Understanding admixture: haplodiploidy to the rescue

Pierre Nouhaud[1], Alexandre Blanckaert[2], Claudia Bank[2] & Jonna Kulmuni[1]

[1] Organismal & Evolutionary Biology Research Programme, University of Helsinki, Finland

[2] Instituto Gulbenkian de Ciência, Oeiras, Portugal

Corresponding author: pierr3.nouhaud@gmail.com (PN), https://www.helsinki.fi/en/researchgroups/evolution-sociality-behaviour

ORCID IDs: 0000-0003-4923-0993 (PN), 0000-0003-2735-7491 (AB), 0000-0003-4730-758X (CB), 0000-0002-8852-0793 (JK)

Twitter handles: @prrnhd (PN), @claudia_bank (CB)

Keywords

Hybridization; adaptive introgression; Bateson-Dobzhansky-Muller incompatibility; haplodiploidy; population genomics

Abstract

Hybridization has broad evolutionary consequences from fueling or counteracting speciation to facilitating adaptation to novel environments. Hybridization and subsequent introgression appear widespread along the tree of life. However, our understanding of how distinct evolutionary forces shape admixed genomes and the fate of introgressed genetic variants remains scarce. Most admixture research in animals has focused on diploid organisms. We propose that haplodiploid organisms can help resolve open questions about the genomic consequences of hybridization in natural populations. The ploidy difference between haploid males and diploid females, the availability of genome-wide male haplotypes, and ongoing cases of admixture make haplodiploid organisms promising models to improve our knowledge with regards to the evolution of hybrid genomes.

Gene flow between divergent lineages is widespread, but its consequences are still poorly understood

Phylogenetics and population genomics have revealed that **hybridization** (see Glossary) is widespread across the tree of life [1]. Despite its pervasiveness, there remain many open questions (Box 1). Answering these questions appears crucial as hybridization can have broad evolutionary consequences. For example, it can promote (e.g., **reinforcement** [2]) or reverse speciation [3] after secondary contact between diverging lineages. In some cases, hybridization could lead to the extinction of one or even both parental species [4]. Hybridization can also lead to the formation of new species through the evolution of reproductive isolation between the hybrid population and its parental sources (hybrid speciation [3,5]).

Through the **introgression** of novel DNA tracts into a population, hybridization enables the transfer of genetic variation across divergent lineages. This can allow the receiving lineage to adapt to novel environments (**adaptive introgression** [6,7], Box 1, Open Question (OQ) 1-2), and in some cases fuel **adaptive radiations** [8]. However, the relative contributions of adaptive and deleterious variants transferred by introgression remains unclear. In addition to having pervasive evolutionary effects, as detailed above, hybridization is a powerful asset to studying speciation (i.e., [9–11]). Specifically, genes and mechanisms underlying reproductive isolation between parental lineages can be identified through genome-wide polymorphism surveys of parents and hybrids [e.g., 12] (Box 1, OQ7). Studying the genomic footprints left by past hybridization events has improved our understanding of key biological phenomena such as speciation and adaptation, as well as early modern human history [13]. Yet, new questions arise e.g., about the recombination

54 landscapes in hybrid genomes and their influence on genome-wide patterns of
55 **admixture** (e.g., [14–16], Box 1, OQ5-6).
56 Moreover, admixture research could help us face contemporary challenges. First, the
57 use of first-generation hybrids in plant and animal breeding has transformed the
58 agricultural industry. Indeed, many such F1 hybrids exhibit transgressive phenotypes
59 leading to higher crop yield thanks to **heterosis** (Box 1, OQ7). Deciphering hybrid
60 genomics coupled with genome editing techniques could provide a more efficient
61 approach to hybrid breeding in the future. Second, anthropogenic activity leads to
62 new hybridization events by bringing previously isolated lineages into contact. For
63 example, numerous alien species have been introduced into new areas, either
64 purposefully or accidentally [17], and climate change is resulting in range shifts in
65 species distributions [18]. Thus, quantifying the outcome of contemporary admixture
66 events and their repeatability is of growing interest in biodiversity management (Box
67 1, OQ8-9). On one hand hybridization could rescue native populations with limited
68 genetic variation through adaptive introgression [e.g., 19], but on the other hand it
69 could lead to the extinction of local populations [20,21] (Box 1, OQ10).
70 In order to better comprehend the various outcomes of hybridization, we urgently
71 need to know how the interplay between evolutionary forces shapes the fate of
72 introgressed variants and the evolution of hybrid genomes.
73

74 ## Haplodiploidy to the rescue

75 Most research on admixture has so far focused on diploid organisms [e.g., 6,7].
76 However, in this opinion piece we argue that the study of hybridizing, haplodiploid
77 (HD) organisms could shed additional light on evolutionary processes governing
78 admixture events in all organisms. **Haplodiploidy** is widespread (approximately 15%

of arthropod species [22], with members of e.g., Hemiptera, Hymenoptera, Thysanoptera and Trombidiformes, Box 2) and **arrhenotoky**, one type of haplodiploidy, has originated about 17 times in animals [23]. Despite this common occurrence, HDs have been understudied in speciation and hybridization research (but see [24]), partly owing to a lack of available genomic resources. Many HD arthropod genomes have now been sequenced [e.g., 25], and more high-quality assemblies will be released in the future thanks to collaborative sequencing efforts (e.g., i5K initiative [26], Global Ant Genomic Alliance [27]). In addition, the build-up of reproductive barriers could be more rapid in HDs compared to diploid organisms [28,29], making HDs a useful study case. Time is therefore ripe for haplodiploidy to contribute to our understanding of admixture. Below we discuss several advantages of haplodiploidy for admixture research.

Inferring selection in hybrids

In HDs one sex is haploid and the other diploid. These differences in ploidy levels provide increased power to infer selection genome-wide in haplodiploid compared to diploid hybrids. Coupled with genome-wide single-nucleotide polymorphism (SNP) data, this feature could help identify selection against alleles involved in **Bateson-Dobzhansky-Muller incompatibilities** (BDMIs), and/or adaptive introgression events (Box 1, OQ1-2). Moreover, measuring allele frequency change during development (within a single generation) could help validate candidate barrier loci identified through F_{ST}- or LD-based approaches.

Identification of putative BDMI loci

BDMIs have been extensively studied as barriers to gene flow between diverging lineages (for a review, see [30]), but their impact on the evolution of hybrid genomes and the outcome of admixture events is less known (but see e.g., [12,31,32]). BDMIs often underlie hybrid sterility and/or inviability (Box 1, OQ1-2 on the interplay between evolutionary forces in admixed genomes) and are challenging to detect in diploids if recessive. Indeed, both theory [33] and empirical work [e.g., 9] suggest most BDMI alleles should be recessive. This recessivity of BDMIs could be the underlying cause of **Haldane's rule**, which applies genome-wide in HDs. As a consequence, diploids are not an ideal system to map BDMIs, since multiple backcrosses are necessary to create homozygous BDMIs that reveal the effect of recessive alleles in hybrids. These challenges are overcome in HDs, both in natural [34] and lab [35] populations. In several hybridizing HDs, hybrid females suffer less from hybridization than hybrid males [34–36]. This pattern is consistent with recessive BDMI alleles being selected against in haploid, hybrid males by causing hybrid breakdown during development. In hybrid diploid females, recessive BDMIs can escape negative selection by being masked in heterozygotes. Hence, the signature of recessive BDMIs in HDs should be sex-specific, with BDMI alleles displaying larger frequency decrease in males during development compared to females (Fig. 1). Thus, comparing allele frequency changes in both sexes during development within a generation along the genome (see, e.g., [37] for an example of such approach using SNP genotyping) could help pinpoint candidate BDMI alleles in natural and lab populations of HDs [34].

Additionally, BDMIs give rise to non-random associations between physically distant loci, since some combinations of alleles will be lacking in hybrids due to selection against unfit combinations at these loci. Candidate BDMIs can then be characterized

126 by locus pairs which positively covary by ancestry, as measured in ancestry
127 disequilibrium (AD) scans [38]. While such AD scans represent a promising tool to
128 identify BDMIs (but see [38] for limitations), it is unclear whether male haploidy will
129 lead to differences in AD patterns in HDs, compared to diploid organisms.
130
131 ## Haploid male haplotypes are an asset for selection scans and genome assembly
132 HDs have another advantage for the study of selection: individual (haploid) male
133 sequencing directly provides complete haplotype information without requiring
134 computational **phasing** ([39,40], but note this does not apply to some HD **paternal**
135 **genome elimination** systems (PGE) [41]). Haplotype-based statistics are the most
136 powerful to detect recent selection events [42]. In haploid males, long haplotypes
137 can be obtained cost-effectively and in large numbers through short-read
138 sequencing. This represents an appealing option to reveal young selective sweeps
139 and signatures of adaptive introgression (Box 1, OQ1 and OQ2 on the interplay
140 between evolutionary forces in admixed genomes, OQ3 on the tempo of adaptive
141 introgression, OQ7 on the genomic basis of fitness differences between hybrids and
142 parents).
143 Owing to male haploidy, HDs also hold promises for *de novo* genome assembly.
144 High quality reference genomes are an asset for population genomic analyses.
145 Building a reference genome usually requires pooling several individuals after
146 several generations of inbreeding, to respectively get enough genetic material and
147 decrease heterozygosity (but see [43]). In HDs, this usually tedious inbreeding task
148 can be avoided using haploid males. Moreover, a low-input protocol recently
149 released by PacBio now allows for long read sequencing libraries from a single, wild-

caught individual [44]. Such protocol applied to haploid male individuals would ensure high assembly contiguity and completeness.

Validation of loci under selection using a combined approach from genotype to fitness

Since dominance is by definition absent in haploids, natural selection acting on fully recessive BDMIs can be studied in controlled conditions by comparing affected (haploid males) and potentially unaffected (diploid females) individuals. The strategy presented above (see *Identification of putative BDMI loci*) could then allow for the validation of candidate barrier loci identified through F_{ST} scans [37], AD scans (which could give spurious signals even in the absence of selection [38]), and/or candidate adaptive introgression events.

This simple strategy does not require genetic mapping *per se*. However, this approach in HDs (or sex chromosomes, see below) is sensitive to any form of selection that is sex-specific [45]. Thus, it will require further testing to distinguish putative recessive BDMIs from sex-specific or sexually antagonistic selection. Still, this experimental framework could provide evidence for positive selection acting on introgressed alleles ([34,35], Box 1, OQ4 on context- and/or environment-dependence of selection on introgressed variation), and help identify the genomic basis of heterosis (Box 1, OQ7 on the genomic basis of fitness differences between hybrids and parents). Finally, quantifying allele frequency changes across life stages enables the inference of selection acting on introgressed variation, and as such provides a link between genotype and fitness [46].

What else can we learn about admixture from HDs?

Testing whether genetic load matters

Differences in the efficiency of **purifying selection** in both parental lineages could have a profound impact on the admixture landscape in hybrids. For example, research on archaic hominin admixture has revealed heterogeneity of Neanderthal introgression in modern human populations. Studies have documented a reduction of Neanderthal ancestry in conserved genic regions [47,48] and an apparent decrease of Neanderthal ancestry through time [49]. It has been suggested that a lower **effective population size** in archaic humans could have led to less efficient purifying selection and accumulation of deleterious alleles (i.e., genetic load, [50,51]). However, after admixture, once in larger populations, these alleles would have been purged more effectively (along with linked neutral variation), resulting in genomic regions devoid of introgressed variation. A recent study showed that the decline of Neanderthal ancestry through time in modern human genomes could be an artifact [52]. However, this increased efficiency in purging deleterious variation could theoretically impact the evolution of hybrid genomes if parental lineages *(i)* accumulate deleterious alleles and *(ii)* show differences in effective population sizes. In HDs, all else being equal (and assuming equal sex ratios), haplodiploids have ¾ the effective population size of diploid autosomes. Moreover, in social Hymenoptera, small effective population sizes can lead to elevated genetic load [53] despite a more efficient purging of deleterious alleles due to male haploidy. As such HD species would be ideal candidates to test whether purifying selection is a pervasive force shaping the admixture landscape (Box 1, OQ1 and OQ2 on the interplay between evolutionary forces in admixed genomes). Ants could represent good models to test this hypothesis, as hybridization is fairly common in this group [54].

Understanding the evolution of barriers to gene flow in sex chromosomes

In many diploid hybridizing systems, sex chromosomes display less admixture than autosomes, being the strongest barriers for gene flow [55–58]. This could be due to either recessivity of BDMIs (Haldane's rule, dominance hypothesis) or sex-specific evolution (Haldane's rule, faster-X and faster-male hypotheses). In HDs, the whole genome follows the inheritance pattern of sex chromosomes of heterogametic sex systems (see [24] for a review of Haldane's rule in HDs). This has two practical consequences. First, all BDMIs in HDs are guaranteed to be between loci situated on chromosomes with a sex-chromosome-like inheritance pattern. Conversely, in diploids, such BDMIs can either be situated both on the sex chromosome (e.g., X-X), both on the autosomes (A-A), or one on an autosome and the other on a sex chromosome (e.g., A-X). In diploid heterogametic sex systems, these BDMIs fully located on the sex chromosome are important for speciation as they are the strongest barriers for gene flow. This is especially true in the presence of dosage compensation (an allele in haploid males has the same effect than its homozygous pair in females) or when migration is male-biased (i.e., with lower effective migration rates of the sex chromosome) [59]. As some HD species harbor diploid males (e.g., *Euodynerus foraminatus* wasps [60]) or they can be artificially made (e.g., *Nasonia* wasps [61]), it could be possible to disentangle whether Haldane's rule is linked to ploidy differences or sex [61]. Second, although effective population sizes differ between sex chromosomes and autosomes in heterogametic sex systems, there is no such difference in HDs. The efficiency of selection and genetic diversity levels are thus comparable genome-wide. Therefore, since the mode of inheritance of sex chromosomes is extended to the whole genome in HDs, results derived from HDs could help us infer the impact of sex-linked BDMIs in heterogametic sex systems

(Box 1, OQ7 on the genomic basis of fitness differences between hybrids and parents).

An integrated framework to study introgression in action

Some of the best-characterized hybridization models arose through ancient admixture events where sampling parental lineages may not always be possible (but see, e.g., [62]). Parental lineages could be unknown (e.g., *Heliconius* butterflies [16,63]), or available in very limited quantities (e.g., archaic humans [64]). As in some diploid systems [e.g., 62], there are several cases of contemporary hybridization already documented in HDs where parental lineages can be studied along with admixed populations (Box 2), which provides several advantages. First, if parental lineages are sufficiently diverged, they could help distinguish admixture from **incomplete lineage sorting** without requiring coalescent simulations [65]. Studying parental lineages also permits controlling for their demographic histories before admixture. Second, this framework from parents to hybrids would reveal how the recombination landscape itself is impacted by admixture (Box 1, OQ6 on the conservation of the recombination landscape between parents and hybrids). Third, comparing parents and hybrids could help document the temporal dynamics of adaptive introgression [66]. This would also help address whether evolution is repeatable, with e.g., recurrent selective sweeps in parental and hybrid populations before and after the admixture event (Box 1, OQ3 on the tempo of adaptive introgression).

Concluding remarks

Admixture appears to be an important mechanism shaping the tree of life, but we still lack a clear understanding on how divergent genomes evolve in admixed individuals. Studying hybridizing HDs is important on its own, as they represent a significant fraction of biodiversity. In addition, here we argue that HD organisms present several interesting features which would help answer many open questions in admixture research. Their study could help link classical speciation genetic studies on BDMIs in the lab with admixture patterns observed in natural populations. The difference in ploidy levels in HD hybrids is a potential asset to map BDMIs and go beyond the standard description of admixture or divergence patterns. This could help identify groups of genes involved in specific traits of interest (see Outstanding Questions) and could reveal complex epistasic interactions (see [67] for an example in *Nasonia* wasps).

Finally, some of the advantages of studying HDs we highlighted could be used in other systems in which life stages that display ploidy differences can easily be sampled (e.g., brown algae, conifers).

Acknowledgments

We are grateful to three anonymous reviewers, the editor Andrea Stephens, Konrad Lohse and Ina Satokangas for insightful comments on earlier versions of this manuscript. We thank Robin Bagley, Alain Migeon and Liselotte Sundström for providing pictures of *N. lecontei*, *T. evansi* and *Formica* sp., respectively. We also apologize to colleagues whose work could not be cited due to space constraints. J.K. was supported by Marie Curie fellowship and J.K. and P.N. were supported by a HiLIFE fellowship. C.B. is grateful for support by EMBO Installation Grant IG4152.

A.B. and C.B. were supported by ERC Starting Grant 804569 and Fundação Calouste Gulbenkian.

References

1 Mallet, J. *et al.* (2016) How reticulated are species? *BioEssays* 38, 140–149

2 Butlin, R.K. and Smadja, C.M. (2018) Coupling, Reinforcement, and Speciation. *Am. Nat.* 191, 155–172

3 Abbott, R. *et al.* (2013) Hybridization and speciation. *J. Evol. Biol.* 26, 229–246

4 Todesco, M. *et al.* (2016) Hybridization and extinction. *Evol. Appl.* 9, 892–908

5 Blanckaert, A. and Bank, C. (2018) In search of the Goldilocks zone for hybrid speciation. *PLOS Genet.* 14, e1007613

6 Taylor, S.A. and Larson, E.L. (2019) Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nat. Ecol. Evol.* 3, 170–177

7 Suarez-Gonzalez, A. *et al.* (2018) Adaptive introgression: a plant perspective. *Biol. Lett.* 14, 20170688

8 Marques, D.A. *et al.* (2019) A Combinatorial View on Speciation and Adaptive Radiation. *Trends Ecol. Evol.* 34, 531–544

9 Tao, Y. and Hartl, D.L. (2006) Genetic Dissection of Hybrid Incompatibilities Between *Drosophila simulans* and *D. mauritiana*. III. Heterogeneous Accumulation of Hybrid Incompatibilities, Degree of Dominance, and

294 Implications for Haldane's Rule. *Evolution (N. Y).* 57, 2580

295 10 Coyne, J.A. and Orr, H.A. (2004) *Speciation*, Sinauer Associates.

296 11 Matute, D.R. *et al.* (2010) A test of the snowball theory for the rate of evolution
297 of hybrid incompatibilities. *Science* 329, 1518–21

298 12 Trier, C.N. *et al.* (2014) Evidence for Mito-Nuclear and Sex-Linked
299 Reproductive Barriers between the Hybrid Italian Sparrow and Its Parent
300 Species. *PLoS Genet.* 10, e1004075

301 13 Racimo, F. *et al.* (2015) Evidence for archaic adaptive introgression in
302 humans. *Nat. Rev. Genet.* 16, 359–371

303 14 Edelman, N.B. *et al.* (2018) Genomic architecture and introgression shape a
304 butterfly radiation. *bioRxiv* DOI: 10.1101/466292

305 15 Schumer, M. *et al.* (2018) Natural selection interacts with recombination to
306 shape the evolution of hybrid genomes. *Science* 360, 656–660

307 16 Martin, S.H. *et al.* (2019) Recombination rate variation shapes barriers to
308 introgression across butterfly genomes. *PLOS Biol.* 17, e2006288

309 17 Lowry, E. *et al.* (2013) Biological invasions: a field synopsis, systematic review,
310 and database of the literature. *Ecol. Evol.* 3, 182–196

311 18 Scheffers, B.R. *et al.* (2016) The broad footprint of climate change from genes
312 to biomes to people. *Science* 354, aaf7671

313 19 Oziolor, E.M. *et al.* (2019) Adaptive introgression enables evolutionary rescue
314 from extreme environmental pollution. *Science* 364, 455–457

20 Grabenstein, K.C. and Taylor, S.A. (2018) Breaking Barriers: Causes, Consequences, and Experimental Utility of Human-Mediated Hybridization. *Trends Ecol. Evol.* 33, 198–212

21 Kyriazis, C.C. *et al.* (2019) High genetic diversity can contribute to extinction in small populations. *bioRxiv* DOI: 10.1101/678524

22 Blackmon, H. *et al.* (2017) Sex Determination, Sex Chromosomes, and Karyotype Evolution in Insects. *J. Hered.* 108, 78–93

23 Mable, B.K. and Otto, S.P. (1998) The evolution of life cycles with haploid and diploid phases. *BioEssays* 20, 453–462

24 Koevoets, T. and Beukeboom, L.W. (2009) Genetics of postzygotic isolation and Haldane's rule in haplodiploids. *Heredity* 102, 16–23

25 Favreau, E. *et al.* (2018) Genes and genomic processes underpinning the social lives of ants. *Curr. Opin. Insect Sci.* 25, 83–90

26 i5K Consortium (2013) The i5K initiative: Advancing arthropod genomics for knowledge, human health, agriculture, and the environment. *J. Hered.* 104, 595–600

27 Boomsma, J. *et al.* (2017) The Global Ant Genomics Alliance (GAGA). *Myrmecological News* 25, 61–66

28 Patten, M.M. *et al.* (2015) Biased introgression of mitochondrial and nuclear genes: a comparison of diploid and haplodiploid systems. *Mol. Ecol.* 24, 5200–5210

29 Lohse, K. and Ross, L. (2015) What haplodiploids can teach us about hybridization and speciation. *Mol. Ecol.* 24, 5075–5077

30 Maheshwari, S. and Barbash, D.A. (2011) The Genetics of Hybrid Incompatibilities. *Annu. Rev. Genet.* 45, 331–355

31 Rieseberg, L.H. *et al.* (1995) Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature* 375, 313–316

32 Schumer, M. *et al.* (2015) Reproductive Isolation of Hybrid Populations Driven by Genetic Incompatibilities. *PLOS Genet.* 11, e1005041

33 Charlesworth, B. *et al.* (1987) The Relative Rates of Evolution of Sex Chromosomes and Autosomes. *Am. Nat.* 130, 113–146

34 Kulmuni, J. and Pamilo, P. (2014) Introgression in hybrid ants is favored in females but selected against in males. *Proc. Natl. Acad. Sci. U. S. A.* 111, 12805–10

35 Knegt, B. *et al.* (2017) Detection of genetic incompatibilities in non-model systems using simple genetic markers: hybrid breakdown in the haplodiploid spider mite *Tetranychus evansi*. *Heredity* 118, 311–321

36 Koevoets, T. *et al.* (2012) Hybrid incompatibilities in the parasitic wasp genus *Nasonia*: negative effects of hemizygosity and the identification of transmission ratio distortion loci. *Heredity* 108, 302–311

37 Kulmuni, J. *et al.* (2019) Evidence for natural selection and barrier leakage in candidate loci underlying speciation in wood ants. *bioRxiv* DOI: 10.1101/500116

38 Schumer, M. and Brandvain, Y. (2016) Determining epistatic selection in admixed populations. *Mol. Ecol.* 25, 2577–2591

39 Browning, S.R. and Browning, B.L. (2011) Haplotype phasing: existing methods and new developments. *Nat. Rev. Genet.* 12, 703–714

40 Lohse, K. *et al.* (2010) Quantifying the pleistocene history of the oak gall parasitoid *Cecidostiba fungosa* using twenty intron loci. *Evolution* 64, 2664–2681

41 de la Filia, A.G. *et al.* (2015) Haplodiploidy and the reproductive ecology of Arthropods. *Curr. Opin. Insect Sci.* 9, 36–43

42 Nielsen, R. (2005) Molecular Signatures of Natural Selection. *Annu. Rev. Genet.* 39, 197–218

43 Chin, C.-S. *et al.* (2016) Phased diploid genome assembly with single-molecule real-time sequencing. *Nat. Methods* 13, 1050–1054

44 Kingan, S. *et al.* (2019) A High-Quality De novo Genome Assembly from a Single Mosquito Using PacBio Sequencing. *Genes* 10, 62

45 Ghenu, A.H. *et al.* (2018) Conflict between heterozygote advantage and hybrid incompatibility in haplodiploids (and sex chromosomes). *Mol. Ecol.* 27, 3935–3949

46 Barrett, R.D.H. *et al.* (2019) Linking a mutation to survival in wild mice. *Science* 363, 499–504

47 Sankararaman, S. *et al.* (2014) The genomic landscape of Neanderthal

379 ancestry in present-day humans. *Nature* 507, 354–357

380 48 Vernot, B. and Akey, J.M. (2014) Resurrecting Surviving Neandertal Lineages
381 from Modern Human Genomes. *Science* 343, 1017–1021

382 49 Fu, Q. *et al.* (2016) The genetic history of Ice Age Europe. *Nature* 534, 200–
383 205

384 50 Harris, K. and Nielsen, R. (2016) The Genetic Cost of Neanderthal
385 Introgression. *Genetics* 203, 881–91

386 51 Juric, I. *et al.* (2016) The Strength of Selection against Neanderthal
387 Introgression. *PLOS Genet.* 12, e1006340

388 52 Petr, M. *et al.* (2019) Limits of long-term selection against Neandertal
389 introgression. *Proc. Natl. Acad. Sci. U. S. A.* 116, 1639–1644

390 53 Romiguier, J. *et al.* (2014) Population genomics of eusocial insects: the costs
391 of a vertebrate-like effective population size. *J. Evol. Biol.* 27, 593–603

392 54 Feldhaar, H. *et al.* (2008) Lifelong commitment to the wrong partner:
393 hybridization in ants. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 2891–2899

394 55 True, J.R. *et al.* (1996) A genome-wide survey of hybrid incompatibility factors
395 by the introgression of marked segments of *Drosophila mauritiana*
396 chromosomes into *Drosophila simulans*. *Genetics* 142, 819–37

397 56 Jiggins, C.D. *et al.* (2001) Reproductive isolation caused by colour pattern
398 mimicry. *Nature* 411, 302–305

399 57 Payseur, B.A. *et al.* (2004) Differential patterns of introgression across the X

chromosome in a hybrid zone between two species of house mice. *Evolution* 58, 2064–2078

58 Runemark, A. *et al.* (2018) Hybridization, sex-specific genomic architecture and local adaptation. *Philos. Trans. R. Soc. B Biol. Sci.* 373, 20170419

59 Höllinger, I. and Hermisson, J. (2017) Bounds to parapatric speciation: A Dobzhansky–Muller incompatibility model involving autosomes, X chromosomes, and mitochondria. *Evolution* 71, 1366–1380

60 Cowan, D.P. and Stahlhut, J.K. (2004) Functionally reproductive diploid and haploid males in an inbreeding hymenopteran with complementary sex determination. *Proc. Natl. Acad. Sci. U. S. A.* 101, 10374–9

61 Beukeboom, L.W. *et al.* (2015) Hybrid incompatibilities are affected by dominance and dosage in the haplodiploid wasp *Nasonia*. *Front. Genet.* 6, 140

62 Runemark, A. *et al.* (2018) Variation and constraints in hybrid genome formation. *Nat. Ecol. Evol.* 2, 549–556

63 Martin, S.H. *et al.* (2013) Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* 23, 1817–1828

64 Sankararaman, S. *et al.* (2012) The Date of Interbreeding between Neandertals and Modern Humans. *PLoS Genet.* 8, e1002947

65 Joly, S. *et al.* (2009) A Statistical Approach for Distinguishing Hybridization and Incomplete Lineage Sorting. *Am. Nat.* 174, E54–E70

66 Jagoda, E. *et al.* (2018) Disentangling Immediate Adaptive Introgression from

Selection on Standing Introgressed Variation in Humans. *Mol. Biol. Evol.* 35, 623–630

67 Cohen, L.B. *et al.* (2019) Genetic, morphometric, and molecular analyses of interspecies differences in head shape and hybrid developmental defects in the wasp genus *Nasonia*. *bioRxiv* DOI: 10.1101/663732

68 de la Filia, A.G. *et al.* (2019) No evidence for an intragenomic arms race under paternal genome elimination in *Planococcus* mealybugs. *J. Evol. Biol.* 32, 491–504

69 Yang, X.-M. *et al.* (2012) Invasion Genetics of the Western Flower Thrips in China: Evidence for Genetic Bottleneck, Hybridization and Bridgehead Effect. *PLoS One* 7, e34567

70 Tay, W.T. *et al.* (2012) Will the Real *Bemisia tabaci* Please Stand Up? *PLoS One* 7, e50550

71 Elfekih, S. *et al.* (2018) Genome-wide analyses of the *Bemisia tabaci* species complex reveal contrasting patterns of admixture and complex demographic histories. *PLoS One* 13, e0190555

72 Beresford, J. *et al.* (2017) Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown. *Mol. Ecol.* 26, 4013–4026

73 Kulmuni, J. *et al.* (2010) Segregation distortion causes large-scale differences between male and female genomes in hybrid ants. *Proc. Natl. Acad. Sci. U. S. A.* 107, 7371–6

74 Werren, J.H. and Loehlin, D.W. (2009) The parasitoid wasp *Nasonia*: an emerging model system with haploid male genetics. *Cold Spring Harb. Protoc.* 2009, pdb.emo134

75 Breeuwer, J.A.J. and Werren, J.H. (1990) Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. *Nature* 346, 558–560

76 Dittmer, J. *et al.* (2016) Disentangling a Holobiont – Recent Advances and Perspectives in *Nasonia* Wasps. *Front. Microbiol.* 7, 1478

77 Bendall, E.E. *et al.* (2017) Oviposition traits generate extrinsic postzygotic isolation between two pine sawfly species. *BMC Evol. Biol.* 17, 26

78 Linnen, C.R. and Farrell, B.D. (2007) Mitonuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (Hymenoptera: Diprionidae) sawflies. *Evolution* 61, 1417–1438

79 Wachi, N. *et al.* (2012) Speciation history of three closely related oak gall wasps, *Andricus mukaigawae*, *A. kashiwaphilus*, and *A. pseudoflos* (Hymenoptera: Cynipidae) inferred from nuclear and mitochondrial DNA sequences. *Mol. Ecol.* 21, 4681–4694

80 Migeon, A. *et al.* (2010) Spider Mites Web: A comprehensive database for the Tetranychidae. In *Trends in Acarology* pp. 557–560

81 Boubou, A. *et al.* (2012) Test of Colonisation Scenarios Reveals Complex Invasion History of the Red Tomato Spider Mite *Tetranychus evansi*. *PLoS One* 7, e35601

Figure legends

Figure 1. Roadmap to detecting BDMIs in haplodiploids. A: Schematic representation of all possible genotypes grouped by fitness (displayed as signs below genotypes and/or genotype groups, see Online Supplementary Material Table S1 for the full fitness scheme) in a HD population with a recessive BDMI (blue and yellow are derived alleles, grey is the ancestral allele) following an initial admixture event. Note that hybrid males are only formed at F2 (red asterisk). B and C: Expected change in ancestral frequency of SNPs (frequency after selection *minus* frequency before selection) within a generation is displayed along an autosome for five (panel B) and 20 (panel C) generations after an admixture event (isolated hybrid population created at 50/50 admixture proportions, with random mating after admixture). In HD males (plain red line), the two-locus BDMI (vertical black lines) creates a much stronger signal of allele frequency change than in HD females (plain blue line) or diploids (D, dashed lines). In addition, this signal decays more slowly after admixture in haplodiploids than in diploids. See also Online Supplementary Material Figures S1 and S2.

Figure I. Promising haplodiploid models for the study of admixture. A. Mating pair of *Neodiprion lecontei* pine sawflies (male is smaller, on the left side, picture by Robin K. Bagley). B. *Bemisia tabaci* whiteflies (picture by Stephen Ausmus, source: Wikimedia Commons). C. Hybrid *Formica aquilonia* × *F. polyctena* sexuals (queen on the left, male on the right, picture by Liselotte Sundström). D. *Tetranychus evansi* spider mites (male is smaller, on top of the female, picture by Alain Migeon). All images used with permission.

Text boxes

Text Box 1. Open questions in admixture research

1. How (rapidly) are advantageous, deleterious and neutral variants sorted in hybrid genomes?
2. What are the relative contributions of adaptive introgression, selection against deleterious introgressed variants, and genetic drift to admixture patterns along the genome?
3. Do introgressed alleles tend to have a history of positive selection in their ancestral genome? Alternatively, do they tend to become beneficial only in hybrid or foreign genomes and if so, how much time after the admixture event?
4. To which extent is selection acting on introgressed variants context- or environment-dependent?
5. How does recombination rate variation affect selective processes and evolution of hybrid genomes?
6. To which extent do parental recombination landscapes predict recombination in hybrid genomes?
7. What are the phenotypic consequences of hybridization and what is the genomic basis of (sex-specific) fitness differences between hybrid and parental lineages (e.g., heterosis, hybrid breakdown)?
8. How repeatable are the hybridization process and the admixture landscape?
9. In the light of global change, will hybridization allow populations to adapt to rapidly changing or novel environmental conditions, purge genetic load or rescue inbred populations?

514 10. Does hybridization between native and alien lineages threaten species diversity
515 by leading to lineage homogenization and/or extinction?
516
517 **Text Box 2.** Promising haplodiploid systems for the study of admixture.
518 In this non-exhaustive list, we highlight various HD systems (Fig. I) that could
519 increase our understanding of genome-wide patterns of admixture (see also, e.g.,
520 *Planococcus* mealybugs (Insecta, Hemiptera) [68] and *Frankliniella occidentalis*
521 western flower thrips (Insecta, Thysanoptera) [69]).
522 ***Bemisia tabaci* whiteflies (Insecta, Hemiptera)**
523 Whiteflies are a major crop pest and vectors for over 300 plant viruses. *B. tabaci* is a
524 complex of at least 34 genetic lineages [70]. There is evidence for admixture
525 between some of these lineages in Africa and in Europe, where gene flow could be
526 human-induced as a consequence of ornamental plant trading [71].
527 ***Formica aquilonia & F. polyctena* wood ants (Insecta, Hymenoptera)**
528 These two wood ant species hybridize in Southern Finland [72]. In one hybrid
529 population, two genetic lineages co-occur and the extent of introgression between
530 lineages is sex-specific [73]. Several putative independent cases of hybridization
531 have been detected [72], enabling studies on the repeatability of hybridization
532 outcomes (Box 1, OQ8).
533 ***Nasonia* jewel wasps (Insecta, Hymenoptera)**
534 *Nasonia* is a complex of four parasitoid wasp species and represents a model genus
535 for evolutionary and developmental genomics [74]. All *Nasonia* species are infected
536 with *Wolbachia*, a bacterium which causes cytoplasmic incompatibility and triggers
537 (almost) complete reproductive isolation in the wild [75]. However, in the lab,
538 antibiotic treatment rescues from hybrid inviability. F2 males display different

539 postzygotic hybrid breakdown traits, depending on environmental factors and cross
540 direction [76].

541 ***Neodiprion lecontei* & *N. pinetum* sawflies (Insecta, Hymenoptera)**

542 These two sister species of pine specialists differ in host use, but mate and produce
543 viable and fertile offspring [77]. Introgression has been detected in the wild, where
544 mitochondrial gene flow is higher than nuclear gene flow [78]. A simulation study
545 suggested this asymmetric introgression pattern could be a general feature of HDs
546 linked to their transmission genetics ([28], see [79] for another example in wasps).

547 ***Tetranychus evansi* spider mites (Arachnida, Trombidiformes)**

548 The tomato red spider mite *T. evansi* is a Solanaceae specialist native from South
549 America that has spread almost all over the globe [80]. There are two genetic
550 lineages which hybridize in the field [81]. This system is easily tractable in the lab
551 and BDMI mapping in F2 individuals revealed strong male hybrid breakdown along
552 with some evidence for heterosis [35].

553

554

Glossary

Adaptive introgression: transfer of a segment of nuclear DNA from one lineage to another, providing a fitness advantage to the recipient individual, leading to its spread and possible fixation in the recipient population due to natural selection.

Adaptive radiation: quick diversification of an ancestral population into multiple, ecologically and phenotypically divergent lineages.

Admixture: mixing of genomes from genetically distinct populations within the same individual.

Arrhenotoky: a type of haplodiploidy where haploid males develop from unfertilized eggs.

Bateson-Dobzhansky-Muller incompatibility: combination of alleles from divergent lineages that leads to a lower than expected fitness value when together in the same individual.

Effective population size: hypothetical size of a population under the Wright–Fisher model (panmictic and of constant size) that best reproduces the observed population genetic statistics (e.g., genetic diversity).

Haplodiploidy: life cycle (or sex-determination system) in which one sex is haploid and the other is diploid.

Haldane's rule: Haldane's rule states that in hybrids, the heterogametic sex (i.e., XY or ZW individuals) is more likely to be inviable or sterile than the homogametic sex (i.e., XX or ZZ individuals). The dominance hypothesis posits that Haldane's rule occurs because the heterogametic sex expresses both dominant and recessive sex-linked genetic incompatibilities whereas the homogametic sex only those which are dominant. The Faster-X hypothesis states that since X-linked genes evolve faster than autosomal ones because of hemizygosity, they have a disproportionate effect

on hybrid fitness. The faster-male hypothesis states that male genes evolve faster due to sexual selection, which would underlie hybrid male sterility or inviability in male heterogametic taxa.

Heterosis: greater fitness of hybrid offspring compared to that of their parents.

Hybridization: reproduction between members of genetically distinct lineages, producing offspring of mixed ancestry (i.e., admixed).

Incomplete lineage sorting: discrepancy between gene and species trees, due to the segregation of ancestral polymorphism in divergent lineages.

Introgression: transfer of genetic material between genetically distinct populations, mediated by hybridization and backcrossing.

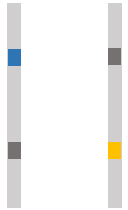
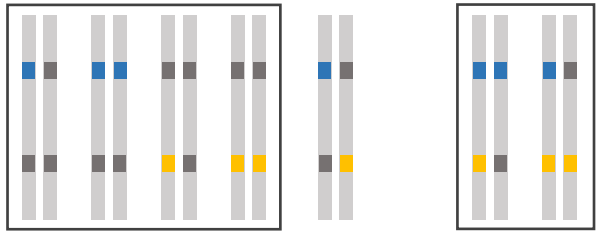
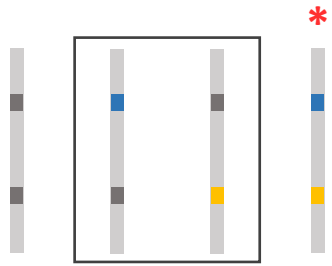
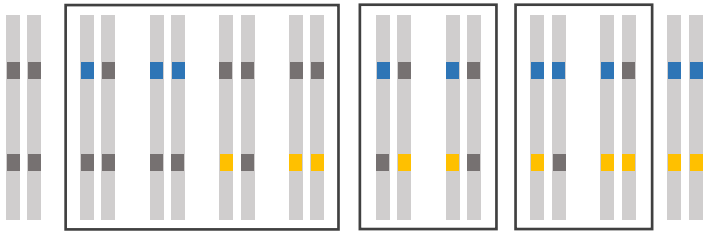
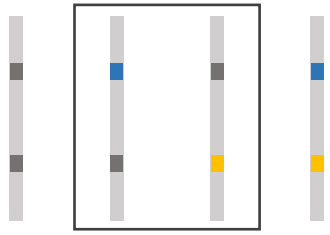
Paternal genome elimination: a type of haplodiploidy where males develop from fertilized eggs, but lose the father's genome from soma and/or germline. In species with diploid PGE, where the parental genome is only lost from the germline, males are diploid. However gene expression in males of some species is haplodized through transcriptional repression of the father's genome.

Phasing: process by which alleles carried by a diploid individual are assigned to the paternal and maternal chromosomes. Computational phasing estimates haplotype phase from genotype data either from parent-offspring trios (when available), or by modelling haplotype frequencies in a population.

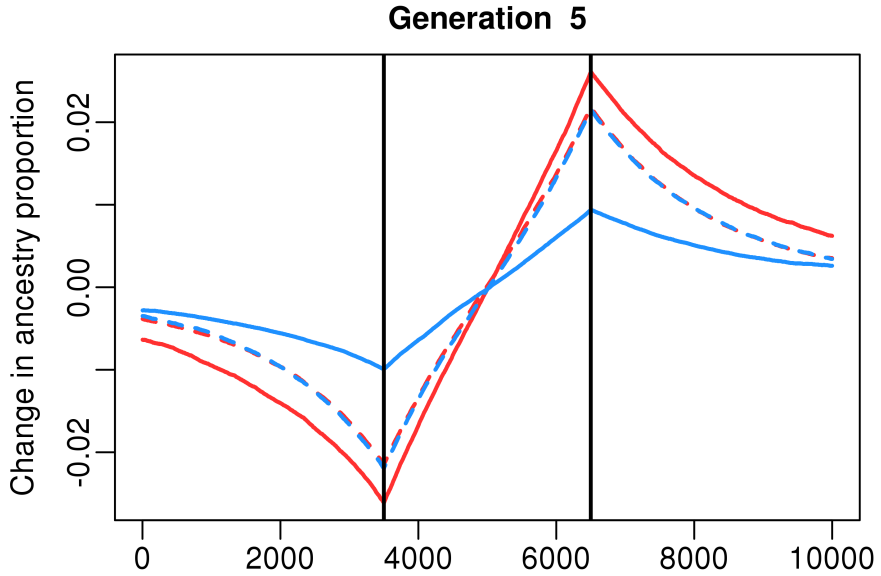
Purifying selection: Removal (so-called purging) of deleterious alleles in a population.

Reinforcement: indirect selection for stronger reproductive isolation between divergent lineages in secondary contact to mediate the cost of producing unfit hybrids.

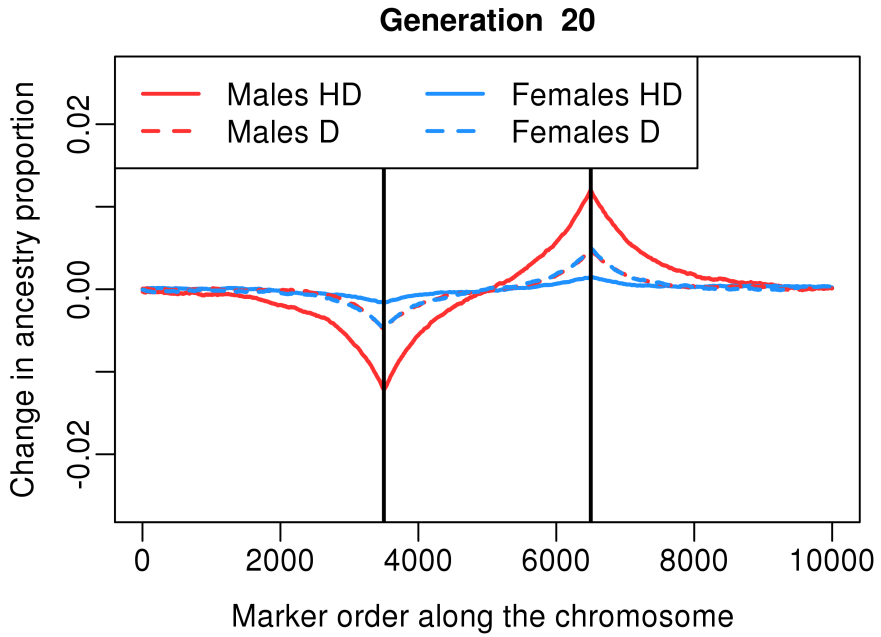
A.

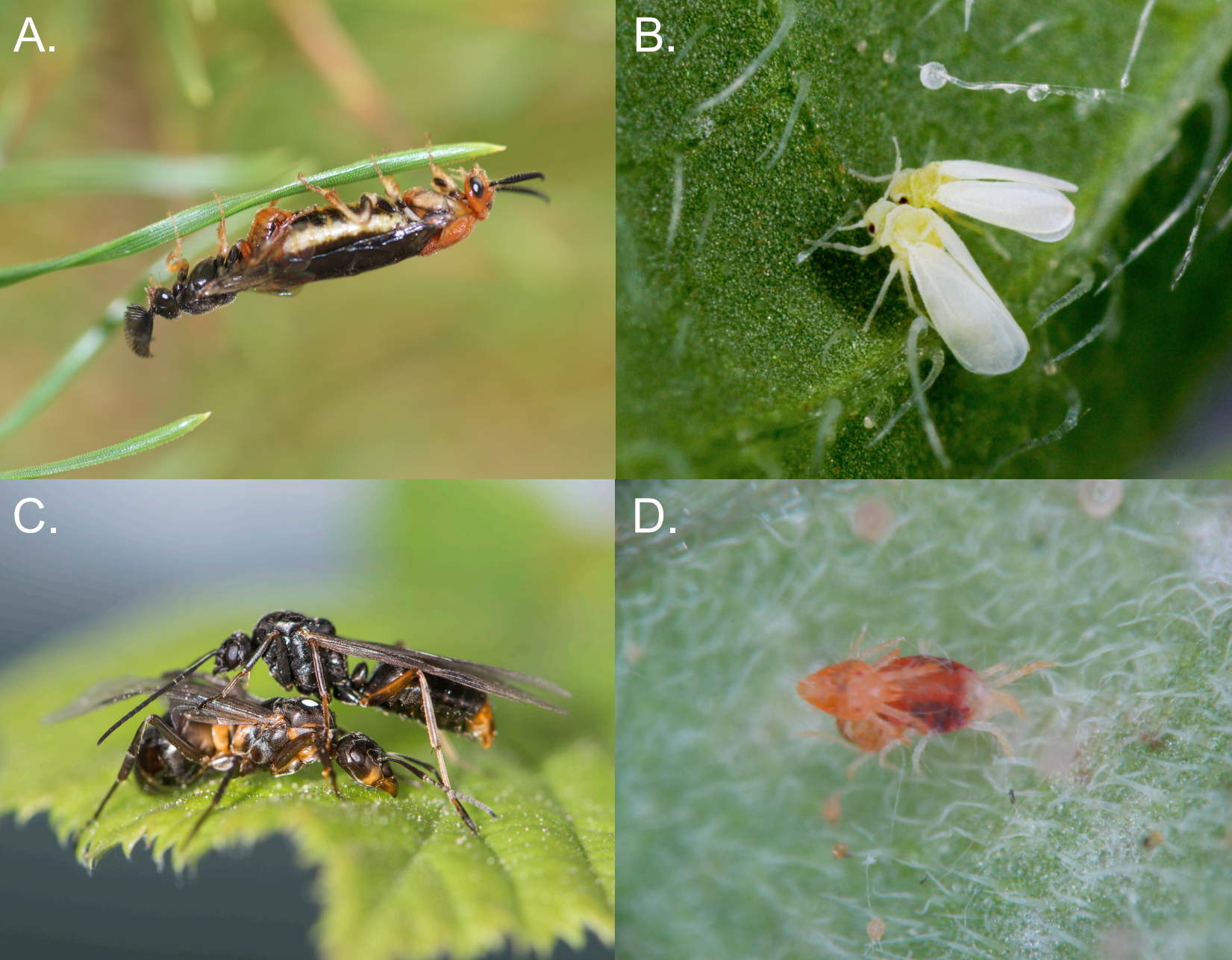
Generation	All possible genotypes in	
	diploid females	haploid males
F1		
F2		
Fitness:	+ ++ -	= + --
F3 onwards		
Fitness:	= + ++ - --	= + --

B.



C.




A.
B.
C.
D.

Outstanding Questions

- To what extent is admixture selected against in functional genomic regions? In haploid males, any introgressed allele that would disrupt a gene network would be selected against independently of its dominance. This feature could be used to map groups of genes which are involved in the same biological process and, coupled with gene expression data, move towards determining the functional basis of interacting genomic regions in natural populations.
- Empirical work and simulations suggest that mitochondrial DNA introgresses at higher rates than nuclear DNA in haplo-diploid organisms (HDs). How general is the asymmetric cytonuclear introgression pattern in HDs and what are its consequences? For instance, are mito-nuclear incompatibilities more prone to act as barriers to gene flow in hybridizing HDs than in diploids?
- There is no recombination in haploid males, where all recessive alleles are exposed to selection. What are the consequences of these two factors on the evolution of admixed genomes?

Supplementary Material for

Understanding admixture: haplodiploidy to the rescue

Pierre Nouhaud[1], Alexandre Blanckaert[2], Claudia Bank[2] & Jonna Kulmuni[1]

[1] Organismal & Evolutionary Biology Research Programme, University of Helsinki, Finland
[2] Instituto Gulbenkian de Ciência, Oeiras, Portugal
Corresponding author: pierr3.nouhaud@gmail.com (P. Nouhaud),
https://www.helsinki.fi/en/researchgroups/evolution-sociality-behaviour

Simulation framework

We aimed to identify the signature of selection against recessive Bateson-Dobzhansky-Muller incompatibilities (BDMIs) in admixed populations, comparing haplodiploid (HD) and diploid life cycles. We simulated a 10,000 bp autosome carrying a BDMI which involves two diallelic loci A and B (resp. the 3500th and 6500th markers). Alleles A1 and B0 are fixed in population 1 and alleles A0 and B2 are fixed in population 2. Alleles A0 and B0 decrease the survival probability by a factor 1+*s* (here *s* = -0.02). There is negative epistasis between allele A1 and B2, decreasing the survival probability of an individual by a factor 1+*e* (here *e* = -0.4). The whole fitness table is given in Online Supplementary Material Table S1. All markers here are equidistant with a recombination rate $r = 10^{-4}$ between adjacent markers.

We followed the sex-specific change in ancestry proportion of SNPs (allele frequency after selection *minus* allele frequency before selection) for the first 100 generations after an admixture event (50/50 proportions, random mating after admixture). Results are averaged over 100 independent replicates.

SLiM code [1] used for simulations is available at https://gitlab.com/evoldyn/haplodiploidy.

Table S1. Fitness table used in our simulations. The BDMI involves two diallelic loci A and B, with alleles A1 and B0 fixed in population 1 and alleles A0 and B2 fixed in population 2. Alleles A0 and B0 decrease the survival probability by a factor 1+*s* (here *s* = -0.02). There is negative epistasis between allele A1 and B2, decreasing the survival probability of an individual by a factor 1+*e* (here *e* = -0.4). All markers here are equidistant with a recombination rate $r = 10^{-4}$ between adjacent markers. Fitness is defined as the probability of surviving until adulthood. Fitness values are indicated in bold.

Female	A0B0	A1B0	A0B2	A1B2
A0B0	$(1+ s_{A0})^2 (1+ s_{B0})^2$ **0.922**	$(1+ s_{B0})^2$ **0.960**	$(1+s_{A0})^2$ **0.960**	**1**
A1B0		$(1+ s_{B0})^2$ **0.960**	**1**	$(1 + e)^2$ **0.420**
A0B2			$(1+s_{A0})^2$ **0.960**	$(1 + e)^2$ **0.420**
A1B2				$(1 + e)^4$ **0.130**
Male	$(1+ s_{A0})^2 (1+ s_{B0})^2$ **0.922**	$(1+ s_{B0})^2$ **0.960**	$(1+s_{A0})^2$ **0.960**	$(1 + e)^4$ **0.130**

[see gif file]

Figure S1. Evolution of the signature of selection against recessive BDMIs (vertical lines) within admixed populations in diploid and haplodiploid populations. The change in ancestry proportion of SNPs (allele frequency after selection *minus* allele frequency before selection) within a generation is displayed along an autosome sequentially for a hundred generations. The generation number is indicated on top of the graph. Results shown are averaged over 100 independent replicates.

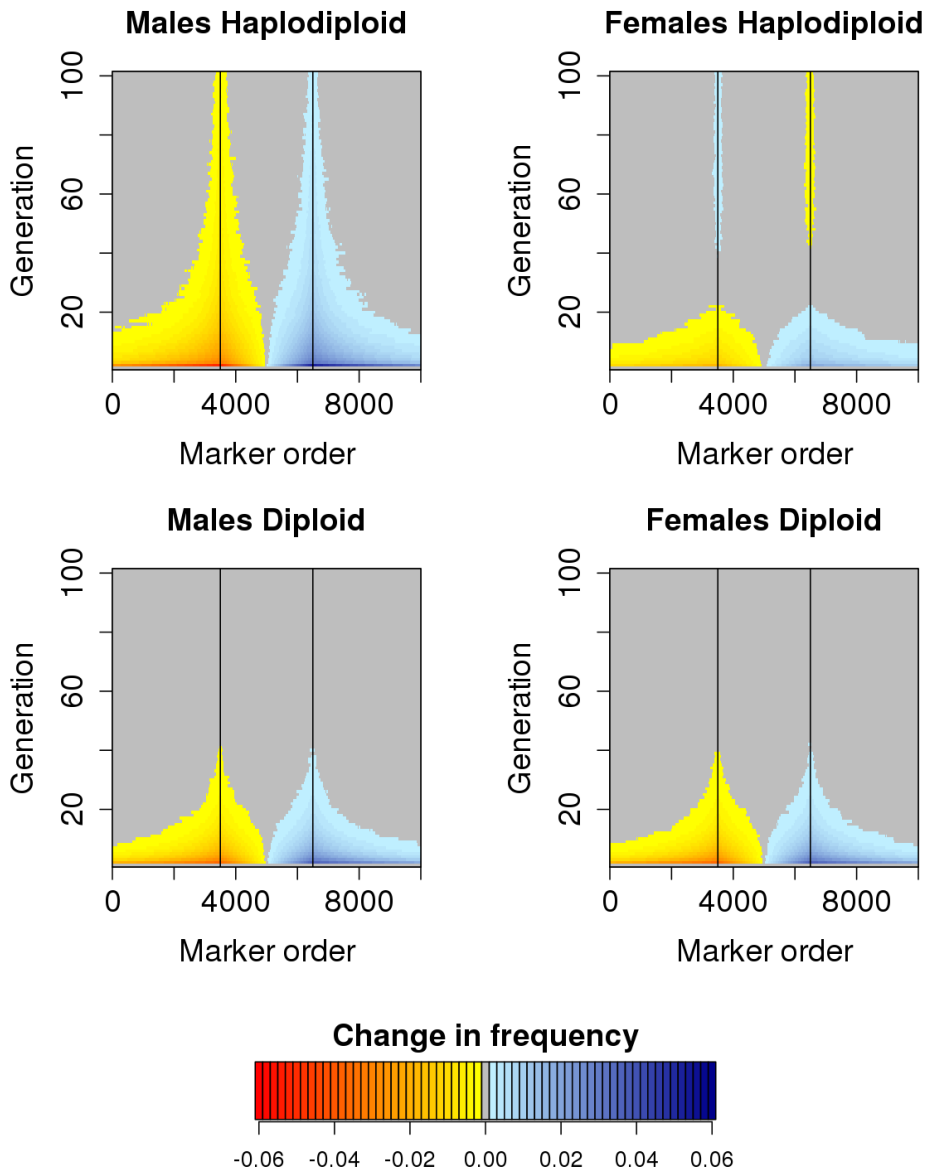


Figure S2. Heatmap displaying the evolution of the signature of selection against recessive BDMIs (vertical lines) for each sex after an admixture event at generation 0 in diploid and haplodiploid populations. The change in ancestry proportion of SNPs (allele frequency after selection *minus* allele frequency before selection, see color scale) is shown along an autosome (x-axis) for a hundred generations (y-axis). Results displayed are averaged over 100 independent replicates.

Reference

1 Haller, B.C. and Messer, P.W. (2019) SLiM 3: Forward Genetic Simulations Beyond the Wright–Fisher Model. *Mol. Biol. Evol.* 36, 632–637