* **21**, 1201-1219 (2008).

- 1198 168. Chenoweth, S.F., Rundle, H.D. & Blows, M.W. The Contribution of Selection and Genetic
 1199 Constraints to Phenotypic Divergence. *American Naturalist* **175**, 186-196 (2010).
- 1200 169. Roff, D.A. & Fairbairn, D.J. A test of the hypothesis that correlational selection generates
 1201 genetic correlations. *Evolution* **66**, 2953-2960 (2012).
- 1202 170. Jones, A.G., Arnold, S.J. & Bürger, R. Stability of the G-matrix in a population experiencing
 1203 pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* **57**, 1747-1760 (2003).
- 1204 171. Agrawal, A.F., Brodie, E.D. & Rieseberg, L.H. Possible consequences of genes of major effect:
 1205 transient changes in the G-matrix. *Genetica* **112**, 33-43 (2001).
- 1206 172. Jones, A.G., Arnold, S.J., Bürger, R., Hohenlohe, P.A. & Uyeda, J.C. The effects of stochastic
 1207 and episodic movement of the optimum on the evolution of the G-matrix and the response
 1208 of the trait mean to selection. *Journal of Evolutionary Biology* **25**, 2210-2231 (2012).
- 1209 173. Yang, J. et al. Genome partitioning of genetic variation for complex traits using common
 1210 SNPs. *Nature Genetics* **43**, 519-U44 (2011).
- 1211 174. Kopp, A. Metamodels and phylogenetic replication: a systematic approach to the evolution of
 1212 developmental pathways. *Evolution* **63**, 2771-2789 (2009).
- 1213 175. Eroukhmanoff, F. & Svensson, E.I. Evolution and stability of the G-matrix during the
 1214 colonization of a novel environment. *Journal of Evolutionary Biology* **24**, 1363-1373 (2011).
- 1215 176. Nolte, A.W. & Tautz, D. Understanding the onset of hybrid speciation. *Trends in Genetics* **26**,
 1216 54-58 (2010).
- 1217 177. Selz, O.M., Lucek, K., Young, K.A. & Seehausen, O. Relaxed trait covariance in interspecific
 1218 cichlid hybrids predicts morphological diversity in adaptive radiations. *Journal of Evolutionary
 1219 Biology*, 10.1111/jeb.12283 (in Press).
- 1220 178. Macaya-Sanz, D. et al. Genetic analysis of post-mating reproductive barriers in hybridizing
 1221 European *Populus* species. *Heredity* **107**, 478-486 (2011).
- 1222 179. Seehausen, O. Conditions when hybridization might predispose populations for adaptive
 1223 radiation. *Journal of Evolutionary Biology* **26**, 279-281 (2013).
- 1224 180. Bolnick, D.I. & Fitzpatrick, B.M. Sympatric speciation: Models and empirical evidence. *Annual
 1225 Review of Ecology Evolution and Systematics* **38**, 459-487 (2007).
- 1226 181. Papadopoulos, A.S.T. et al. A comparative analysis of the mechanisms underlying speciation on
 1227 Lord Howe Island. *Journal of Evolutionary Biology* **26**, 733-745 (2013).
- 1228 182. Stern, D.L. & Orgogozo, V. The loci of evolution: How predictable is genetic evolution ?
 1229 *Evolution* **62**, 2155-2177 (2008).
- 1230 183. Hoekstra, H.E. & Coyne, J.A. The locus of evolution: Evo devo and the genetics of adaptation.
 1231 *Evolution* **61**, 995-1016 (2007).
- 1232 184. Eriksson, A. & Manica, A. Effect of ancient population structure on the degree of
 1233 polymorphism shared between modern human populations and ancient hominins.
 1234 *Proceedings of the National Academy of Sciences of the United States of America* **109**, 13956-
 1235 13960 (2012).
- 1236 185. Bazin, E., Dawson, K.J. & Beaumont, M.A. Likelihood-Free Inference of Population Structure
 1237 and Local Adaptation in a Bayesian Hierarchical Model. *Genetics* **185**, 587-602 (2010).
- 1238 186. Lawson, D.J., Hellenthal, G., Myers, S. & Falush, D. Inference of Population Structure using
 1239 Dense Haplotype Data. . *PLoS Genetics* **8**, e1002453 (2012).
- 1240 187. Slate, J. Quantitative trait locus mapping in natural populations: progress, caveats and future
 1241 directions. *Molecular Ecology* **14**, 363-379 (2005).
- 1242 188. Buerkle, C.A. & Lexer, C. Admixture as the basis for genetic mapping. *Trends in Ecology &
 1243 Evolution* **23**, 686-694 (2008).
- 1244 189. Gompert, Z. & Buerkle, C.A. Bayesian estimation of genomic clines. *Molecular Ecology* **20**,
 1245 2111-2127 (2011).
- 1246 190. Trier, C.N., Hermansen, J.S., Sætre, G.P. & Bailey, R.I. Evidence for mito-nuclear and sex-
 1247 linked incompatibilities between the hybrid Italian sparrow and its parent species. *PLoS
 1248 Genetics*, 10.1371/journal.pgen.1004075 (in press).

- 1249 191. Barrett, R.D.H., Rogers, S.M. & Schluter, D. Natural selection on a major armor gene in
1250 threespine stickleback. *Science* **322**, 255-257 (2008).
- 1251 192. Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. Mapping complex disease
1252 traits with global gene expression. *Nature Reviews Genetics* **10**, 184-194 (2009).
- 1253 193. Orr, H.A. The Population-Genetics of Speciation - the Evolution of Hybrid Incompatibilities.
1254 *Genetics* **139**, 1805-1813 (1995).
- 1255 194. Vonlanthen, P. et al. Eutrophication causes speciation reversal in whitefish adaptive
1256 radiations. *Nature* **482**, 357-U1500 (2012).
- 1257 195. Rosenblum, E.B. et al. Goldilocks Meets Santa Rosalia: An Ephemeral Speciation Model
1258 Explains Patterns of Diversification Across Time Scales. *Evolutionary Biology* **39**, 255-261
1259 (2012).
- 1260 196. Butlin, R.K., Bridle, J.R. & Schluter, D. Speciation and Patterns of Diversity (Cambridge
1261 University Press, 2009).
- 1262 197. Sanderson, M.J. & Donoghue, M.J. Reconstructing shifts in diversification rates on
1263 phylogenetic trees. *Trends in Ecology & Evolution* **11**, 15-20 (1996).
- 1264 198. Seehausen, O., vanAlphen, J.J.M. & Witte, F. Cichlid fish diversity threatened by
1265 eutrophication that curbs sexual selection. *Science* **277**, 1808-1811 (1997).
- 1266 199. Taylor, E. et al. Speciation in reverse: morphological and genetic evidence of the collapse of a
1267 three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology* **15**, 343-
1268 355 (2006).
- 1269 200. Seehausen, O., Takimoto, G., Roy, D. & Jokela, J. Speciation reversal and biodiversity
1270 dynamics with hybridization in changing environments. *Molecular Ecology* **17**, 30-44 (2008).
- 1271 201. Etges, W.J., de Oliveira, C.C., Noor, M.A.F. & Ritchie, M.G. Genetics of incipient speciation in
1272 *Drosophila mojavensis*. III. Life-history divergence in allopatry and reproductive isolation.
1273 *Evolution* **64**, 3549-3569 (2010).
- 1274
- 1275
- 1276

1277 **Box 1: Genomic tools for studying speciation**

1278 Next-generation sequencing (NGS) is rapidly expanding the tool box available for studying speciation.

1279 **Patterns of genomic divergence:** Several methods can be used to investigate genome-wide divergence along
 1280 the speciation continuum. These methods include: genome scans using SNP arrays⁷⁸, RAD-seq^{72, 77} or related
 1281 genotyping-by-sequencing (GBS) methods, whole exome or transcriptome sequencing⁷⁶ and whole genome re-
 1282 sequencing¹¹³ of population samples.. Patterns in genome-wide divergence can be visualized and compared by
 1283 means of F_{ST} kernel density plots (Fig. 2) and Manhattan plots⁹⁸.

1284 **Testing for signatures of introgression:** Various approaches are available to assess if genetic variants are
 1285 shared between incipient species as a result of hybridization or due to incomplete lineage sorting⁹⁰. The ABBA-
 1286 BABA test¹⁸⁴ is particularly applicable to genome-scale datasets. It relies on the frequencies of two specific
 1287 patterns of allele sharing among a group of four species.

1288 **Identifying signatures of selection:** Genome scans can reveal genomic regions that show evidence of divergent
 1289 selection between incipient species using **F_{ST} -outlier analysis** or related approaches, which can be applied to
 1290 individual SNPs⁷⁷ or to smoothed average F_{ST} ⁷² within windows or regions of the genome. The latest methods
 1291 can account for demographic and other sources of variation (e.g. ^{105, 185}) and make improved use of high-
 1292 density marker information¹⁸⁶.

1293 **Mapping genes that are involved in reproductive isolation:** Genome scans of incipient species pairs along the
 1294 speciation continuum are a logical first step in the search for candidate RI genes^{69, 72, 74, 98}. A range of genetic
 1295 mapping tools are available to identify links between divergent genomic regions and the phenotypic traits that
 1296 contribute to RI. **Quantitative trait locus (QTL)** mapping is one powerful such method¹⁸⁷. In short, a genome-
 1297 wide set of markers is genotyped in a phenotypically variable population with known pedigree to statistically
 1298 associate markers (QTLs) with phenotypes of interest (in this case traits associated with RI). With functional
 1299 information on genes in the vicinity of a QTL, candidate RI genes can be identified.

1300 **Admixture mapping:** If pedigree data are not available, it is possible to take advantage of the phenotypic and
 1301 genetic differences that exist between hybridizing taxa and use admixture as the basis for genetic mapping of
 1302 phenotypes that contribute to RI^{109, 188} using samples from wild hybrid populations. Intrinsic and extrinsic
 1303 postzygotic barriers involve alleles that are selected against in hybrids and a variety of methods can be used to
 1304 identify such alleles in hybrid zones or in other situations where admixture occurs. Genomic cline analysis¹⁸⁹ is
 1305 one such method in which candidate RI loci with low levels of introgression relative to most of the genome can
 1306 be identified^{79, 190}.

1307 **Manipulative selection experiments:** QTL and admixture mapping have an unfortunate bias toward detecting
 1308 loci of large effect¹⁴⁸. Alternatively, alleles affecting fitness and RI can be located using manipulative selection
 1309 experiments which track allelic changes or genome-wide responses^{86, 191}. Estimates of these effects can be
 1310 ascertained by measuring selection and introgression in the wild. To date very few studies have taken this
 1311 approach and none has measured effects on reproductive isolation.

1312 **Gene expression studies:** To further investigate the significance of candidate RI-loci, expression QTL (eQTL)
 1313 analysis can be useful. It identifies genomic loci that regulate expression levels of mRNAs¹⁹². Systematically
 1314 generated eQTL information can provide insight into the mechanism underlying reproductive isolation in
 1315 regions identified through genome-wide association studies, and can help to identify networks of genes and
 1316 the role of gene interaction (including epistasis in DMIs) in reproductive isolation.

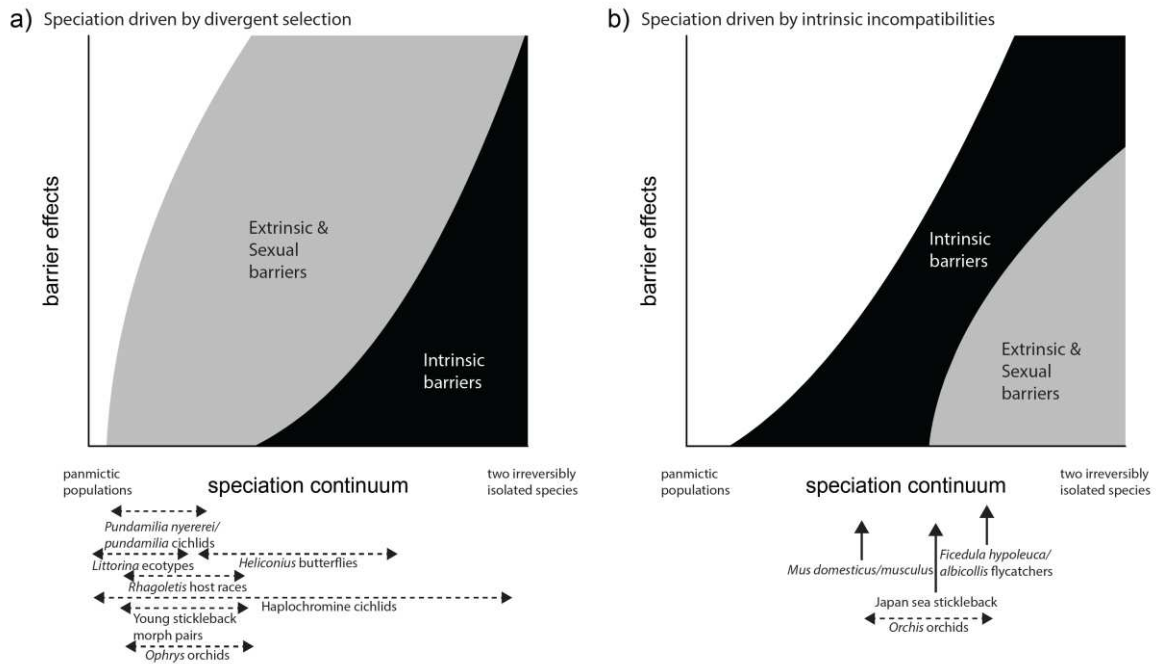
1317 **Box 2. Evolution of reproductive isolation**

1318 Reproductive isolation (RI) can usefully be divided into three forms: i) Extrinsic forms of postzygotic isolation
 1319 result from divergent ecological or sexual selection and depend on interaction with the environment or with
 1320 other individuals (e.g. reduced viability or fertility of migrants and hybrids due to ecological or behavioral
 1321 factors). ii) Intrinsic forms of postzygotic isolation are due to environment independent genetic
 1322 incompatibilities (e.g. Bateson-Dobzhansky-Muller incompatibilities). iii) Finally, prezygotic isolation includes
 1323 phenological isolation, habitat isolation, and sexual isolation due to assortative mating or fertilization.

1324 In speciation driven by divergent ecological or sexual selection, extrinsic and prezygotic forms of isolation
 1325 evolve first, and often interact, to produce reproductive isolation, and intrinsic forms of isolation will often only
 1326 evolve later in the speciation process (Panel A). In contrast, speciation driven by intrinsic barriers often results
 1327 from epistatic incompatibilities, which may (though do not necessarily¹⁹) accumulate in an accelerating
 1328 “snowball” fashion^{61, 193} as a by-product of selection or due to genetic drift (the latter only slowly). Extrinsic
 1329 postzygotic and prezygotic barriers may accumulate later, facilitating ecological coexistence between sibling
 1330 species and reinforcement of reproductive isolation (Panel B).

1331 In both panels the x-axis depicts the position of a diverging taxon pair on the “speciation continuum” (in terms
 1332 of relative time) and the y-axis represents the strength of reproductive isolation (RI) between sister taxa. Curve
 1333 shapes are hypothetical, and reflect the idea that in speciation driven by divergent selection, extrinsic
 1334 postzygotic and sexual barriers arise rapidly early in speciation. Classes of barriers within each panel are not
 1335 necessarily additive or interactive, and the emergence of RI via either of these barrier types should be viewed
 1336 as independent trajectories. Movement along the speciation continuum, from weakly isolated species to
 1337 irreversibly isolated ones, is not constant, speciation can go back and forth, or be arrested at intermittent
 1338 stages, and the average timescales for speciation via the processes contrasted here (Panels A & B) may vary.

1339 Arrows along the x-axis indicate the position(s) of model systems (studied by the authors of this paper) along
 1340 the speciation continuum. These organisms vary in the strength and types of barriers isolating incipient and
 1341 sister species. Studies of the genomics of speciation at different points on the speciation continuum are
 1342 emerging in several systems, mainly where speciation is driven by divergent selection (as indicated by the
 1343 dashed arrows showing timespans along the speciation continuum). In many cases strong reproductive
 1344 isolation may never evolve, particularly in ecological speciation (e.g. ¹²²). Incomplete reproductive isolation may
 1345 facilitate cases of “speciation reversal” (e.g. ¹⁹⁴) and “ephemeral” speciation (e.g. ¹⁹⁵).



1346

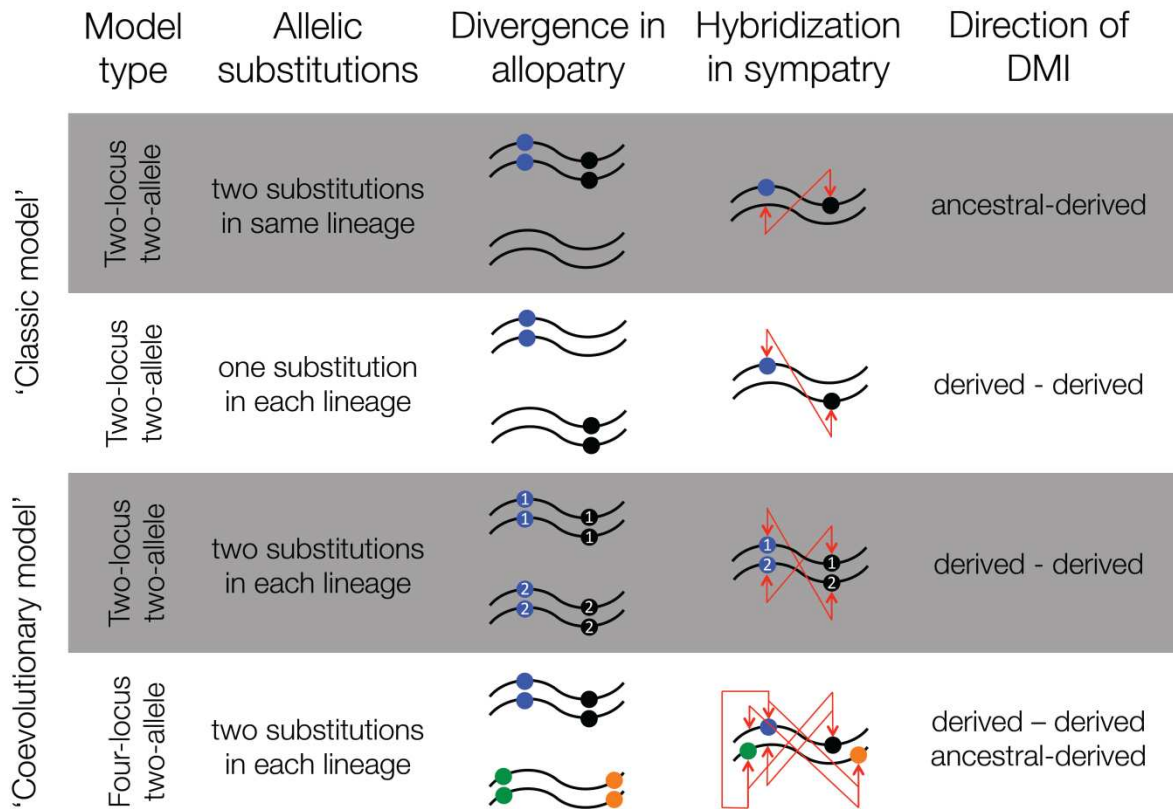
1347 **Box 3: New data for new theory: speciation genomics and patterns in biodiversity**

1348 As speciation produces the raw material for biodiversity patterns, connecting speciation processes to these
 1349 patterns in biodiversity is an important goal¹⁹⁶. We envisage that speciation genomics can make important and
 1350 unique contributions to elucidating these connections. Study of the distribution of species richness among
 1351 clades provides evidence for non-uniform diversification rates among taxa, which can arise from differences in
 1352 speciation and/or extinction rate (e.g.¹⁹⁷). Speciation rates estimated from the fossil record are far slower than
 1353 those predicted from mathematical models and observed in studies of recent diversification, and one
 1354 explanation for this discrepancy is a high frequency of “ephemeral speciation,” in which taxa that have recently
 1355 undergone speciation have high rates of extinction¹⁹⁵. This has been documented in cases of “speciation
 1356 reversal”^{194, 198, 199} which is possible when speciation does not reach “completion”^{122, 200}.

1357

1358 A better understanding of the genomic basis of speciation might help us to understand the influence of
 1359 speciation on species persistence and patterns of species diversity. For instance, ecological speciation readily
 1360 and rapidly produces divergent, partially isolated ecotypes and species that may immediately be able to coexist
 1361 without competitive exclusion. Ecological speciation might thereby contribute disproportionately to the
 1362 buildup of biodiversity compared to non-ecological mechanisms¹⁹⁶. However, isolation between young
 1363 ecologically differentiated species is often extrinsically based and contingent upon the persistence of divergent
 1364 selection (see Box 2). The species that arise most rapidly may therefore be those species that are most
 1365 vulnerable to extinction early in their histories²⁰⁰. In contrast, speciation via intrinsic mechanisms may produce
 1366 species that are less prone to ephemerality because speciation reversal may be less likely. However, speciation
 1367 rates might be slower in these lineages than in lineages where ecological speciation is common, and ecological
 1368 differences must evolve after speciation in order for closely related taxa to coexist. Progress in connecting
 1369 speciation to broader-scale patterns of species richness will require attention to how speciation mechanisms,
 1370 and their genomic basis, influence rates of speciation and the persistence and coexistence of young species. If
 1371 mechanisms of speciation leave distinctive genomic signatures, correlation between genomic patterns and
 1372 disparity in species richness among clades could be tested quantitatively using comparative phylogenetic
 1373 approaches.

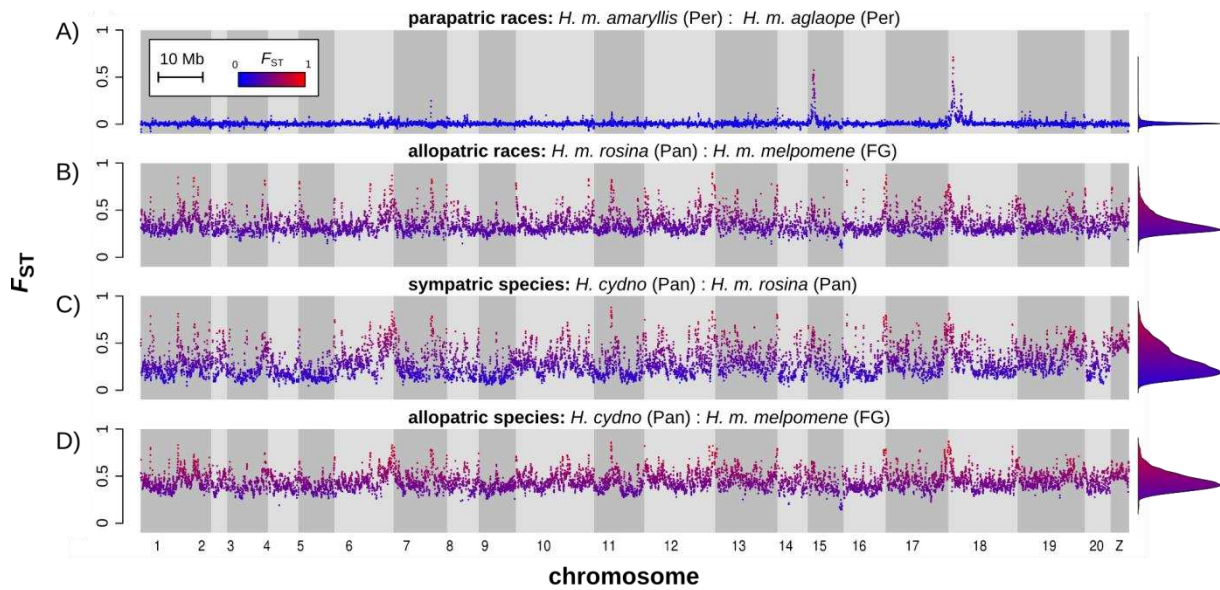
1374



1375

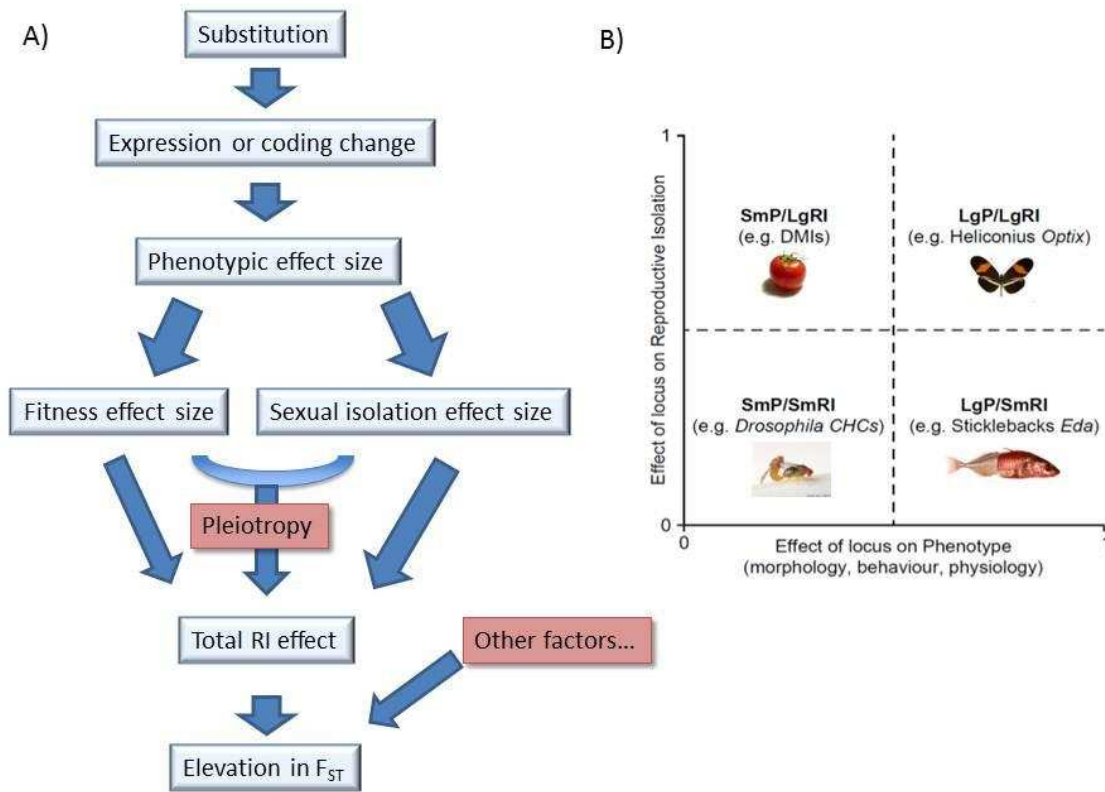
1376

1377 **Fig. 1. ‘Classic’ and coevolutionary models of hybrid incompatibility in a genomic conflict scenario.** In the
 1378 ‘classic model’, Bateson-Dobzhansky-Muller incompatibilities (DMIs) are envisioned as two-locus, two-allele
 1379 interactions, in which incompatibilities arise between an ancestral allele and an allele derived in one lineage
 1380 (1st row) or between alleles derived in two separate lineages (2nd row); a special case of the latter model can
 1381 refer to maternal-effect selfish loci in which maternal "poison" and zygotic "antidote" are due to
 1382 developmental expression divergence of the same locus. In the coevolutionary models, DMIs are continually
 1383 fixed at the same loci (3rd row) or at different loci (4th row). In all examples with two substitutions in a lineage,
 1384 the selfish locus (left) drives the evolution of the restorer locus (right). Red arrows indicate negative epistatic
 1385 interactions between complimentary loci. In all models, the ancestral state is wild-type except for row three. In
 1386 this row, the ancestral state is a coevolving selfish element-restorer system. Insight into the role of genomic
 1387 conflict in speciation reveals the potential for further development of models of hybrid incompatibility. Models
 1388 that incorporate the possibility for increased lag-load due to ongoing coevolution predict successively more
 1389 severe incompatibilities. Additional theoretical work is needed to investigate such coevolutionary models.



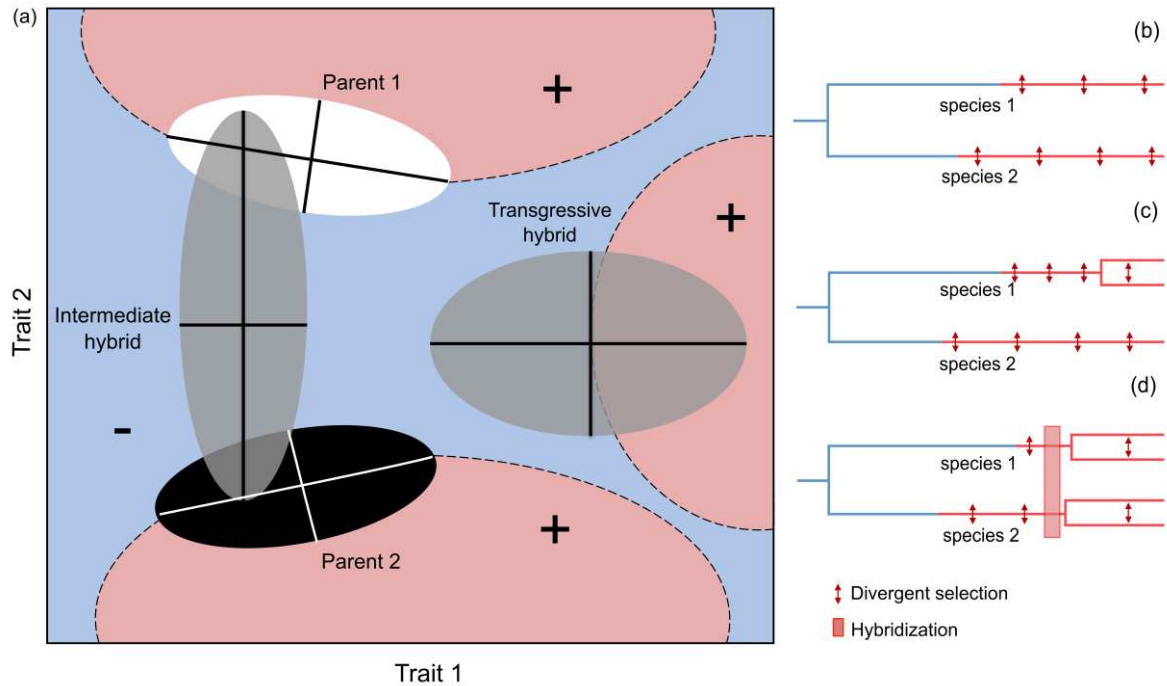
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399

Fig. 2. Genomic patterns of divergence along the speciation continuum in *Heliconius* butterflies. The top panel shows the patterns of differentiation between hybridizing parapatric races (A) and sympatric species (C) and between geographically isolated races (B) and species (D) along the genome. Divergence is highly heterogeneous even between allopatric populations of the same species (B). The shape of the frequency distribution of locus-specific F_{ST} values (bottom panel) clearly differs between the different stages in the continuum and between geographic scenarios with, for example, the greater variance in (C) consistent with gene flow between species in sympatry. However, the challenge is to distinguish between speciation with (A, C) versus without (B, D) gene flow.



1400

1401 **Fig. 3. Effect sizes of substitutions on phenotype and on reproductive isolation.** (A) Effects of variation
 1402 at different levels, and connections between those levels. The size of effect can vary at each step from zero or
 1403 quite small to very large. A substitution can alter gene expression or protein coding, which in turn has some
 1404 effect on a phenotype. This phenotype can have effects of varying size on environment-dependent fitness (and
 1405 hence possibly extrinsic postzygotic isolation), environment-independent fitness (hence possibly intrinsic
 1406 postzygotic isolation) and on prezygotic isolation. Alternatively a phenotype may pleiotropically affect both
 1407 fitness and prezygotic isolation. All these effects combine to generate total RI, which will likely elevate F_{ST} ,
 1408 although other factors can alter F_{ST} as well. (B) The lack of correlation between the effect of a locus on
 1409 phenotype (P) and on reproductive isolation (RI). An example for each of the four relationships is shown to
 1410 illustrate that phenotypic effect size does not necessarily predict RI effect size: loci with small effect on
 1411 phenotype and large effect on reproductive isolation (SmP/LgRI: DMIs in *Solanum*²⁷); loci with large effect on
 1412 phenotype and large effect on reproductive isolation (LgP/LgRI: *Optix* in *Heliconius*¹⁶⁰); loci with small effect on
 1413 phenotype and small effect on reproductive isolation (SmP/SmRI: CHCs in *Drosophila*²⁰¹); loci with large effect
 1414 on phenotype and small effect on reproductive isolation (LgP/SmRI: *Eda* in stickleback¹⁹¹). The relationships
 1415 between phenotypic and RI effect size and F_{ST} are largely unknown at present.



1416

1417

1418 **Fig. 4. Influence of genetic constraints on speciation.** (A) With the help of NGS, it is now feasible to infer
 1419 relatedness of individuals in any given natural population and thus to estimate a **G**-matrix without the use of
 1420 pedigree-data¹⁷³. The **G**-matrix (represented here as an ellipse in the space of two quantitative traits) can bias
 1421 evolution in certain directions, and depending on the adaptive landscape (represented by regions of higher (+;
 1422 red) and lower (-, blue) fitness than the parental populations (white, black)), might constrain adaptive
 1423 divergence and speciation. Hybridization events may facilitate speciation by aligning the **G**-matrix in the
 1424 direction of divergence between parental species (intermediate hybrid), or by giving rise to novel phenotypes
 1425 (transgressive hybrid) in new regions of positive fitness that cannot be reached through gradual evolution in
 1426 either of the parental species.

1427

1428 (B-D) The influence of genetic constraints on speciation can be tested at the phylogenetic level. (B) Constraints
 1429 may persist over evolutionary time as a result of the inability of divergent selection to change genetic
 1430 architecture, preventing speciation from happening. (C) Alternatively, other forms of selection (e.g.
 1431 correlational selection) can alter the structure and orientation of the **G**-matrix and potentially facilitate
 1432 divergence and speciation over moderate time scales. (D) Hybridization and gene flow can dramatically alter **G**
 1433 in just a few generations, fueling adaptive divergence and resulting in sudden bursts of speciation. Note that
 1434 hybridization between sister species is shown here for illustration, but hybridization that facilitates divergence
 1435 may occur more widely among related taxa.

1436

1437 **Biographies**

1438 Ole Seehausen studied speciation and hybridization since his PhD at the University of Leiden in the 1990s.
 1439 Adaptive radiations receive his particular attention, such as the cichlid fishes of Lake Victoria and, more
 1440 recently, the whitefish of prealpine European lakes, stickleback and trout. He is a professor in the Institute of
 1441 Ecology & Evolution of the University of Bern and head of a research department at EAWAG, the Swiss Federal
 1442 Institute of Aquatic Science and Technology. His lab combines ecological and behavioral research with genetics
 1443 and genomics to investigate processes and mechanisms implicated in adaptation, speciation, species
 1444 coexistence and extinction.

1445 Roger Butlin has studied speciation since his postdoctoral work with Godfrey Hewitt in the 1980s. He is
 1446 interested in the processes generating reproductive isolation and its genetic basis. Reinforcement has been a
 1447 particular focus of study. Current projects are examining the role of chemosensory genes in aphid host race
 1448 formation and the genetic basis of parallel local adaptation and speciation in periwinkles. He is a professor of
 1449 evolutionary biology at the University of Sheffield in the UK and currently holds the 2013 Tage Erlander guest
 1450 professorship at the University of Gothenburg in Sweden.

1451 Irene Keller is a bioinformatician at the Department of Clinical Research of the University of Bern (Switzerland).
 1452 She received her PhD from the University of Bern and worked as a postdoctoral fellow with Richard Nichols at
 1453 Queen Mary University of London and with Jukka Jokela and Ole Seehausen at Eawag and University of Bern
 1454 (Switzerland). Her interests focus on the application of molecular and bioinformatics tools to understand the
 1455 genetic basis of adaptation, speciation and human disease.

1456 Catherine E. Wagner is an evolutionary biologist with interests in speciation and the origins of diversity, and the
 1457 relationships between diversity-generating processes and macroevolutionary patterns. Her research uses
 1458 population genetic, phylogenetic, and comparative methods to study diversification. She is currently a
 1459 postdoctoral researcher at Eawag, the Swiss Federal Institute of Aquatic Science and Technology and the
 1460 University of Bern, Switzerland, where her work focuses primarily on African cichlid fishes. She earned her
 1461 Ph.D. in ecology and evolutionary biology from Cornell University in 2011.

1462 Janette Boughman and her lab study the selective forces causing speciation in threespine sticklebacks, with
 1463 particular focus on sexual selection and its interaction with natural selection to generate reproductive isolation.
 1464 She uses lab and field behavioral experiments to understand the subtle yet powerful action of these forces on
 1465 phenotypic and genetic evolution and how this transmits to the genome. She has studied both the
 1466 accumulation of reproductive isolation and its loss through reverse speciation. Recent work investigates fitness
 1467 landscapes at both the phenotypic and genetic level and their role in diversification. She is Associate Professor
 1468 at Michigan State University.

1469 Paul A. Hohenlohe is an Assistant Professor in the Department of Biological Sciences and the Institute for
 1470 Bioinformatics and Evolutionary Studies at the University of Idaho. He earned his Ph.D. in zoology at the
 1471 University of Washington in 2000 and subsequently worked as a conservation biologist and postdoctoral
 1472 researcher. His research focus is on evolutionary genetics and genomics, including RAD sequencing and other
 1473 tools for population genomics and conservation in non-model organisms, experimental evolution, and
 1474 evolutionary quantitative genetics theory.

1475 Catherine Peichel earned her Ph.D. in the area of developmental genetics at Princeton University in 1998.
 1476 During this time, she became intrigued by the genetic basis of phenotypic differences between species. Thus,
 1477 during her postdoctoral fellowship with David Kingsley at Stanford University, she helped to develop the
 1478 threespine stickleback as a genetic and genomic model system. She has led a research laboratory at the Fred
 1479 Hutchinson Cancer Research Center in Seattle, Washington since 2003. Her lab takes a number of approaches
 1480 to investigate the genetic and genomic changes that underlie adaptation and speciation in sticklebacks.

1481 Glenn-Peter Sætre is a professor in evolutionary biology at the University of Oslo, Norway. He obtained his
1482 doctorate also in Oslo and worked several years at Uppsala University, Sweden as a post doc and assistant
1483 professor before returning to Oslo in 2003 as full professor. He studies speciation, hybridization and adaptive
1484 evolution, mainly in birds, combining genomic analysis and population genetics with behavioural and ecological
1485 studies. His research lab is currently mainly focusing on the genomics of hybrid speciation.

1486

1487

1488 **Highlighted references**

- 1489 Andrew, R.L. & Rieseberg, L.H. Divergence is focused on few genomic regions early in speciation: incipient speciation of
 1490 sunflower ecotypes *Evolution* **67**, 2468-2482 (2013).
 1491 *Applying genomic analysis at different points on the speciation continuum is important for understanding how reproductive*
 1492 *isolation develops. This study shows how differentiation can be focused on a small proportion of the genome early in*
 1493 *speciation.*
 1494
 1495 Barrett, R. D. H., Rogers, S. M. & Schluter, D. Natural selection on a major armor gene in threespine stickleback. *Science*
 1496 **322**, 255-257 (2008).
 1497 *This study showed unexpected complexity in the response to selection for alleles of the Ectodysplasin locus, probably*
 1498 *because of pleiotropic effects on other fitness-related traits. Such pleiotropic effects may be widespread and have major*
 1499 *impacts on the progress of adaptation and speciation.*
 1500
 1501 Barton, N. H. & de Cara, M. A. R. The evolution of strong reproductive isolation. *Evolution* **63**, 1171-1190 (2009).
 1502 *An important, theoretical paper considering the conditions under which selection can overcome recombination to bring*
 1503 *together multiple reproductive barriers and so generate strong reproductive isolation.*
 1504
 1505 Bimova, B. V. et al. Reinforcement selection acting on the European house mouse hybrid zone. *Mol. Ecol.* **20**, 2403-2424
 1506 (2011).
 1507 *Using powerful analyses of the exceptionally well-studied mouse hybrid zone, the authors provide clear evidence for the*
 1508 *operation of reinforcement and also for the limits on its effectiveness in reducing gene flow.*
 1509
 1510 Coyne, J. & Orr, H. *Speciation.* (Sinauer Associates, Sunderland, M.A., 2004).
 1511 *Coyne and Orr's comprehensive book is a must-read for every student of speciation. It provides excellent background*
 1512 *information and review of all facets of research on the speciation process.*
 1513
 1514 Coyne, J. A. & Orr, H. A. "Patterns of speciation in *Drosophila*" revisited. *Evolution* **51**, 295-303 (1997).
 1515 *A key comparative study on the rate of evolution of reproductive isolation that was a model for similar studies in other taxa*
 1516
 1517 Ellegren, H. et al. The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* **491**, 756-760 (2012).
 1518 *A truly genome-wide view of differentiation in a system with outstanding ecological and behavioural information available.*
 1519 *Strikingly heterogeneous levels of divergence were observed, including major divergence peaks at telomeres, that are, as*
 1520 *yet, unexplained.*
 1521
 1522 Felsenstein, J. Skepticism Towards Santa Rosalia, or Why Are There So Few Kinds of Animals. *Evolution* **35**, 124-138 (1981).
 1523 *A key paper introducing the antagonism between recombination and the build-up of linkage disequilibrium which lies at the*
 1524 *heart of the speciation process.*
 1525
 1526 Feder, J. L., Egan, S. P. & Nosil, P. The genomics of speciation-with-gene-flow. *Trends in Genetics* **28**, 342-350 (2012).
 1527 *In this review, the authors bring together empirical studies with theory on the impact of divergent selection on gene flow*
 1528 *elsewhere in the genome to examine how reproductive isolation might spread through the genome as speciation proceeds.*
 1529
 1530 Heliconius Genome Consortium. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species.
 1531 *Nature* **487**, 94-98 (2012).
 1532 *An excellent illustration of how annotated reference genomes together with population-level NGS data can address*
 1533 *fundamental questions about the genetic basis for speciation. The Heliconius Genome Consortium provide evidence that*
 1534 *genes for mimicry in color pattern have been exchanged via hybridization among Heliconius species, thus implicating*
 1535 *adaptive introgression as a crucial component of speciation in this system.*
 1536
 1537 Jones, F. C. et al. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55-61 (2012).
 1538 *An excellent example of how genomic data can inform researchers about the genetic basis for repeated adaptation to*
 1539 *similar environments. They show that stickleback fish populations throughout the northern hemisphere have repeatedly*
 1540 *utilized the same genetic elements to adapt to freshwater environments.*
 1541
 1542 Kirkpatrick, M. & Barton, N. Chromosome inversions, local adaptation and speciation. *Genetics* **173**, 419-434 (2006).
 1543 *This paper gives a thorough theoretical background and new insights into the role of chromosome inversions in adaptation*
 1544 *and speciation.*
 1545
 1546 Lindtke, D., González-Martínez, S. C., Macaya-Sanz, D. & Lexer, C. Admixture mapping of quantitative traits in *Populus*
 1547 *hybrid zones: power and limitations.* *Heredity*, doi:10.1038/hdy.2013.69 (2013).
 1548 *One of the first studies to use admixture mapping in a natural hybrid population to examine the genetic basis of traits*
 1549 *contributing to reproductive isolation. Application of this powerful approach is likely to make important contributions to*
 1550 *speciation research.*

- 1551
 1552 Nolte, V., Pandey, R.V., Kofler, R. & Schlötterer, C. Genome-wide patterns of natural variation reveal strong selective
 1553 sweeps and ongoing genomic conflict in *Drosophila mauritiana*. *Genome Research* **23**, 99-110 (2013).
 1554 *Genome-sequence data here show how widespread strong selection due to genomic conflict can be, suggesting that it may*
 1555 *be a potent source of incompatibilities between previously isolated populations.*
 1556
 1557 Nosil, P. Ecological Speciation (Oxford University Press, Oxford and New York, 2012).
 1558 *An in-depth treatment of speciation by divergent natural selection.*
 1559
 1560 Orr, H. A. & Turelli, M. The evolution of postzygotic isolation: Accumulating Dobzhansky-Muller incompatibilities. *Evolution*
 1561 55, 1085-1094, (2001).
 1562 *Orr and Turelli's paper provides excellent background on the evolution of reproductive isolation via postzygotic intrinsic*
 1563 *barriers. It considers the dynamics of accumulation of Dobzhansky-Muller incompatibilities in diverging lineages.*
 1564
 1565 Presgraves, D. C. The molecular evolutionary basis of species formation. *Nature Reviews Genetics* **11**, 175-180, (2010).
 1566 *Presgraves' review summarizes empirical evidence on the genetic basis of hybrid dysfunction, focusing on work in genetic*
 1567 *model organisms where key genes have been identified.*
 1568
 1569 Schluter, D. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766-1774 (1996).
 1570 *A key conceptual paper showing how the structure of genetic variances and covariances among quantitative traits can*
 1571 *influence the direction of evolution and so the progress of adaptive radiation.*
 1572
 1573 Seehausen, O. et al. Speciation through sensory drive in cichlid fish. *Nature* **455**, 620-626 (2008).
 1574 *This study shows how divergent female preferences, initially with an ecological basis, can impose selection on male signal*
 1575 *traits and generate reproductive isolation in the face of gene flow, given the right balance between intensity of selection and*
 1576 *distribution of habitats.*
 1577
 1578 Wu, C. I. The genic view of the process of speciation. *Journal of Evolutionary Biology* **14**, 851-865 (2001).
 1579 *Wu's thought-provoking review discusses the process of speciation from a genetic perspective, by highlighting the difference*
 1580 *between thinking of the evolution of reproductive isolation as a whole-genome process, versus understanding the influence*
 1581 *of specific loci on reproductive isolation/gene exchange. Wu's point that genes, and not whole genomes, are the unit of*
 1582 *species differentiation is a seminal perspective, critical to much of the current work in speciation genetics.*

1583 **Online key points:**

- 1584
- 1585
- 1586
- 1587
- 1588
- 1589
- 1590
- 1591
- 1592
- 1593
- 1594
- 1595
- 1596
- 1597
- 1598
- 1599
- 1600
- 1601
- 1602
- 1603
- 1604
- 1605
- 1606
- 1607
- 1608
- Speciation is a central process in evolution that is fundamentally about the origin of reproductive isolation. The latest generation of genomic approaches provides remarkable opportunities to describe speciation and learn about speciation mechanisms.
 - Genome scans, now truly genome-wide and at base-pair resolution, reveal substantial genomic divergence among incipient species even in the face of gene flow, with extensive genomic heterogeneity in the extent of differentiation, especially at early stages of speciation, both in sympatry and in allopatry.
 - The sources of this heterogeneity remain incompletely understood. Combining genome scans with sophisticated population genetic modeling, QTL, and admixture analysis has the potential to isolate the influence of selection from demographic, historical and structural effects and to link these sources of genomic divergence to phenotypes and to reproductive isolation.
 - Available empirical data suggest that differentiation between parapatric populations can be restricted to few genomic islands, whereas incipient species that coexist in sympatry show differentiation widely distributed across the genome. This may suggest that genomically widespread selection is required to permit the maintenance and perhaps the buildup of genetic differentiation in sympatry.
 - Recent genomic studies reveal that the genetic basis of reproductive isolation is often complex. The effects of pleiotropy, genetic correlations, and patterns of recombination need to be considered, alongside effects of ecological and sexual selection as well as genomic conflict.
 - A surprising recent discovery has been the re-use of ancient gene variants in speciation, acquired from standing genetic variation or by introgressive hybridization.
 - We propose a roadmap for the development of speciation genomics towards answering classical as well as emerging questions in speciation research.

1609