Parental divergence as a predictor of hybridization outcomes

Poikela N*,**[1], Tunström K*[2], Stelkens R[3], Runemark A[2], Kulmuni J**[1,4].

[1)] Organismal and Evolutionary Biology, University of Helsinki, Finland
[2)] Department of Biology, Lund University, Sweden
[3)] Department of Zoology, Stockholm University, Sweden
[4)] Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, The Netherlands

* Shared first authorship
** Corresponding authors:
Jonna Kulmuni, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, The Netherlands, j.k.kulmuni@uva.nl
Noora Poikela, Organismal and Evolutionary Biology, University of Helsinki, Finland, noora.poikela@helsinki.fi

Abstract

Hybridization can generate both fitness benefits and genetic incompatibilities, yet predicting how these opposing effects shape hybrid fitness and genome evolution remains a major challenge in evolutionary genomics. Parental genetic divergence is expected to play a central role, as increasing divergence can promote heterosis and adaptive novelty while simultaneously increasing incompatibilities and barriers to gene flow. Here, we review the theoretical and empirical evidence linking parental divergence to homoploid hybridization outcomes and use meta-analyses to identify general patterns across taxa. We find that heterosis tends to decline with increasing parental divergence, consistent with accumulation of incompatibilities with divergence. Furthermore, we show that greater parental divergence is associated with lower minor parent ancestry in hybrid genomes, potentially explained by stronger constraints on hybrid genome formation set by accumulating incompatibilities. We also find a significant association between mitochondrial and major parent ancestry, consistent with mitonuclear incompatibilities constraining hybrid genome evolution. In contrast, our results concerning minor parent ancestry on sex chromosomes are inconclusive. Together, these findings suggest that parental divergence is an important predictor of both hybrid fitness and the retention of introgressed ancestry, while highlighting substantial variation in hybridization outcomes across taxa. Our analyses also reveal major limitations arising from inconsistent reporting and gaps in available data, underscoring the need for more standardized approaches to understanding and predicting the evolutionary consequences of hybridization.

Keywords: Hybridization, divergence, ancestry, incompatibilities, heterosis

Significance statement

Hybridization between species and populations has diverse outcomes that are challenging to predict. We show that genetic divergence between hybridizing populations predicts heterosis and ancestry variation in hybrid genomes, while knowledge gaps exist in synthesizing the role of environmental and genomic features in hybrid genome evolution.

Introduction

Hybridization, the interbreeding between genetically divergent populations producing offspring of mixed ancestry (Barton & Hewitt 1985), is an important evolutionary process that can generate novel genetic variation and facilitate adaptation, but can also reduce population fitness and threaten species persistence. Comparative genomic studies have revealed evidence of historical and contemporary introgression across a large diversity of eukaryotic lineages, including plants, fungi, insects, fishes, birds, and mammals (e.g. Martin et al. 2013; Edelman & Mallet 2021; Dagilis et al. 2022). Anthropogenic environmental change, habitat modification, biological invasions, and climate-driven range shifts are increasingly bringing previously isolated populations and species into secondary contact, generating new opportunities for hybridization (Hoffmann & Sgrò 2011; Grabenstein & Taylor 2018; Ottenburghs 2021; Sánchez-Guillén et al. 2024). Understanding the factors that determine whether hybridization facilitates adaptation and diversification or contributes to population decline and extinction is therefore increasingly important in a rapidly changing world.

Yet, hybridization can have remarkably diverse evolutionary outcomes (Fig. 1). Some hybridization events produce viable and fertile offspring, potentially facilitating genetic rescue (Aguilar-Gómez et al. 2025), adaptive introgression (Suarez-Gonzalez et al. 2018), the establishment of stable hybrid populations persisting for many generations (Nouhaud et al. 2022), or even the emergence of hybrid species (Mavárez et al. 2006). Other hybridization events produce inviable or sterile offspring, resulting in outbreeding depression and, in some cases, population decline (Bomblies et al. 2007; Todesco et al. 2016). Extensive hybridization and introgression can also erode species boundaries through genetic swamping or species fusion (Todesco et al. 2016; Howard-McCombe et al. 2023). Even among successful hybrid populations, genomes often differ markedly in the amount and genomic distribution of introgressed ancestry, reflecting variation in how selection removes or retains parental genetic material. On the other hand, comparative studies of independently formed (‘replicate’) hybrid populations have also revealed surprising repeatability in the outcomes of hybridization in terms of ancestry purging and retention, suggesting that hybrid genome evolution can be predictable despite the diversity of possible outcomes (Runemark et al. 2018; Chaturvedi et al. 2020; Langdon et al. 2022; Nouhaud et al. 2022; Owens et al. 2025). Understanding what determines this variation, and under which conditions hybridization outcomes can be predicted, remains a central challenge in evolutionary genomics, with important implications for biodiversity conservation, agricultural breeding, and biotechnology.

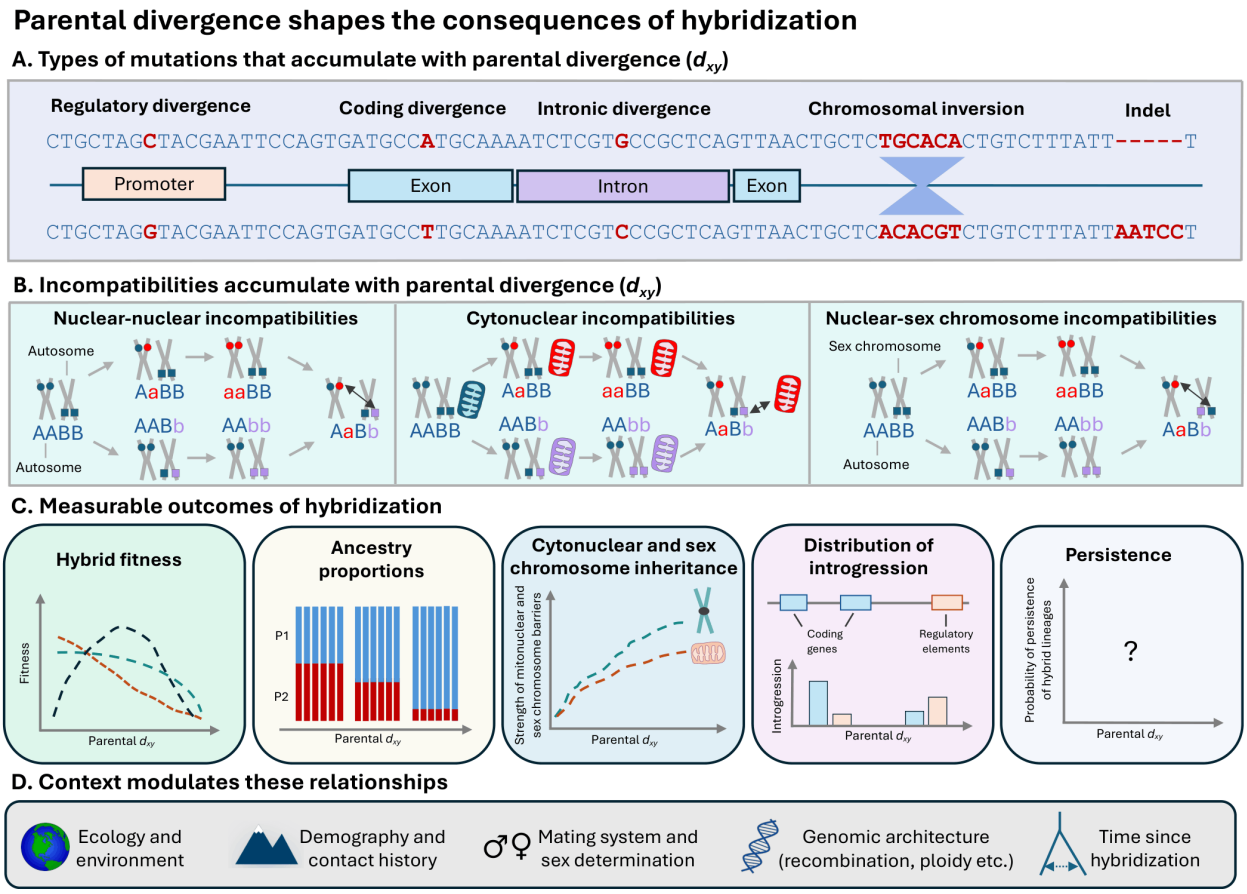


Figure 1. Parental divergence shapes consequences of hybridization. During divergence different types of mutations accumulate (A) that can lead to incompatibilities (B) but also genetic novelty. Parental divergence impacts the measurable outcomes of hybridization (C), incompatibilities increasingly constraining hybrid genome evolution with growing divergence. Context (D) modulates relationships between parental divergence and outcomes of hybridization.

Among the many factors influencing hybridization outcomes, the extent of evolutionary divergence between parental lineages has long been considered fundamental. As lineages diverge, they accumulate genetic differences through nucleotide substitutions, new structural variants, regulatory changes, and chromosomal rearrangements, some of which give rise to genetic incompatibilities when combined in hybrid genomes (Fig. 1). Theory predicts that increasing parental divergence can have opposing consequences for hybrid fitness and genome evolution. On the one hand, hybrids between more divergent lineages are expected to carry a greater number of Bateson-Dobzhansky-Muller incompatibilities (BDMIs; Orr 1995; Turelli & Orr 2000) including disrupted regulatory interactions and cytonuclear mismatches, as well as other incompatibilities involving sex chromosomes or other genomic regions involved in reproductive isolation. These incompatibilities can reduce hybrid fitness, limit introgression, and accelerate the removal of incompatible ancestry from hybrid genomes. On the other hand, greater divergence between hybridizing populations brings together increasingly distinct genetic variants, potentially generating heterosis, transgressive phenotypes, and novel allelic combinations in coding genes or other regulatory and functional elements, that enhance adaptive potential (Rieseberg et

al. 1999; Stelkens & Seehausen 2009). Thus, the same evolutionary divergence that increases the risk of genetic incompatibilities may also expand the genetic and phenotypic raw material upon which natural selection can act. Rather than implying a universally monotonic relationship between parental divergence and hybridization success, current theory (Fraïsse et al. 2016; Simon et al. 2018; Wei & Zhang 2018; Würschum et al. 2023) suggests that hybridization outcomes reflect a balance between opposing evolutionary forces, whose relative importance may vary among taxa, individuals carrying different allelic combinations (Barton 2001), and across stages of the speciation continuum.

These opposing theoretical predictions are reflected in the diversity of empirical observations. Some studies report declining hybrid fitness with increasing parental divergence (Stelkens et al. 2015; Turissini et al. 2017), while others document successful hybridization between deeply diverged lineages (Shapira et al. 2014; Bernardes et al. 2017). Similarly, hybrid genomes vary considerably in their ancestry composition, ranging from relatively balanced contributions of both parental lineages to strongly asymmetric genomic mosaics dominated by one parent. Introgressed ancestry may be retained in some genomic regions but largely excluded from others, reflecting the combined effects of selection, recombination, and genomic architecture. Cytonuclear interactions may favor the retention of nuclear alleles from the mitochondrial parent, while sex chromosomes often exhibit reduced introgression relative to autosomes (Martin et al. 2013; Fraïsse & Sachdeva 2021), although the strength and direction of these patterns vary among taxa (Fraïsse & Sachdeva 2021). Despite these observations, variation in hybrid fitness and genomic ancestry is often examined separately, even though both may be shaped by the same underlying process: the extent of evolutionary divergence between parental genomes.

Despite considerable theoretical and empirical progress, our understanding of how parental divergence shapes hybridization outcomes remains fragmented. Most studies focus on individual taxa, particular forms of reproductive isolation, or specific genomic features, making it difficult to identify general patterns and principles across biological systems. Recent reviews have synthesized the prevalence of hybridization (Taylor & Larson 2019), the adaptive consequences of introgression (Edelman & Mallet 2021), and the evolutionary mechanisms shaping hybrid genomes (Runemark et al. 2019; Moran et al. 2021; Schumer & Rieseberg 2026). However, a quantitative synthesis examining how parental divergence influences multiple hybridization outcomes across taxa is still lacking.

Here, we synthesize published genomic and fitness data to evaluate the extent to which parental genetic divergence predicts hybridization outcomes across eukaryotic taxa. We examine to what extent parental divergence (measured as d_{xy}) can predict heterosis, and asymmetries in parental genomic contributions across the genome and in sex chromosomes and mitochondria. By integrating evidence across diverse taxa, we aim to identify general evolutionary patterns, assess their consistency across biological systems, and highlight important exceptions and remaining knowledge gaps (Table 1). Ultimately, we ask whether parental divergence provides a common framework for understanding and predicting the fitness and genomic consequences of hybridization.

Table 1. Outstanding questions on the impact of parental divergence (d_{xy}) on the evolution of hybrid populations and genomes. The questions addressed in our meta-analyses are highlighted in grey.

Hybrid fitness	What is the relationship between parental divergence (d_{xy}) and hybrid vigour (heterosis) and does it vary across taxa?
Ancestry proportions	Are ancestry proportions in hybrids more skewed with higher parental d_{xy} consistent with accumulation of incompatibilities with increasing divergence?
Ancestry proportions	Is the fixation of parental ancestry in hybrid genomes accelerated with increasing parental d_{xy}?
Cytonuclear and sex chromosome inheritance	Is hybrid ancestry skewed towards the parent which has contributed the mitochondrion?
Cytonuclear and sex chromosome inheritance	How does parental d_{xy} impact the asymmetry of introgression on sex chromosomes compared to autosomes?
Distribution of introgression	Does the amount of introgression in genic regions differ from regulatory regions, and does this relationship vary depending on parental d_{xy}?
Persistence of hybrid lineages	How does parental d_{xy} affect the long-term persistence of resulting hybrid lineages?
Persistence of hybrid lineages	How does context (ecology, demography etc.) modulate the relationship between parental d_{xy} and the persistence of hybrid lineages?

The impact of parental divergence on hybrid fitness

Hybridization can produce offspring with lower, similar or higher fitness relative to their parental lineages (Arnold & Hodges 1995; Atsumi et al. 2021). The extent of genetic divergence between hybridizing populations is expected to be an important predictor of these fitness outcomes. Parental divergence is commonly quantified using d_{xy}, a measure of absolute nucleotide divergence that, unlike relative measures such as F_{st}, is not normalized by within-population diversity and is therefore more suitable for comparisons across taxa (Charlesworth 1998; Noor & Bennett 2009; Cruickshank & Hahn 2014). An alternative is to use D_a, or net nucleotide divergence, which subtracts within-population diversity from d_{xy}, to estimate divergence accumulated since the populations split (Hahn 2018).

Increasing parental genetic divergence is generally associated with stronger reproductive isolation and the accumulation of genetic incompatibilities, which can reduce hybrid fitness (Coyne & Orr 1989; Coughlan & Matute 2020). Theoretical work predicts that incompatibilities accumulate faster than linearly as lineages diverge over time, producing a "snowball" effect (Orr 1995; Orr & Turelli 2001). However, empirical tests of this prediction have yielded mixed results (Matute et al. 2010; Moyle & Nakazato 2010; Presgraves 2010). Nevertheless, broader comparative evidence indicates that increasing genetic divergence is associated with a pronounced decline in gene flow between lineages. A meta-analysis using an approximate Bayesian computation framework found that irrespective of species' life history traits and ecology, the probability of gene flow between diverging lineages decreases sharply between 0.5% to 2% of molecular divergence (Roux et al. 2016).

Interestingly, at comparable levels of genetic divergence, the probability of gene flow appears to be significantly lower in plants than in animals suggesting that barriers to gene flow may become established more rapidly in plants (Monnet et al. 2025). This finding is somewhat counterintuitive, given that hybridization is generally thought to occur more frequently in plants than in animals (Mallet 2005), although the occurrence of hybridization does not necessarily imply successful introgression. In another meta-analysis, a d_{xy} threshold of 0.20%-0.27% predicted current taxonomic status with more than 90% accuracy in large placental mammals with broadly comparable life history traits and generation times exceeding four years (de Jong et al. 2025). Such estimates, however, should not be interpreted as universal thresholds beyond which reproductive isolation is complete. The relationship between genetic divergence and reproductive isolation varies among taxa (Cerca & Stelkens 2026) and incompatibilities can arise at substantially lower levels of divergence, including between populations within the same species.

Combining divergent genomes through hybridization can generate incompatibilities, but it can also increase genetic variation and produce novel allelic combinations that can enhance fitness and provide new raw material for adaptation (Buerkle et al. 2000; Dagilis et al. 2019; Krapf et al. 2026). Individual hybrids may therefore carry both beneficial and deleterious allelic combinations, resulting in considerable fitness variation within hybrid populations. The magnitude of potential fitness benefits may also depend on parental genetic divergence: greater divergence can increase the range of genetic and phenotypic variation, potentially generating adaptive phenotypic novelty in hybrid offspring. However, these benefits may be counteracted by the accumulation of genetic incompatibilities, particularly at higher levels of divergence. Parental divergence may thus simultaneously increase both the potential costs and benefits of hybridization, with their relative contributions determining the net fitness consequences (Dagilis et al. 2019).

In addition to selection against incompatibilities and for novel beneficial allele combinations, several evolutionary mechanisms can contribute to fitness gains following hybridization. Heterosis or more generally hybrid vigor (Shull 1948), refers to increased hybrid performance relative to parental lineages and can arise through several genetic mechanisms (Li et al. 2024). It has been documented across diverse taxa, including microbes (Shapira et al. 2014; Bernardes et al. 2017), plants (Yu et al. 1997) and animals (Fitzpatrick & Shaffer 2007). One of the underlying mechanisms of heterosis is the masking of recessive deleterious alleles that have accumulated in one or both parental populations, for instance, through genetic drift in populations with small effective population sizes. In early generation hybrids, these deleterious alleles can be complemented by functional alleles inherited from the other parent, increasing hybrid fitness. In later hybrid generations, recombination and selection can remove these deleterious alleles, potentially favoring ancestry from the parent carrying the lower deleterious load at affected genomic regions (Nouhaud et al. 2022). Adaptive introgression can further increase fitness by introducing advantageous alleles into a different genetic background (Suarez-Gonzalez et al. 2018; Edelman & Mallet 2021). For a detailed discussion on different factors shaping hybrid genomes see BOX 1 and recent reviews by Moran et al. (2021) and Schumer & Rieseberg (2026).

Empirical studies have revealed contrasting relationships between parental divergence and hybrid fitness. For example, Stelkens et al. (2015) found that increasing parental genetic distance was associated with reduced F_2 viability across seven pairs of cichlid populations, while Turissini et al. (2017) reported a decline

in the ability of F_1 hybrids to locate food with increasing Nei's genetic distance (D) across 94 pairs of *Drosophila* species. Similarly, a meta-analysis across various animal taxa found that interspecific hybrids generally showed lower fitness than intraspecific hybrids, although divergence time calculated using TIMETREE did not significantly predict hybrid performance (Muraro et al. 2022).

In contrast, studies in yeast suggest that increasing parental divergence can enhance hybrid fitness, at least over relatively low levels of genetic differentiation. Because these studies compare individual strains rather than populations, parental divergence is typically quantified using *p*-distance, the proportion of nucleotide sites that differ between two sequences, which is conceptually similar to d_{xy}, but calculated between individual genomes. Wei & Zhang (2018) found that heterosis was highest at intermediate parental divergence (p-distance ~0.004), a non-monotonic relationship in which the fitness benefits of divergence may eventually be outweighed by the accumulation of incompatibilities. Strauss et al. (2026) similarly found increasing hybrid fitness with parental divergence, although their analysis was restricted to a narrower range (*p*-distance up to ~0.0015). Together, these findings highlight that the relationship between parental genetic divergence and hybrid fitness may depend on the range of divergence examined, the fitness components measured, and the biological system under study. Whether general patterns emerge across broader divergence ranges and taxonomic groups remains unresolved.

Removal of incompatibilities and variation in hybrid fitness across generations

When hybrid incompatibilities are sufficiently severe to cause inviability or sterility, they can prevent hybrids from realizing any potential fitness benefits of hybridization. However, many incompatibilities have weaker or context-dependent fitness effects (Arnold 1997; Arnegard et al. 2014; Thompson et al. 2022; Heidbreder et al. 2026). In hybrid populations that persist and reproduce, recombination progressively breaks up parental haplotypes, allowing selection to remove deleterious ancestry combinations while retaining or favoring beneficial ones. This process, referred to as genome stabilization or sorting, can substantially reshape both the genomic composition and fitness of hybrid populations over successive generations. Consequently, the fitness of early-generation hybrids may differ considerably from that of later-generation hybrids, as selection and recombination alter the frequencies of incompatible and favorable allele combinations (Pereira et al. 2021; Kulmuni et al. 2024; Kuijt et al. 2026).

In the simple case of a dominant two-locus incompatibility between derived alleles D1 and D2, selection can rapidly favor compatible allele combinations over successive hybrid generations. Consider two parental populations fixed for A1A1/D2D2 and D1D1/A2A2, respectively, where A1 and A2 represent the ancestral alleles at the two loci. Both parental genotypes are compatible because the derived alleles (D1 and D2) occur in separate genetic backgrounds, whereas F_1 hybrids (A1D1/A2D2) carry both derived alleles and therefore express the incompatibility. In subsequent generations, segregation and recombination generate a range of new genotypes, some of which continue to carry both incompatible derived alleles and consequently suffer reduced fitness. Selection can therefore favor genotypes lacking at least one of these alleles, including the original parental combinations A1A1/D2D2 and D1D1/A2A2 and the fully ancestral combination A1A1/A2A2. As compatible combinations increase in frequency, the incompatibility may become less prevalent in the hybrid population, potentially allowing hybrid fitness to recover.

However, the efficiency of incompatibility sorting depends strongly on its underlying genetic architecture (Xiong & Mallet 2022). Incompatibilities may be recessive, vary in strength, or involve interactions among

more than two loci, all of which can impact the likelihood and rate of their purging. Recessive incompatibilities may be purged more slowly, because their deleterious effects are masked in heterozygotes and become exposed only in particular homozygous genetic backgrounds (Blanckaert & Bank 2018). Multilocus incompatibilities can create more complex barriers to gene flow (Blanckaert et al. 2020), and the recombination and segregation events required to generate compatible allele combinations may further constrain their resolution, thereby slowing their purging (Blanckaert et al. 2020; Satokangas et al. 2020; Heidbreder et al. 2026). Consequently, the fitness effects of hybridization can change substantially across generations as recombination and selection reshape the frequencies of incompatible and beneficial genetic combinations. Although this process may facilitate fitness recovery and the establishment of persistent hybrid populations, its outcome depends on the genetic architecture of incompatibilities and the evolutionary conditions under which sorting and hybrid genome stabilization occurs (Pereira et al. 2021; Schneemann & Welch 2025; Kuijt et al. 2026).

Potential for heterosis declines with parental divergence but persists across a broad divergence range

We tested how increasing parental divergence impacts the potential for heterosis in hybrids. Previous meta- and comparative analyses have surveyed specific taxonomic groups (e.g. yeast; (Bernardes et al. 2017), animals; (Muraro et al. 2022)). We surveyed the literature for taxa pairs with genome-wide d_{xy} estimates and estimates of hybrid and parental fitness to test this. These measures were obtained for 13 species pairs and their hybrids, with multiple generations or multiple fitness measures for each species pair. The species represented in the data included birds (2), insects (4), arachnids (1), crustaceans (2), and plants (4). Our analysis covers taxa pairs ranging from d_{xy} of 0.002 to 0.026 divergence. We find that mid-parent heterosis decreased with increasing parental divergence (Fig. 2; LMM: $z = -3.26$, $P = 0.001$). The association remained significant when hybrid generation was included as a covariate (LMM: $z = -3.32$, $P < 0.001$), and heterosis did not differ between F_1 and later-generation hybrids ($z = -0.70$, $P = 0.483$).

Our results of declining heterosis with parental divergence are consistent with increasing impact of incompatibilities as divergence accumulates. Yet, there is considerable variation in the extent of heterosis and it persists until d_{xy} of 0.024. Our dataset covers parental d_{xy} from 0.002 to 0.026, encompassing divergence levels associated with the grey zone of speciation (0.5–2% net divergence, D_a; Roux et al. 2016) and reaches its upper limit. However, direct comparisons should be made with caution, as D_a accounts for within-population diversity and is therefore generally lower than d_{xy}. Further comparisons involving population pairs with greater parental divergence would be valuable for understanding the relationship between heterosis and parental genetic divergence across a broader range of evolutionary divergence, particularly beyond the grey zone of speciation. For example, heterosis has been reported in yeast at genetic distances exceeding 0.10 (Bernardes et al. 2017). Although our analysis did not find a difference in heterosis between early generation and later generation hybrids, this needs further systematic study. Furthermore, studies on yeast have demonstrated the impact of the environment on hybrid fitness (Bernardes et al. 2017), which also needs to be investigated systematically in future studies (see BOX 1 for other factors impacting hybrid fitness).

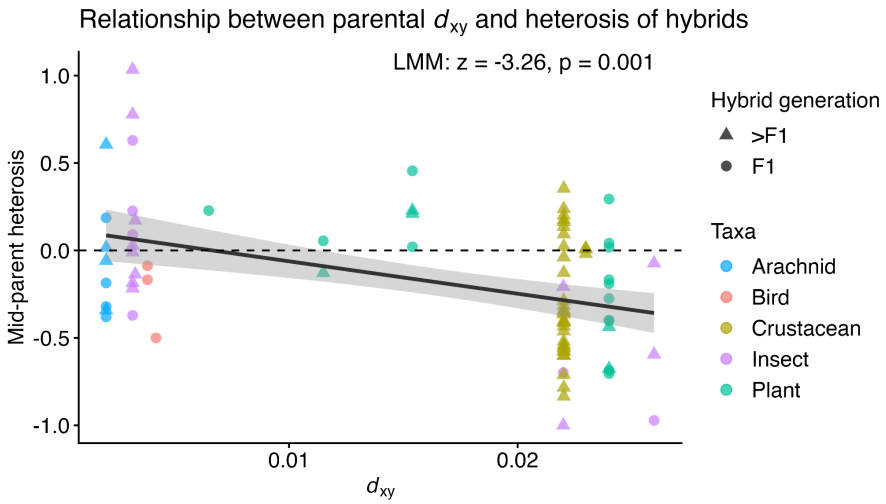


Figure 2. Relationship between genome-wide absolute genetic divergence (d_{xy}) between parental lineages and heterosis in hybrids. The hybrid generations include F_1 and later-generation hybrids (>F_1: BC_1, F_2 and later). Heterosis values were calculated using the fitness values of both parental and hybrid lineages. Points above the dashed line indicate the presence of heterosis. Generally, higher genetic divergence is associated with reduced heterosis in hybrids.

BOX 1. Factors impacting hybrid fitness and genome stabilization.

In this review we focus on parental divergence as a driver of hybrid fitness and genome stabilization. Other potential factors impacting the efficacy of selection on incompatibilities and genome stabilization are mentioned below. We do not discuss the impact of recombination rate variation here as it is extensively discussed in other sections.

Ploidy. Increased ploidy can mask recessive incompatibilities and deleterious alleles, restore fertility in otherwise sterile hybrids, and alter recombination and segregation, with direct consequences for genome stabilization. The level of parental divergence can affect the probability of homoploid hybrid speciation, and Chapman & Burke (2007) found that genetic divergence between the parents of homoploid hybrid species was approximately half that of polyploid hybrids. On the other hand, haploid life stages and haplodiploidy (Nouhaud et al. 2020) may speed up genome stabilization by favoring rapid purging of recessive incompatibilities, as these are unmasked in haploid individuals and visible to selection.

Environment. The fitness consequences of hybridization depend on environmental context, which can alter the effects of both deleterious and beneficial allelic combinations (Arnold et al. 2012; Schneemann et al. 2020; Brice et al. 2021; Thompson et al. 2022; Thompson et al. 2024; Krapf et al. 2026; Poikela et al. 2026). Allelic combinations that are deleterious in parental environments may be neutral or even advantageous under different environmental conditions. Theoretical models incorporating polygenic adaptation suggest that later-generation hybrids can achieve higher fitness than their parental lineages in novel environments, even when early-generation hybrids experience fitness costs (Kulmuni et al. 2024). This outcome can arise under different histories of divergent and parallel adaptation, although its likelihood depends on the genetic architecture of adaptation and the environmental conditions considered (Kulmuni et al. 2024).

During genome stabilization, a skew in retention of parental alleles could arise due to environmental selection. If the hybrid lineage inhabits an ecological niche more similar to one of the parent species, hybrids may retain more ancestry from this parent (Heil et al. 2019), especially if adaptation to the environment is polygenic and genome-wide. Hybrids may be able to exploit a niche that either parent cannot, which may favor retention of novel combinations of parental alleles if they convey adaptation to the novel environment. Although environmental effects are not the primary focus of this review, they are increasingly relevant as climate change and other anthropogenic pressures alter the conditions under which hybridization occurs, and hybrid fitness is expressed.

Population size. In small hybrid populations, drift can overwhelm selection, allowing weak incompatibilities to persist or beneficial combinations to be lost. Large population sizes should make selection on hybrid ancestry and removal of incompatible allele combinations more efficient.

Load of deleterious alleles in parental genomes. An alternative driver of asymmetric ancestry retention in hybrids would be that the major parent ancestor is the species with the lowest load of deleterious alleles (Bierne et al. 2002; Harris & Nielsen 2016; Juric et al. 2016; Schumer et al. 2018). If one parental species carries a greater genetic load, selection may preferentially remove its deleterious alleles from hybrid genomes, potentially reducing its overall ancestry contribution. Related patterns have been described in allopolyploids, where subgenome dominance can result in gene expression being biased towards the genome with the lowest number of accumulated deleterious variants. This pattern is consistent with genetic load in parental genomes shaping allopolyploid hybrid transcription profiles, although other mechanisms, including differences in regulatory architecture and transposable element content, may also play important roles.

Mating system. Selfing versus outcrossing affects how quickly recessive incompatibilities and deleterious alleles become homozygous and exposed to selection, and therefore how quickly they can be purged. Likewise, clonality can preserve successful hybrid genotypes without recombination breaking them apart (in yeast for example). Differences in mating systems could contribute to differences in the extent of introgression and genome stabilization among taxa.

Demography. The demographic and geographical context of hybridization can strongly impact the persistence and sorting of hybrid incompatibilities and consequently, hybrid fitness and genome stabilization. Continued gene flow from parental populations into the hybrid population can repeatedly introduce parental alleles that recreate incompatible combinations, slowing down the purging process. Hybrid zones represent a classic example of this tension between gene flow and selection against incompatible combinations (Endler 1977; Barton & Hewitt 1985; Harrison & Larson 2014). Repeated backcrossing with either parental lineage can further shift genome-wide ancestry towards that parent, again modifying the genomic backgrounds in which incompatibilities are expressed, selected against and purged. Over successive generations, this process may reduce the genomic contribution of one parent while retaining particular alleles that confer a fitness advantage, potentially facilitating adaptive introgression.

The impact of parental divergence on hybrid genome stabilization

Hybrid genomes that have persisted through many generations of recombination and selection provide valuable insights into the evolutionary forces shaping ancestry retention. Following the sorting of incompatible and beneficial allele combinations over tens to hundreds or more generations, these genomes may exhibit patterns of parental ancestry that reflect the cumulative effects of selection (see BOX 2). In particular, the genomic distribution and relative contributions of parental ancestry can provide insights into the extent to which genetic incompatibilities constrain hybrid genome evolution. Here we focus primarily on the role of incompatibilities and their predicted relationship with parental divergence.

At a genome-wide scale, patterns of sorting of parental alleles in hybrids provide insights into the frequency at which incompatible combinations arise. Hybridization between closely related lineages is expected to involve relatively few incompatible allele combinations (Orr 1995; Orr & Turelli 2001; Matute et al. 2010; Satokangas et al. 2020), potentially allowing substantial contributions from both parental genomes. As parental divergence increases, however, the accumulation of genetic incompatibilities may impose stronger constraints on the retention of introgressed ancestry. Consequently, we predict that hybrid genomes derived from more divergent parental lineages will exhibit lower proportions of minor-parent ancestry, defined as ancestry originating from the parent contributing the smaller fraction of the hybrid genome (Runemark et al. 2019; Moran et al. 2021). This prediction assumes that selection against

incompatible combinations ultimately produces asymmetric ancestry retention, with one parental genome becoming predominant. However, the extent of such asymmetry may also depend on factors other than overall parental divergence, including cytonuclear interactions, sex chromosome incompatibilities, recombination, and demographic history.

To evaluate these predictions, we conducted comparative analyses of published genomic data, including reanalyses of existing datasets, addressing three questions: 1) Does increasing parental divergence predict a reduction in genome-wide minor-parent ancestry? 2) Is the parent contributing the mitochondrial genome more likely to contribute the majority of nuclear ancestry, as predicted by hypotheses of mito-nuclear compatibility? 3) Do sex chromosomes exhibit reduced minor-parent ancestry relative to autosomes, consistent with their proposed disproportionate contribution to reproductive isolation?

BOX 2: Genome stabilization.

Here we define genome stabilization (Buerkle & Rieseberg 2008) as the result of selective removal of incompatible and fixation of beneficial alleles resulting in stable ancestry proportions within a hybrid lineage. These ancestry proportions can vary from even to highly skewed. One of the emerging patterns in hybridization research is that removal of minor parent ancestry varies across generations, with initial fast removal followed by slower purging. This was coined as a principle by Moran et al. (2021): "A combination of rapid and slower removal of foreign ancestry stabilizes admixed genomes". Variation in ancestry is expected to be highest at the early stages following hybridization, with subsequent recombination and selective sorting of haplotypes shaping hybrid genomes. Low recombination rate (Schumer et al. 2018) and a high number of incompatibilities (Veller et al. 2023) leads to faster sorting. Selective sorting of incompatibilities at early stage following hybridization is expected to remove large linkage blocks (Veller et al. 2023), leading to rapid changes in ancestry proportions. As recombination breaks up linkage blocks, the ancestry blocks removed by selection become smaller over time. Moreover, as strong incompatibilities are removed by selection, the effect sizes of the incompatible alleles and thus the selection for purging is reduced. The ancestry proportion in the hybrid genomes thus becomes more stable over time and eventually stabilizes when there no longer is parent-of-origin specific selection on alleles. By analogy with the concept of genetic load (Charlesworth et al. 1990), genome stabilization can be considered complete when the genome of a hybrid lineage no longer harbors any incompatible combinations of parental alleles that result in lower fitness. The proportion of minor parent ancestry is one outcome of the genome stabilization process, and fixation of compatible and beneficial parent of origin alleles is another.

What can minor parent ancestry tell us about incompatibilities?

Identifying the genomic regions and allelic interactions responsible for reduced hybrid fitness has been a long-standing goal of hybridization research (Barton 2001). The distribution of parental ancestry in hybrid genomes can provide important clues about the genomic regions giving rise to incompatibilities. However, directly identifying incompatible epistatic combinations often requires large sample sizes or systematic screening of allelic combinations in controlled genetic backgrounds (Puchta et al. 2016), approaches that remain empirically intractable in most study systems. Established hybrid populations and species offer an alternative perspective, as their genomes reflect the cumulative effects of selection and recombination over successive generations and may provide some insights into genomic regions involved in incompatibilities. For example, classic studies in *Xipophorous* swordtail fish (Schumer et al. 2018) and *Heliconius* butterflies (Martin et al. 2019) have revealed reduced minor-parent ancestry in regions of low

recombination, consistent with the accumulation of multiple incompatibilities in regions of low recombination. However, the correlation between recombination and introgression does not hold at low levels of parental divergence based on *Drosophila* (Dagilis & Matute 2023) and *Formica* wood ants (Satokangas et al. 2026), consistent with parental divergence playing a role in hybrid genome evolution.

A major challenge is distinguishing if the fixation of specific parental ancestry at a locus is driven by selection or instead by demographic and stochastic processes. Comparisons among independently formed hybrid populations from the same species pair, can potentially help identify genomic regions showing repeated ancestry biases, potentially indicating predictable effects of selection (Runemark et al. 2018; Langdon et al. 2022; Nouhaud et al. 2022). To date, such studies have been performed for *Formica* wood ants (Nouhaud et al. 2022) where variants from the parent species with lowest effective population size were disproportionately purged, swordtail fish (Langdon et al. 2022), and *Passer* sparrows (Runemark et al. 2018) where mitonuclear loci were preferentially inherited from the parent species contributing the mitochondrion. Another promising approach is experimental evolution as it allows changes in parental ancestry to be followed across successive hybrid generations under controlled conditions. Studies in *Saccharomyces* yeast hybrids have revealed environment-dependent retention of parental alleles (Martínez & Lang 2026) and parent-of-origin biases in chromosome loss (Antunovics et al. 2005). Experimental crosses in *Drosophila* have demonstrated predictable changes in parental genome contributions over generations (Matute et al. 2020; Zhang et al. 2020) and reduced retention of variants from insular parent species (Matute et al. 2020). However, repeatable patterns of purging and fixation could also arise if one of the parental alleles is weakly deleterious or universally beneficial (Juric et al. 2016). To confirm the role of incompatibilities in fixation of specific parental alleles, additional approaches are needed (see Moran et al. 2024). Recent progress has been made in using large samples of haplotypes or phased genomes to detect BDMIs through the underrepresentation of particular two-locus genotype combinations in natural hybrid populations of swordtail fish, rice and wood ants (Li et al. 2022; Li & Lee 2023; Heidbreder et al. 2026). Although such patterns can identify candidate incompatibilities, distinguishing epistatic selection from other causes of ancestry disequilibrium remains challenging (Szabo & Cutter 2024).

A promising venue is analyzing patterns across more than one cross, as has been done in swordtail fish, identifying similarities in ancestry composition across two different hybridizing species pairs (Langdon et al. 2024). Extending comparative work across species pairs spanning different levels of genetic divergence would allow us to test whether minor-parent ancestry becomes increasingly constrained as divergence accumulates. Such analyses can also help establish whether the genomic consequences of hybridization become more predictable with increasing parental divergence as has been observed in swordtail fish (Langdon et al. 2024). Furthermore, comparative analyses can reveal if the pattern of minor parent ancestry retention reflects the expected “snowball” pattern of accumulation of incompatibilities (Orr 1995; Orr & Turelli 2001; Matute et al. 2010).

To investigate whether parental genetic divergence predicts genome-wide ancestry composition in hybrids, we compiled published estimates of absolute genetic divergence (d_{xy}) between parental species and minor parent ancestry in their hybrids from 19 species pairs (birds: 5, fish: 5, insects: 4, plants: 4, mammals: 1), with one to eight independent hybrid populations or hybrid zones per species pair (see

Materials and Methods). Minor parent ancestry decreased with increasing parental divergence as we predicted (Fig. 3A; Linear mixed model (LMM): z = −3.76, P < 0.001). In other words, hybrids between more divergent parents retained less ancestry from the minor parent. The association remained significant when we excluded the systems with d_{xy} > 0.01 (Fig. 3B; LMM: z = −2.60, P = 0.009) and when we included hybrid type (hybrid lineage or hybrid zone) as a covariate (LMM: z = -3.55, P < 0.001), whereas hybrid type itself had no significant effect on minor parent ancestry (z = −0.23, P = 0.818). These findings suggest that parental divergence is an important predictor of ancestry asymmetry across hybrid systems, consistent with selection against incompatible genetic combinations becoming stronger as divergence increases. However, ancestry proportions are also influenced by demographic history, patterns of gene flow, variation in recombination rate, and environmental conditions, which are not accounted for in our analysis (see BOX 1 and discussion below).

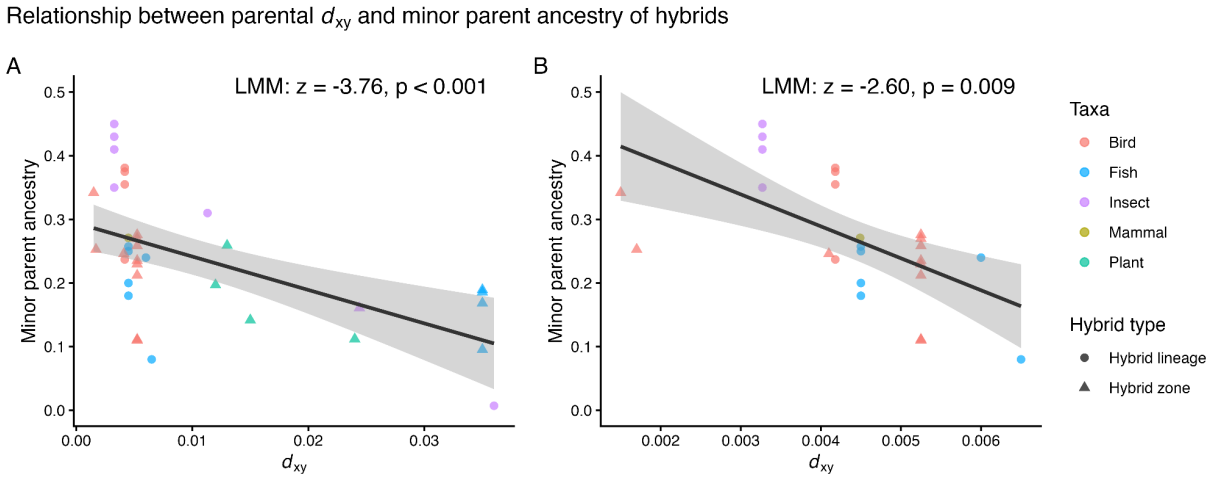


Figure 3. The effect of absolute genetic divergence (d_{xy}) between parental lineages on minor parent ancestry in hybrids A) across all data points and B) across data points with d_{xy} less than 0.01. The higher the divergence, the smaller the proportion of minor parent ancestry.

Other drivers of ancestry variation in hybrids

Although parental genetic divergence may predict genome-wide ancestry proportions, other factors can also influence ancestry retention in hybrid genomes. Here we focus on cytonuclear interactions (Runemark et al. 2018; Moran et al. 2024) and sex chromosome divergence. Mitochondrial function depends on nuclear-encoded proteins, creating opportunities for co-evolution of mitochondrial and nuclear genomes. Hybridization can disrupt these coadapted interactions, potentially favoring the retention of nuclear alleles from the parent contributing the mitochondrial genome. Such ancestry biases are expected particularly at nuclear loci involved in mitochondrial function, although selection for mito-nuclear compatibility could also influence broader patterns of genomic ancestry (Hill 2017). Several studies have identified ancestry patterns consistent with selection against mito-nuclear incompatibilities in hybrid populations or species, including swordtail fish and Italian sparrows (Trier et al. 2014; Runemark et al. 2018; Healy & Burton 2020; Hirase et al. 2021; Moran et al. 2024).

Sex chromosomes may also play a disproportionate role in shaping ancestry retention in hybrid genomes. An open question is whether parental sex chromosome divergence predicts the proportion of minor

parent ancestry more strongly than overall parental divergence, as could be expected if their contribution to epistatic interactions underlying reproductive isolation is disproportionate. Comparing minor-parent ancestry on the sex chromosomes compared to the autosomes can also shed light on whether incompatibilities are more frequent on the sex chromosomes compared to autosomes than expected based on their effective population sizes, as postulated by Haldane's rule (Haldane 2022) and the faster X hypothesis (Rice 1984; Charlesworth et al. 1987; Meisel & Connallon 2013). Haldane's rule describes the tendency for the heterogametic sex to exhibit greater hybrid sterility or inviability, while faster-X evolution may contribute to the accumulation of sex-linked incompatibilities through the increased exposure of recessive alleles to selection in hemizygous individuals. Comparisons of the distribution of minor parent ancestry among species with different sex determination systems could also provide insights into the role of drift and various selection pressures in driving divergence. The pattern of faster Z is thought to arise mainly from drift (Mank et al. 2007; Mank et al. 2009) which contrasts with faster X which may be driven by faster adaptation of hemizygous genes (Saunders & Muyle 2024) and/or sexual selection (Koevoets & Beukeboom 2009). Moreover, by documenting the extent to which introgression of minor parent variants in specific genes on homogametic/recombining sex chromosomes or chromosomes are preferentially purged, it is possible to contribute to the understanding of the basis for Haldane's rule (Haldane 2022). Together, comparative analyses of sex chromosome ancestry may help clarify whether sex-linked incompatibilities contribute disproportionately to the genomic barriers that shape hybrid ancestry retention.

Here, we conducted comparative analyses across published hybrid systems to investigate the role of mtDNA in genome stabilization and driver of major parent ancestry in hybrid genomes. Across established hybrid lineages, the parent that contributed the mtDNA (the mitoparent) also contributed on average 64% of the nuclear genome (95% CI 54–74%; N = 15 hybrid populations from 7 taxon pairs), significantly more than the 50% expected if mtDNA origin were independent of nuclear ancestry (Fig. 4A; one-sided one-sample t-test: $t = 3.00$, $df = 14$, $P = 0.005$). In 11 of the 15 populations, mtDNA derived from the major parent. All four exceptions came from two systems, *Formica* wood ants and *Passer* sparrows, in which independent hybrid populations differed in their mitoparent (two populations with major-parent and two with minor-parent mtDNA in each system).

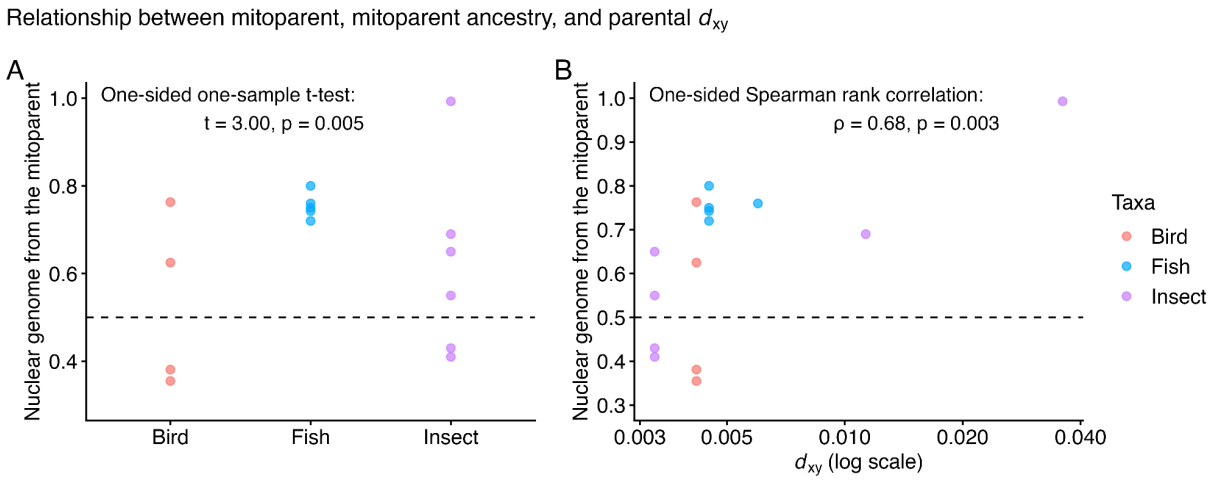


Figure 4. Relationship between mitochondrial parent (=mitoparent; the parental species that contributed the mtDNA), the nuclear genome from the mitoparent in hybrids, and absolute genetic divergence (d_{xy}) between parental lineages. (A) The influence of the mitoparent on the proportion of the nuclear genome derived from the mitoparent in hybrids. (B) The impact of d_{xy} (displayed on a log scale for clarity) between parental lineages on the proportion of the nuclear genome derived from the mitoparent in hybrids.

The proportion of nuclear ancestry from the mitoparent increased with parental divergence (Fig. 4B; one-sided Spearman rank correlation, ρ = 0.68, P = 0.003). To test whether this correlation was driven solely by the two most divergent systems (*Heliconius* and *Aricia*, d_{xy} > 0.01), we repeated the analysis without them, and the correlation remained significant (ρ = 0.63, P = 0.010). This pattern is consistent with cytonuclear incompatibilities, which accumulate as parental species diverge and favor combining the mtDNA of one parent with a nuclear genome derived mainly from that same parent. However, the remaining systems span a narrow range of divergence and only five taxon pairs. This relationship should be confirmed with more hybrid lineages, particularly from more divergent parental species.

Finally, we investigated if sex chromosomes carry less minor parent ancestry compared to autosomes and how this relationship depends on parental divergence. As consistent and comparable measures of genomic divergence and introgression between parent species of hybrid species proved to be difficult to find due to inconsistent reporting, we reanalyzed genomic data from previous studies, focusing on taxa that the authors have described as homoploid hybrids. We estimated the proportion of minor parent ancestry separately for the autosomes and the sex chromosome of each system (see Materials and Methods). In four of the six ZW hybrid populations, the Z carried clearly less minor parent ancestry than the autosomes (Fig. 5). In *Heliconius elevatus* the proportion fell from 0.17 on the autosomes to 0.04 on the Z, and in the *Aricia* hybrid from 0.32 to 0.12. In the Cretan and Corsican Italian sparrows, it fell from 0.33 to 0.18 and from 0.44 to 0.32. The Sicilian Italian sparrow showed no detectable difference (0.35 and 0.36), although its confidence interval was wide. In *Heliconius heurippa*, the autosomes were equally close to both parents. Its Z carried 0.25 ancestry from *Heliconius melpomene*, the parent described as the minor contributor in the literature. The XY hybrids did not show consistent reduction in minor parent ancestry on the sex chromosome. The X of *Xiphophorus nezahualcoyotl* carried slightly less minor parent ancestry than its autosomes (0.37 and 0.41), but this difference was not robust to larger bootstrap blocks. In *Ursus thibetanus*, the ancestry of X differed little from its autosomes (0.51 and 0.50). In *Vicugna pacos*, the

autosomes were closer to *Lama glama*, however, X carried mostly *Vicugna vicugna* ancestry (0.77, against 0.44 on the autosomes). The ratio of sex chromosome to autosomal minor parent ancestry showed no clear relationship with parental divergence (Fig. 5A). Although some systems contradict our prediction by showing greater minor-parent ancestry on the sex chromosomes than on the autosomes, these systems tend to result from parental species with a relatively low divergence.

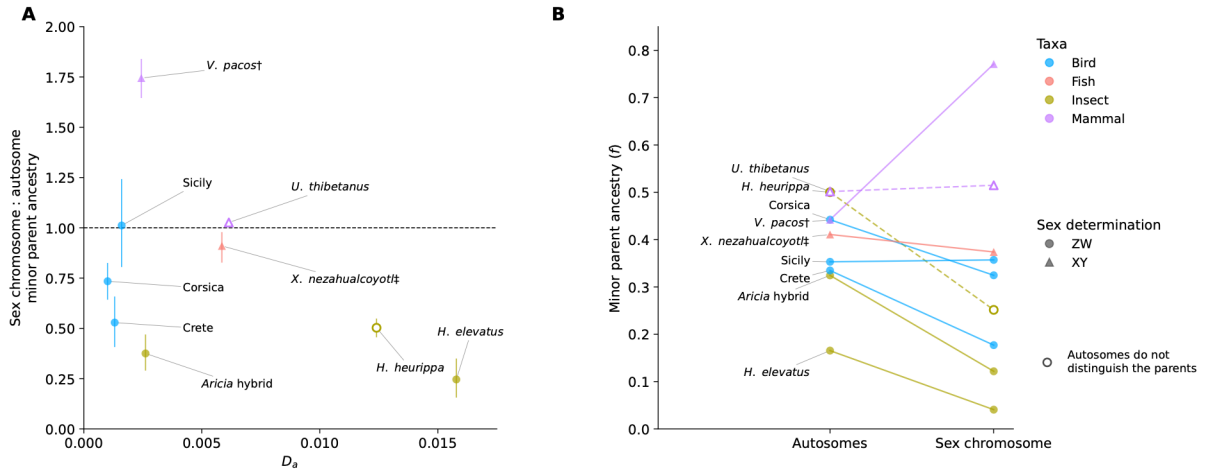


Figure 5. Minor parent ancestry on the sex chromosomes compared to autosomes. (A) Ratio of minor parent ancestry on the sex chromosome to minor parent ancestry on the autosomes as a function of autosomal parental divergence (D_a). The minor parent is the parent contributing less than half of the autosomal ancestry. The dashed line marks equal minor parent ancestry, and values below one indicate proportionally less minor-parent ancestry on the sex chromosome compared to autosomes. Vertical bars are 95% confidence intervals from a bootstrap of 1 Mb genomic blocks. Corsica, Crete and Sicily are populations of the Italian sparrow. They share the same parents and are offset horizontally so that their intervals can be seen. (B) Proportion of minor parent ancestry (*f*) on the autosomes and on the sex chromosome of each hybrid lineage. Color denotes taxonomic group and shape shows sex-determination system. Hollow symbols and dashed lines mark *H. heurippa* and *U. thibetanus* for which minor parent is inferred from the literature.

These comparisons have three limitations. Sex-determination system and taxonomic group are confounded in our sample, because all ZW hybrids are birds or insects and all XY hybrids are fish or mammals. The three sparrow populations share their parental species, and the two *Heliconius* hybrid species both have *H. melpomene* as parent species, implying that these comparisons are not independent. Finally, divergence that accumulates in a hybrid lineage after admixture can make the Z of a ZW hybrid appear to carry less minor parent ancestry than it does, and we could not estimate the size of this bias (see Materials and Methods). We therefore describe each system separately and do not test for an effect of sex-determination system.

Challenges and opportunities for comparative analyses of hybrid fitness and genome evolution

Comparative analyses on the impact of parental divergence on hybrid fitness and genome composition can shed light on the evolutionary consequences of hybridization. However, robust inference requires sufficiently large and representative data sets, and the availability of comparable data remains a major limitation. Although several studies have quantified the fitness of early-generation hybrids, differences in

experimental design, fitness measures, and data reporting complicate comparisons across studies, as highlighted also by other comparative frameworks (Stankowski et al. 2024; Walker et al. 2026) and studies on introgression (Dagilis et al. 2022). In particular, access to estimates of both hybrid and parental fitness, ideally accompanied by raw fitness data, is essential for evaluating heterosis across systems. Furthermore, our analysis on the impact of parental divergence on heterosis was constrained by the limited availability of comparable estimates of parental genetic divergence (d_{xy}). Even studies employing whole-genome sequencing frequently did not report mean absolute nucleotide divergence (d_{xy}), preventing their inclusion in our analyses. Future comparative studies would therefore benefit from more consistent reporting of hybrid and parental fitness as well as parental divergence on autosomes and sex chromosomes (preferably measured in multiple ways, including d_{xy}, F_{st} and D_a).

To address outstanding questions on ancestry proportions in hybrids and distribution of introgression within the genome across a range of parental divergence values (Table 1), detailed ancestry inference is required. Inference of genomic ancestry in hybrid genomes face multiple methodological challenges (see also Dagilis et al. 2022). Differences in genome assembly and annotation quality can affect the proportion of the genome accessible to ancestry inference, potentially introducing systematic biases across taxa. Reference-mapping bias (Lunter & Goodson 2011) may further distort ancestry estimates when sequencing data from divergent parental backgrounds are aligned to a single reference genome. Similarly, the divergence of available outgroup and sampled lineages can influence the reliability of site identification and allele-sharing statistics, including those used in analyses based on D-statistic (Dagilis et al. 2022; Tricou et al. 2022; Koppetsch et al. 2024). Fortunately, large-scale genome sequencing initiatives are currently paving the way for producing reference genomes using standardized pipelines (e.g. The Darwin Tree of Life Project Consortium 2022) and pangenomic approaches (Garrison et al. 2018) hold great promise to overcome reference bias based challenges. An additional complication arises when the sampled parental populations are imperfect proxies for the ancestral lineages that originally contributed to a hybrid genome. This can bias estimates of parental divergence and ancestry, particularly in ancient hybrid lineages. Advances in methods for detecting historical introgression and reconstructing ancestral relationships, including approaches based on ancestral recombination graphs (Brandt et al. 2024; Wong et al. 2024; Nielsen et al. 2025), may help address some of these limitations (Jiang et al. 2020). Further improvements are gained with reference-free ancestry inferences (Baird et al. 2023; Setter et al. 2026).

Despite these challenges, the development of standardized, openly accessible databases using transparent reporting hold great promise for comparative hybridization research. We advocate for making key parameters available in a comparable format, in accordance with FAIR principles (Wilkinson et al. 2016). These should include information on sampling design, hybrid generation, sex determination system, ploidy, reference genome quality, ancestry inference methods, mitochondrial and chloroplast ancestry, introgression estimates for autosomes and sex chromosomes to further improve the interpretation and comparability of published estimates. Community efforts to establish standardized terminology and reporting frameworks comparable to the Reproductive Isolation Ontology framework (Walker et al. 2026; https://speciation-network.pages.ist.ac.at/resources/) are important in facilitating the integration of data across studies. A shared database of hybrid fitness and genomic ancestry could enable more comprehensive tests of the factors shaping hybrid genome evolution, including the relative contributions of parental divergence, cytonuclear interactions, and sex chromosome evolution.

Continued advances in genome assembly, long-read sequencing, and pangenomic approaches will also make it possible to conduct comparative analyses on aspects of hybrid genome evolution that remain poorly understood (Table 1). For instance, comparing the retention of minor-parent ancestry across coding and regulatory regions, structural variants, and other genomic features could help identify which classes of variation contribute disproportionately to reproductive isolation. Stronger purging of minor parent ancestry in genomic regions with higher density of functional elements has been reported (e.g. Sankararaman et al. 2016; Langdon et al. 2024), with reduced introgression at regulatory sites in humans (Telis et al. 2020). In contrast, meta-analysis revealed increased introgression of MHC alleles across hybrid zones (Gaczorek et al. 2024). Similarly, examining the parental origins of centromeric regions across chromosomes and independently formed hybrid lineages could reveal whether centromeres from one parental species are preferentially retained, potentially providing insights into chromosome segregation and genome stability.

Persistence of hybrid lineages over long evolutionary time scales

Hybridization has contributed to the evolutionary histories of lineages across the tree of life, with genomic signatures of ancient admixture persisting over long evolutionary timescales. However, the persistence of introgressed ancestry does not necessarily imply the long-term survival of an independently evolving hybrid lineage. While some hybrid populations establish distinct evolutionary trajectories and contribute to biodiversity, others disappear through extinction or repeated backcrossing with parental populations. What determines whether hybrid lineages persist over thousands of generations, and how might parental divergence influence their long-term evolutionary success?

One important factor contributing to persistence of hybrid lineages is reproductive isolation from the parental lineages (Mallet 2007). Such isolation may arise through genetic incompatibilities (Schumer et al. 2015), ecological differentiation (Moore 1977; Arnold 1997; Rieseberg et al. 2003), or geographical barriers (Hermansen et al. 2011). If reproductive isolation is lacking, continued gene flow from parental populations can erode the genetic distinctiveness of hybrid lineages through repeated backcrossing, potentially leading to genetic swamping or the reintroduction of incompatible allele combinations. Comeault & Matute (2018) used experimental evolution in *Drosophila* to show that both parental divergence and the number of hybridizing species affect the probability that hybrids evolve reproductive isolation from their parental species. Hybrid populations derived from parental species with intermediate levels of divergence were more likely to evolve assortative mating and discrimination against their parental species than those generated from weakly or strongly diverged parental species. Furthermore, hybrid populations derived from three parental species were more likely to evolve reproductive isolation than those derived from two species. These findings suggest that hybridization can generate genetic combinations that facilitate the evolution of reproductive barriers, although the likelihood of this outcome depends on the divergence and composition of the parental lineages.

Furthermore, for a lineage to persist over evolutionary time scales, it needs to be able to survive environmental fluctuations and adapt to changing conditions. The general consensus is that adaptation requires genetic variation (although the amount of genetic variation does not always correlate with adaptive potential; see Abson et al. 2026). Hybridization increases the amount of genetic variation available for selection to act on (Buerkle et al. 2000), and hybridization between more divergent parental

603 populations generates more phenotypic diversity and novelty. If hybrid lineages are able to maintain
604 higher genetic variation compared to parental lineages, they may be able to persist across changing
605 environmental conditions and even outcompete parental lineages over longer evolutionary time scales.
606 However, the process of hybrid genome stabilization removes incompatibilities and fixes beneficial allele
607 combinations, potentially removing much of the initial genetic variation segregating in hybrid populations
608 (Kulmuni et al. 2024). To what extent hybrid lineages can maintain their increased genetic variation, is
609 unclear and will depend on several factors including the genomic architecture of incompatibilities, their
610 strength, and the genomic landscape of recombination. For example, purging of incompatibilities located
611 in low recombining regions will remove larger blocks of ancestry compared to purging in high recombining
612 regions (Sachdeva & Barton 2018a; Sachdeva & Barton 2018b) as discussed in BOX 2. Thus, purging of
613 incompatibilities in low recombining regions removes potentially more neutral and beneficial genetic
614 variation than incompatibilities located in high recombining regions. In contrast, recombination in other
615 genomic regions may facilitate the removal of incompatible combinations while preserving advantageous
616 variation from both parental lineages.

617 The long-term consequences of these processes are also likely to depend on environmental variability.
618 The rate of environmental change will impact the rate at which both hybrid and parental populations can
619 maintain and accumulate new genetic variation, and the evolution of adaptive responses (Ketola &
620 Kronholm 2023). Rapid or unpredictable environmental shifts may challenge the ability of populations to
621 adapt, potentially increasing extinction risk (Botero et al. 2015). Hybrid lineages could benefit under such
622 conditions if they retain genetic variation or novel allele combinations that provide adaptive advantages
623 over their parental lineages. However, whether hybridization consistently enhances resilience to
624 environmental change remains unclear. An important question for future research is therefore whether
625 the genetic variation retained following hybrid genome stabilization contributes to long-term adaptive
626 potential, and whether this relationship depends on the evolutionary divergence between parental
627 lineages.

Conclusions

Hybridization can generate evolutionary novelty and facilitate adaptation, but it can also produce genetic incompatibilities that reduce fitness and threaten population persistence. Here, we combined a synthesis of existing research with comparative analyses across eukaryotic taxa to assess how parental genetic divergence relates to hybrid fitness and genomic ancestry. Our analyses revealed three general patterns: i) Heterosis tends to decline with increasing parental divergence, ii) the proportion of minor parent ancestry in hybrid genomes decreases with parental divergence, consistent with increasing constraints imposed by incompatibilities, and iii) mitochondrial parent tends to contribute the majority of ancestry in hybrid genomes, the relationship growing stronger with higher parental divergence. Our results regarding minor parent ancestry in sex chromosomes are inconclusive. Our analyses should be extended with data from species pairs with higher divergence.

The genomic consequences of hybridization are also shaped by processes that cannot be explained by parental divergence alone. Ploidy, mating system, sizes of hybridizing populations, demography and environment can all play a role in fixation and retention of parental ancestry in hybrid genomes (see BOX 1). Our analyses also highlight the limitations of current comparative datasets, particularly the small number of independent hybrid systems and inconsistencies in the reporting of parental divergence, fitness, and genomic ancestry. More standardized and accessible data will be essential for testing the generality of these relationships and distinguishing the contributions of different evolutionary mechanisms, as pointed out by previous studies (Dagilis et al. 2022; Stankowski et al. 2024; Walker et al. 2026). Ultimately, parental divergence provides a useful starting point for predicting hybridization outcomes, but understanding when hybridization generates evolutionary novelty, establishes persistent lineages, or contributes to reproductive isolation requires integrating divergence with the genetic, demographic, and ecological context in which hybrid genomes evolve.

Materials and methods

We investigated how parental divergence, measured as absolute nucleotide divergence (d_{xy}), influences heterosis, major and minor parent ancestry in hybrid genomes, ancestry proportions on autosomes and sex chromosomes, cytonuclear ancestry, and the influence of contrasting sex-determination systems.

Initial literature search and screening

We identified candidate studies through a literature search of Web of Science designed to capture studies of genetic divergence between hybridizing taxa. We first defined a set of thematic search blocks (hybrid, early generation, fitness, divergence, ancestry, architecture, cytonuclear, sex chromosome, expression, context, and environment), each consisting of keywords joined with the Boolean operator OR to maximize sensitivity. Blocks were then combined with AND to form six targeted searches: hybrid + fitness + divergence; hybrid + divergence + ancestry; hybrid + architecture; hybrid + cytonuclear; hybrid + sex chromosome; and hybrid + expression + speciation. Together these returned 9,910 unique records. Titles, abstracts, and journal names were exported and screened with a rule-based keyword classifier that matched each record against curated vocabularies for evolutionary ecology, human population genetics, crop and livestock breeding, medical/clinical research, forensics, and non-biological uses of "hybrid" (e.g. hybrid algorithms or materials). Each record was assigned to a domain and, within on-topic domains, to a relevance tier (high, medium, or low) according to the number of directly on-topic terms it contained (e.g. parental divergence, d_{xy}, hybrid incompatibility, Bateson-Dobzhansky-Muller incompatibility, Haldane's rule, hybrid sterility or breakdown, genome stabilization, transgressive segregation) relative to broader hybridization terms (e.g. introgression, admixture, gene flow, polyploidy, phylogeography). We retained papers classified as evolutionary ecology at any relevance tier, together with human population genetics papers with a hybridization context (n = 4,318) and restricted these to papers published from 2010 onwards (n = 3,183).

We then retained only papers with evidence of deposited sequence data, defined as an accession from an INSDC/NCBI sequence repository (SRA, ENA, DDBJ, GenBank, BioProject, or GEO/ArrayExpress) identified in the paper's full text or data-availability statement, or a link between the paper and a record in one of these repositories. Papers whose data were deposited only in general-purpose repositories (Dryad, Zenodo, Figshare) were not retained at this step. This left 513 papers, which were sorted by text-pattern matching on titles and abstracts into three hybrid-stage categories, applied in order: (i) hybrid zone or cline studies ("hybrid zone" or "cline"; n = 65); (ii) early-generation hybrids, including F1, F2, and backcross individuals and experimental crosses ("F1", "F2", "backcross", "early generation", or "experimental"; n = 64); and (iii) later-generation or established hybrids (all remaining papers; n = 384). A paper matching both of the first two categories was assigned to category (i). Finally, all 513 papers were manually screened for relevance before data extraction.

This initial screening revealed that we either had too strict filtering criteria, or that the initial search failed to capture some relevant studies. To complement the broad search, we performed targeted manual searches for relevant papers, especially to identify papers studying the effect on fitness for hybrid taxa.

Relationship between heterosis and absolute genetic divergence (d_{xy})

Proxies for fitness for both parental lineages and their hybrids were sourced from known hybridizing systems for which genome-wide mean d_{xy} between the parental lineages was also available (for one species pair we used D_a, which is calculated for a pair of samples instead of being a population-level measure; D_a is usually slightly lower than d_{xy}; Table S1). We included traditional fitness proxies like number of fertility or fecundity but also other traits relevant for survival like host plant acceptance and heat tolerance. In total, we obtained fitness estimates and d_{xy} for 12 species pairs and their hybrids: birds (2), insects (4), arachnids (1), crustaceans (2) and plants (4). Most species pairs contributed several data points, from different hybrid generations (F_1, F_2, BC_1, or later generations), different fitness proxies, or both. Some of the crosses were generated and studied in the laboratory, while others came from naturally hybridizing populations studied in the field or in the laboratory.

Mid-parent heterosis (MPH) was calculated as the relative difference between hybrid fitness and the mean fitness of the two parental lineages. Here, F_1 represents hybrid fitness, and P_1 and P_2 represent the fitness of the two parental lineages. MPH was calculated as:

$MPH = \frac{F_1 - MP}{MP}$, where MP is the mid-parent value, calculated as the average fitness of the two parental lineages:

$MP = \frac{P_1 + P_2}{2}$.

Positive values indicate that hybrids outperform the parental mean, and negative values indicate that they underperform it.

We tested whether MPH decreases with increasing parental d_{xy} using a linear mixed model (LMM) with a Gaussian error distribution. MPH was the response variable and parental d_{xy} the explanatory variable, with fitness measure nested within species pair as random effect to account for the non-independence of multiple data points from the same system. To test whether hybrid generation affected the results, we fitted a second model with hybrid generation as an additional fixed effect, contrasting F_1 hybrids with later-generation hybrids (BC_1, F_2 and later), in which heterozygosity is likely reduced and recessive incompatibilities can be exposed but also selected against. A Shapiro-Wilk test indicated a slight deviation of the model residuals from normality (W = 0.970, P = 0.047). However, simulation-based residual diagnostics (DHARMa; Hartig et al. 2026) showed no deviations from model assumptions (Kolmogorov-Smirnov test: P = 0.570; dispersion test: P = 0.360; outlier test: P = 1.000), and refitting the model with a Student-t error distribution, which is less sensitive to heavy tails, gave identical estimates. We therefore retained the Gaussian model. Models were fitted with the glmmTMB package (Brooks et al. 2017) in R v4.4.1 and RStudio v2024.04.2.

Relationship between ancestry proportions and absolute genetic divergence (d_{xy})

Genome-wide ancestry proportions of hybrid lineages or hybrid zones were collected from well-studied systems for which genome-wide mean d_{xy} between the parental lineages was also available (for one species pair we used D_a, which is calculated for a pair of samples rather than at the population level; Table S2). We use the term "hybrid lineage" to cover both stable hybrid populations and hybrid species. Ancestry proportions and d_{xy} were often taken from different publications, and where raw ancestry proportions

were not provided, we estimated them from published figures (marked in Table S2). In total, we obtained genome-wide ancestry proportions and parental d_{xy} for 19 species pairs and their hybrids, with 1–8 independent hybrid populations or zones per system. These data included species from birds (5), fish (5), insects (4), mammals (1), and plants (4).

We tested whether minor parent ancestry in hybrid genomes decreases with increasing parental d_{xy} using a linear mixed model (LMM) with a Gaussian error distribution. Minor parent ancestry was the response variable, parental d_{xy} the explanatory variable, and species pair a random effect, to account for the non-independence of multiple hybrid populations or zones from the same system. Unlike in the mtDNA analysis (see below), which required mtDNA to be fixed for one parent, minor parent ancestry can be measured in both hybrid lineages and hybrid zones, so both were included. To test whether hybrid type affected the results, we fitted a second model with hybrid type (hybrid lineage or hybrid zone) as an additional fixed effect. To ensure that the results were not disproportionately influenced by the small number of species pairs with high d_{xy} values (>0.01), we repeated the analysis excluding these pairs. Shapiro-Wilk test indicated a slight deviation of the model residuals from normality in the full model ($W = 0.93$, $P = 0.025$). However, simulation-based residual diagnostics (DHARMa; Hartig et al. 2026) showed no deviations from model assumptions (Kolmogorov-Smirnov test: $P = 0.463$; dispersion test: $P = 0.440$; outlier test: $P = 1.000$), and refitting the full model with a Student-t error distribution, which is less sensitive to heavy tails, gave identical estimates. Model residuals were approximately normally distributed in the reduced data set ($W = 0.93$, $P = 0.080$). Models were fitted with the glmmTMB package (Brooks et al. 2017) in R v4.4.1 and RStudio v2024.04.2.

Relationship between mitoparent, ancestry proportions and absolute genetic divergence (d_{xy})

We used the same dataset as in the ancestry proportion – d_{xy} analysis, but restricted it to hybrid lineages with available mtDNA information (Table S3). For each hybrid lineage, we identified the mitoparent, defined as the parental species that contributed the mtDNA, and calculated the proportion of the nuclear genome derived from that parent. If mtDNA tends to be inherited from the major parent, this proportion should exceed 0.5. Hybrid zone systems were excluded from this analysis because the mitoparent could not be determined. Additionally, we excluded one hybrid lineage (*Vicugna vicugna* × *Lama guanicoe*) in which both parental mtDNA haplotypes were still segregating (23/84 vicuña, 61/84 guanaco), so no single mitoparent could be assigned. A hybrid population (*Xiphophorus cortezi* × *Xiphophorus birchmanni*) in which nearly all mtDNA (99.7%) was derived from one parent was treated as fixed for that parent's mtDNA. This resulted in 15 hybrid populations from 7 taxa pairs, with 1–4 hybrid populations per taxa pair. Because populations within the same taxa pair can have different mitoparents, as in *Formica* and *Passer*, we treated each population as an independent observation.

We tested whether the proportion of nuclear ancestry from the mitoparent exceeded 0.5 with a one-sided one-sample t-test. To test whether mtDNA is more likely to derive from the major parent when the parental species are more divergent, we computed a one-sided Spearman rank correlation between the proportion of nuclear ancestry from the mitoparent and parental d_{xy}. All analyses were done in R (v4.4.1) and RStudio (v2024.04.2).

Relationship between sex chromosomes, major parent ancestry and parental species divergence

Published studies did not report minor parent ancestry separately for autosomes and sex chromosomes, and the divergence measures they report were not comparable across systems. We therefore reanalyzed raw sequence data from previous studies.

We included hybrid systems with short-read data for the hybrid, both parents and an outgroup species, and a high-quality assembly for at least one taxon. We restricted the set to taxa the original authors described as homoploid hybrids. For all systems except the Italian sparrow, we used a single individual per species with the highest coverage, preferring the homogametic sex. For the Italian sparrow, we used 9 to 21 individuals per population, with 15 house sparrows, 10 Spanish sparrows and 9 tree sparrows as the outgroup.

Reads were downloaded from the SRA, aligned with bwa-mem2 (Vasimuddin et al. 2019) to the reference genomes (Table S4), and were genotyped with bcftools v1.23.1 (Danecek et al. 2021). The reference was the major parent in four systems, the minor parent in three, the hybrid in one and a congener outside the trio in one, and we did not correct for this heterogeneity.

We calculated d_{xy} and nucleotide diversity (π) with pixy v2.2.3 (Korunes & Samuk 2021) in non-overlapping windows of 50 kb (10 kb for the sparrow), and f_{dM} with Dsuite v0.5 r58 (Malinsky et al. 2021) in windows of 50 biallelic SNPs with a 25 SNP overlap. For each pair of samples among the hybrid and its two parents, we used net divergence, $D_a = d_{xy} - (\pi 1 + \pi 2)/2$. Instead of averaging windows, we summed the counts of differences and comparisons reported by pixy across windows, separately for the autosomes and the X or Z. We kept windows in which all three pairs had at least 1000 sites. We assigned sequences of at least 5 Mb as well all assembled chromosomes in the *Passer* sparrows as autosomes, excluding Y and W chromosomes and the pseudoautosomal regions of the sparrow Z and the *Vicugna* X.

For a hybrid H with parents N and O, we estimated the proportion of ancestry from N with the differenced hybrid index, $f = ½[1 + (D_a(\text{H,O}) - D_a(\text{H,N})) / D_a(\text{N,O})]$, in which the diversity of the hybrid cancels. The minor parent, the species contributing the smaller part of the hybrid genome, was the parent with f below 0.5 on the autosomes. Whenever the confidence interval for autosomal ancestry inherited from the parental species included 0.5 (*H. heurippa* and *U. thibetanus*) we assigned the parent described as minor in literature the minor parent. We compared ancestry retention on sex chromosomes compared to autosomes as the ratio f_{sex} / f_{auto}, where one means equal minor parent ancestry. Confidence intervals (95%) came from 2000 bootstrap resamples of 1 Mb blocks, drawn separately for the autosomes and the sex chromosome.

Divergence that a hybrid lineage accumulates after admixture shrinks f towards 0.5, but to a lower extent in hybrid lineages resulting from crossings of more divergent parent species. In all ZW systems, the parents were 1.25 to 3.0 times more diverged on the Z than on the autosomes. This can make the Z appear to carry less minor parent ancestry than it does, but we could not estimate the magnitude of this potential bias.

For f_{dM} the minor parent was placed as P3, so positive values indicate excess allele sharing between the hybrid and its minor parent. The underlying data and summary statistics, including the numbers of sites, windows and genomic blocks used in each analysis, are provided in Table S5.

Sex chromosomes were genotyped as diploid. The *V. pacos* and *L. glama* samples are males, so their hemizygous X is called homozygous, but an approximate correction changed the *V. pacos* ratio only from 1.74 to 1.70. The *X. nezahualcoyotl* and *X. cortezi* samples are also males, but their X was about as heterozygous as their autosomes. Their X is therefore largely undifferentiated from the Y, so *f* on the X of *X. nezahualcoyotl* may confound X- and Y-linked ancestry and should be interpreted with caution. In the Italian sparrow, Z-linked f_{dM} exceeded autosomal f_{dM} regardless of which parent was assigned as P3 (Table S5). Because a shift in ancestry would be expected to increase f_{dM} in one orientation while decreasing it in the other, this consistent pattern may instead reflect the threefold greater parental divergence on the Z chromosome rather than differences in ancestry. We therefore refrain from interpreting the comparison of f_{dM} between chromosome classes in the sparrow.

Acknowledgements

This review was funded by the Research Council of Finland (Suomen Akatemia) grants 324485 and 346805, awarded to J.K.

Data availability statement

The data underlying the figures and analyses are provided in the Supplementary Material. The code used to perform the statistical analyses and generate the figures will be made publicly available on Zenodo upon publication.

References

Abson KL et al. 2026. Nucleotide diversity is a poor predictor of short-term adaptive potential. *Proc. Natl. Acad. Sci. U.S.A.* 123:e2536181123.

Aguilar-Gómez D et al. 2025. Genetic rescue of Florida panthers reduced homozygosity but did not swamp ancestral genotypes. *Proceedings of the National Academy of Sciences* 122:e2410945122.

Antunovics Z, Nguyen H, Gaillardin C, Sipiczki M. 2005. Gradual genome stabilisation by progressive reduction of the genome in an interspecific hybrid with. *FEMS Yeast Research* 5:1141–1150.

Arnegard ME et al. 2014. Genetics of ecological divergence during speciation. *Nature* 511:307–311.

Arnold ML. 1997. Natural hybridization and evolution. Oxford University Press, USA

Arnold ML, Ballerini ES, Brothers AN. 2012. Hybrid fitness, adaptation and evolutionary diversification: lessons learned from Louisiana Irises. *Heredity* 108:159–166.

Arnold ML, Hodges SA. 1995. Are natural hybrids fit or unfit relative to their parents? *Trends in Ecology & Evolution* 10:67–71.

Atsumi K, Lagisz M, Nakagawa S. 2021. Nonadditive genetic effects induce novel phenotypic distributions in male mating traits of F1 hybrids. *Evolution* 75:1304–1315.

Baird SJE, Petružela J, Jaroň I, Škrabánek P, Martínková N. 2023. Genome polarisation for detecting barriers to geneflow. *Methods in Ecology and Evolution* 14:512–528.

Barton NH. 2001. The role of hybridization in evolution. *Molecular Ecology* 10:551–568.

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics* 16:113–148.

Bernardes JP, Stelkens RB, Greig D. 2017. Heterosis in hybrids within and between yeast species. *j. evol. Biol.* 30:538–548.

Bierne N, Lenormand T, Bonhomme F, David P. 2002. Deleterious mutations in a hybrid zone: can mutational load decrease the barrier to gene flow? *Genet. Res.* 80:197–204.

Blanckaert A, Bank C. 2018. In search of the Goldilocks zone for hybrid speciation. *PLOS Genetics* 14:e1007613.

Blanckaert A, Bank C, Hermisson J. 2020. The limits to parapatric speciation 3: evolution of strong reproductive isolation in presence of gene flow despite limited ecological differentiation. *Philos Trans R Soc Lond B Biol Sci* 375:20190532.

Bomblies K et al. 2007. Autoimmune response as a mechanism for a Dobzhansky-Muller-Type incompatibility syndrome in plants. *PLOS Biology* 5:e236.

Botero CA, Weissing FJ, Wright J, Rubenstein DR. 2015. Evolutionary tipping points in the capacity to adapt to environmental change. *Proc. Natl. Acad. Sci. U.S.A.* 112:184–189.

Brandt DYC, Huber CD, Chiang CWK, Ortega-Del Vecchyo D. 2024. The promise of inferring the past using the ancestral recombination graph. *Genome Biol Evol* 16:evae005.

Brice C, Zhang Z, Bendixsen D, Stelkens R. 2021. Hybridization outcomes have strong genomic and environmental contingencies. *Am Nat* 198:E53–E67.

Brooks ME et al. 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal* 9:378–400.

Buerkle CA, Morris RJ, Asmussen MA, Rieseberg LH. 2000. The likelihood of homoploid hybrid speciation. *Heredity* 84:441–451.

Buerkle CA, Rieseberg LH. 2008. The rate of genome stabilization in homoploid hybrid species. *Evolution* 62:266–275.

Cerca J, Stelkens R. 2026. The identity crisis of speciation genomics. *evol j linn soc* 5:kzag011.

Chapman MA, Burke JM. 2007. Genetic divergence and hybrid speciation. *Evolution* 61:1773–1780.

Charlesworth B. 1998. Measures of divergence between populations and the effect of forces that reduce variability. *Molecular biology and evolution* 15:538–543.

Charlesworth B, Coyne JA, Barton NH. 1987. The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist* 130:113–146.
Charlesworth D, Morgan MT, Charlesworth B. 1990. Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. *Evolution* 44:1469–1489.
Chaturvedi S et al. 2020. Recent hybrids recapitulate ancient hybrid outcomes. *Nat Commun* 11:2179.
Comeault AA, Matute DR. 2018. Genetic divergence and the number of hybridizing species affect the path to homoploid hybrid speciation. *Proc. Natl. Acad. Sci. U.S.A.* 115:9761–9766.
Coughlan JM, Matute DR. 2020. The importance of intrinsic postzygotic barriers throughout the speciation process. *Phil. Trans. R. Soc. B* 375:20190533.
Coyne JA, Orr HA. 1989. Patterns of speciation in *Drosophila*. *Evolution* 43:362–381.
Cruickshank TE, Hahn MW. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology* 23:3133–3157.
Dagilis AJ et al. 2022. A need for standardized reporting of introgression: insights from studies across eukaryotes. *Evol Lett* 6:344–357.
Dagilis AJ, Kirkpatrick M, Bolnick DI. 2019. The evolution of hybrid fitness during speciation. *PLOS Genetics* 15:e1008125.
Dagilis AJ, Matute DR. 2023. The fitness of an introgressing haplotype changes over the course of divergence and depends on its size and genomic location. *PLOS Biology* 21:e3002185.
Danecek P et al. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10:giab008.
Edelman NB, Mallet J. 2021. Prevalence and adaptive impact of introgression. *Annual Review of Genetics* 55:265–283.
Endler JA. 1977. Geographic variation, speciation, and clines. Princeton University Press
Fitzpatrick BM, Shaffer HB. 2007. Hybrid vigor between native and introduced salamanders raises new challenges for conservation. *Proc. Natl. Acad. Sci. U.S.A.* 104:15793–15798.
Fraïsse C, Gunnarsson PA, Roze D, Bierne N, Welch JJ. 2016. The genetics of speciation: insights from Fisher's geometric model. *Evolution* 70:1450–1464.
Fraïsse C, Sachdeva H. 2021. The rates of introgression and barriers to genetic exchange between hybridizing species: sex chromosomes vs autosomes. *Genetics* 217:iyaa025.
Gaczorek T et al. 2024. Widespread adaptive introgression of major histocompatibility complex genes across vertebrate hybrid zones. *Molecular Biology and Evolution* 41:msae201.
Garrison E et al. 2018. Variation graph toolkit improves read mapping by representing genetic variation in the reference. *Nat Biotechnol* 36:875–879.
Grabenstein KC, Taylor SA. 2018. Breaking barriers: causes, consequences, and experimental utility of human-mediated hybridization. *Trends in Ecology & Evolution* 33:198–212.
Hahn MW. 2018. Molecular population genetics. Oxford University Press
Haldane JBS. 2022. Sex ratio and unisexual sterility in hybrid animals. In: Selected Genetic Papers of J.B.S. Haldane (Routledge Revivals). p. 495–503
Harris K, Nielsen R. 2016. The genetic cost of Neanderthal introgression. *Genetics* 203:881–891.
Harrison RG, Larson EL. 2014. Hybridization, introgression, and the nature of species boundaries. *jhered* 105:795–809.
Hartig F, Lohse L, Leite M de S, Werneke C. 2026. DHARMa: residual diagnostics for hierarchical (multi-Level / mixed) regression models. Available from: https://cran.r-project.org/web/packages/DHARMa/index.html
Healy TM, Burton RS. 2020. Strong selective effects of mitochondrial DNA on the nuclear genome. *Proc. Natl. Acad. Sci. U.S.A.* 117:6616–6621.
Heidbreder P et al. 2026. Genomic incompatibilities are persistent barriers when speciation happens with gene flow in Formica ants. *Mol Biol Evol* 43:msag063.

Heil CSS et al. 2019. Temperature preference can bias parental genome retention during hybrid evolution. *PLOS Genetics* 15:e1008383.
Hermansen JS et al. 2011. Hybrid speciation in sparrows I: phenotypic intermediacy, genetic admixture and barriers to gene flow. *Mol Ecol* 20:3812–3822.
Hill GE. 2017. The mitonuclear compatibility species concept. *The Auk* 134:393–409.
Hirase S et al. 2021. Integrative genomic phylogeography reveals signs of mitonuclear incompatibility in a natural hybrid goby population. *Evolution* 75:176–194.
Hoffmann AA, Sgrò CM. 2011. Climate change and evolutionary adaptation. *Nature* 470:479–485.
Howard-McCombe J et al. 2023. Genetic swamping of the critically endangered Scottish wildcat was recent and accelerated by disease. *Current Biology* 33:4761-4769.e5.
Jiang Y et al. 2020. Differentiating homoploid hybridization from ancestral subdivision in evaluating the origin of the D lineage in wheat. *The New Phytologist* 228:409–414.
de Jong MJ et al. 2025. Species delimitation using genome-wide estimates of *Dxy* and *Fst*. *bioRxiv*. Available from: https://www.biorxiv.org/content/10.64898/2025.12.08.692692v1
Juric I, Aeschbacher S, Coop G. 2016. The strength of selection against Neanderthal introgression. *PLoS Genet* 12:e1006340.
Ketola T, Kronholm I. 2023. Experimental evolution of evolutionary potential in fluctuating environments. *Journal of Evolutionary Biology* 36:945–949.
Koevoets T, Beukeboom LW. 2009. Genetics of postzygotic isolation and Haldane's rule in haplodiploids. *Heredity* 102:16–23.
Koppetsch T, Malinsky M, Matschiner M. 2024. Towards reliable detection of introgression in the presence of among-species rate variation. *Syst Biol* 73:769–788.
Korunes KL, Samuk K. 2021. pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Mol Ecol Resour* 21:1359–1368.
Krapf P, Blankers T, Kulmuni J. 2026. Do the benefits of hybridization outweigh the costs under conditions of global change? Available from: https://ecoevorxiv.org/repository/view/13242/
Kuijt M et al. 2026. Hybrid breakdown is temporary and not expressed in a novel environment during 50 generations of experimental evolution in *Tetranychus urticae* hybrids. *bioRxiv*. Available from: https://www.biorxiv.org/content/10.64898/2026.07.16.738867v1
Kulmuni J, Wiley B, Otto SP. 2024. On the fast track: hybrids adapt more rapidly than parental populations in a novel environment. *Evol Lett* 8:128–136.
Langdon QK et al. 2022. Predictability and parallelism in the contemporary evolution of hybrid genomes. *PLOS Genetics* 18:e1009914.
Langdon QK et al. 2024. Swordtail fish hybrids reveal that genome evolution is surprisingly predictable after initial hybridization. *PLOS Biology* 22:e3002742.
Li J, Lee C-R. 2023. The role of gene presence–absence variations on genetic incompatibility in Asian rice. *New Phytologist* 239:778–791.
Li J, Schumer M, Bank C. 2022. Imbalanced segregation of recombinant haplotypes in hybrid populations reveals inter- and intrachromosomal Dobzhansky-Muller incompatibilities. *PLOS Genetics* 18:e1010120.
Li Z, Zhao Y, Luo K. 2024. Molecular mechanisms of heterosis and its applications in tree breeding: progress and perspectives. *Int J Mol Sci* 25:12344.
Lunter G, Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res* 21:936–939.
Malinsky M, Matschiner M, Svardal H. 2021. Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Mol Ecol Resour* 21:584–595.
Mallet J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* 20:229–237.

Mallet J. 2007. Hybrid speciation. *Nature* 446:279–283.
Mank JE, Axelsson E, Ellegren H. 2007. Fast-X on the Z: rapid evolution of sex-linked genes in birds. *Genome Res.* 17:618–624.
Mank JE, Nam K, Ellegren H. 2009. Faster-Z evolution is predominantly due to genetic drift. *Molecular Biology and Evolution* 27:661–670.
Martin SH et al. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research* 23:1817–1828.
Martin SH, Davey JW, Salazar C, Jiggins CD. 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biol* 17:e2006288.
Martínez AA, Lang GI. 2026. Pervasive weak incompatibilities shape phenotypic variation and genome evolution in yeast hybrids. *Nat Commun* 17:10662.
Matute DR et al. 2020. Rapid and predictable evolution of admixed populations between two *Drosophila* species pairs. *Genetics* 214:211–230.
Matute DR, Butler IA, Turissini DA, Coyne JA. 2010. A test of the snowball theory for the rate of evolution of hybrid incompatibilities. *Science* 329:1518–1521.
Mavárez J et al. 2006. Speciation by hybridization in *Heliconius* butterflies. *Nature* 441:868–871.
Meisel RP, Connallon T. 2013. The faster-X effect: integrating theory and data. *Trends Genet* 29:537–544.
Monnet F et al. 2025. Rapid establishment of species barriers in plants compared with that in animals. *Science* 389:1147–1150.
Moore WS. 1977. An evaluation of narrow hybrid zones in vertebrates. *The Quarterly Review of Biology* 52:263–277.
Moran BM et al. 2021. The genomic consequences of hybridization. *eLife* 10:e69016.
Moran BM et al. 2024. A lethal mitonuclear incompatibility in complex I of natural hybrids. *Nature* 626:119–127.
Moyle LC, Nakazato T. 2010. Hybrid incompatibility “snowballs” between *Solanum* species. *Science* 329:1521–1523.
Muraro M, Falaschi M, Ficetola GF. 2022. Patterns of performance variation between animal hybrids and their parents: a meta-analysis. *Evol Biol* 49:482–496.
Nielsen R, Vaughn AH, Deng Y. 2025. Inference and applications of ancestral recombination graphs. *Nat Rev Genet* 26:47–58.
Noor M a. F, Bennett SM. 2009. Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity* 103:439–444.
Nouhaud P, Blanckaert A, Bank C, Kulmuni J. 2020. Understanding admixture: haplodiploidy to the rescue. *Trends in Ecology & Evolution* 35:34–42.
Nouhaud P, Martin SH, Portinha B, Sousa VC, Kulmuni J. 2022. Rapid and predictable genome evolution across three hybrid ant populations. *PLOS Biology* 20:e3001914.
Orr HA. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* 139:1805–1813.
Orr HA, Turelli M. 2001. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution* 55:1085–1094.
Ottenburghs J. 2021. The genic view of hybridization in the Anthropocene. *Evolutionary Applications* 14:2342–2360.
Owens GL et al. 2025. Shared selection and genetic architecture drive strikingly repeatable evolution in long-term experimental hybrid populations. *Mol Biol Evol* 42:msaf014.
Pereira RJ, Lima TG, Pierce-Ward NT, Chao L, Burton RS. 2021. Recovery from hybrid breakdown reveals a complex genetic architecture of mitonuclear incompatibilities. *Mol Ecol* 30:6403–6416.

Poikela N, Snook RR, Kulmuni J, Ritchie MG. 2026. Adaptive introgression between two *Drosophila* species enhances heat tolerance despite barriers to gene flow. *PLOS Biology* 24:e3004028.
Presgraves DC. 2010. Speciation genetics: search for the missing snowball. *Current Biology* 20:R1073–R1074.
Puchta O et al. 2016. Network of epistatic interactions within a yeast snoRNA. *Science* 352:840–844.
Rice WR. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 38:735–742.
Rieseberg LH et al. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* 301:1211–1216.
Rieseberg LH, Archer MA, Wayne RK. 1999. Transgressive segregation, adaptation and speciation. *Heredity* 83:363–372.
Roux C et al. 2016. Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLOS Biology* 14:e2000234.
Runemark A et al. 2018. Variation and constraints in hybrid genome formation. *Nat Ecol Evol* 2:549–556.
Runemark A, Vallejo-Marin M, Meier JI. 2019. Eukaryote hybrid genomes. *PLoS Genet* 15:e1008404.
Sachdeva H, Barton NH. 2018a. Replicability of introgression under linked, polygenic selection. *Genetics* 210:1411–1427.
Sachdeva H, Barton NH. 2018b. Introgression of a block of genome under infinitesimal selection. *Genetics* 209:1279–1303.
Sánchez-Guillén RA, Arce-Valdés LR, Ballén-Guapacha AV, Ordaz-Morales JE, Stand-Pérez M. 2024. Interspecific hybridization in insects in times of climate change. In: Effects of climate change on insects: physiological, evolutionary, and ecological responses. Oxford University Press
Sankararaman S, Mallick S, Patterson N, Reich D. 2016. The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Current Biology* 26:1241–1247.
Satokangas I et al. 2026. Introgression and divergence in a young species group. *Molecular Ecology* 35:e70448.
Satokangas I, Martin SH, Helanterä H, Saramäki J, Kulmuni J. 2020. Multi-locus interactions and the build-up of reproductive isolation. *Philos Trans R Soc Lond B Biol Sci* 375:20190543.
Saunders PA, Muyle A. 2024. Sex chromosome evolution: hallmarks and question marks. *Molecular Biology and Evolution* 41:msae218.
Schneemann H, De Sanctis B, Roze D, Bierne N, Welch JJ. 2020. The geometry and genetics of hybridization. *Evol* 74:2575–2590.
Schneemann H, Welch JJ. 2025. When will hybridization bring long-term benefits? *bioRxiv*. Available from: https://www.biorxiv.org/content/10.1101/2025.01.16.633403v1
Schumer M et al. 2018. Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science* 360:656–660.
Schumer M, Cui R, Rosenthal GG, Andolfatto P. 2015. Reproductive isolation of hybrid populations driven by genetic incompatibilities. *PLOS Genetics* 11:e1005041.
Schumer M, Rieseberg LH. 2026. Genetics and genomics of hybridization. *Nat Rev Genet* 27:840–854.
Setter D, Lohse K, Baird SJE. 2026. diempy: fast and reference-free genome polarization and chromosome painting. *G3 Genes|Genomes|Genetics*:jkag140.
Shapira R, Levy T, Shaked S, Fridman E, David L. 2014. Extensive heterosis in growth of yeast hybrids is explained by a combination of genetic models. *Heredity* 113:316–326.
Shull GH. 1948. What is “heterosis”? *Genetics* 33:439–446.
Simon A, Bierne N, Welch JJ. 2018. Coadapted genomes and selection on hybrids: Fisher’s geometric model explains a variety of empirical patterns. *Evolution Letters* 2:472–498.
Stankowski S et al. 2024. Toward the integration of speciation research. *Evolutionary Journal of the Linnean Society* 3:kzae001.

Stelkens R, Seehausen O. 2009. Genetic distance between species predicts novel trait expression in their hybrids. *Evolution* 63:884–897.
Stelkens RB, Schmid C, Seehausen O. 2015. Hybrid breakdown in cichlid fish. *PLoS ONE* 10:e0127207.
Strauss SK et al. 2026. Quantitative genetics of natural *S. cerevisiae* strains upon mating reveal heritable determinants of fitness and differential mating affinities. *Cell Reports* [Internet] 45. Available from: https://www.cell.com/cell-reports/abstract/S2211-1247(26)00050-1
Suarez-Gonzalez A, Lexer C, Cronk QCB. 2018. Adaptive introgression: a plant perspective. *Biol. Lett.* 14:20170688.
Szabo N, Cutter AD. 2024. Experimental evolution of hybrid populations to identify Dobzhansky-Muller incompatibility loci. *Ecol Evol* 14:e10972.
Taylor SA, Larson EL. 2019. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nat Ecol Evol* 3:170–177.
Telis N, Aguilar R, Harris K. 2020. Selection against archaic hominin genetic variation in regulatory regions. *Nat Ecol Evol* 4:1558–1566.
The Darwin Tree of Life Project Consortium. 2022. Sequence locally, think globally: The Darwin Tree of Life Project. *Proceedings of the National Academy of Sciences* 119:e2115642118.
Thompson KA et al. 2022. Analysis of ancestry heterozygosity suggests that hybrid incompatibilities in threespine stickleback are environment dependent. *PLOS Biology* 20:e3001469.
Thompson KA et al. 2024. The ecology of hybrid incompatibilities. *Cold Spring Harb Perspect Biol* 16:a041440.
Todesco M et al. 2016. Hybridization and extinction. *Evol Appl* 9:892–908.
Tricou T, Tannier E, de Vienne DM. 2022. Ghost lineages highly influence the interpretation of introgression tests. *Syst Biol* 71:1147–1158.
Trier CN, Hermansen JS, Sætre G-P, Bailey RI. 2014. Evidence for mito-nuclear and sex-linked reproductive barriers between the hybrid Italian sparrow and its parent species. *PLoS Genet* 10:e1004075.
Turelli M, Orr HA. 2000. Dominance, epistasis and the genetics of postzygotic isolation. *Genetics* 154:1663–1679.
Turissini DA, Comeault AA, Liu G, Lee YCG, Matute DR. 2017. The ability of *Drosophila* hybrids to locate food declines with parental divergence. *Evolution* 71:960–973.
Vasimuddin Md, Misra S, Li H, Aluru S. 2019. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. In: 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS). p. 314–324 Available from: https://ieeexplore.ieee.org/abstract/document/8820962
Veller C, Edelman NB, Muralidhar P, Nowak MA. 2023. Recombination and selection against introgressed DNA. *Evolution* 77:1131–1144.
Walker JM et al. 2026. A standardised framework for classifying estimates of reproductive isolation. *Evol Lett* 10:441–450.
Wei X, Zhang J. 2018. The optimal mating distance resulting from heterosis and genetic incompatibility. *Sci. Adv.* 4:eaau5518.
Wilkinson MD et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3:160018.
Wong Y et al. 2024. A general and efficient representation of ancestral recombination graphs. *Genetics* 228:iyae100.
Würschum T et al. 2023. Maximization through optimization? On the relationship between hybrid performance and parental genetic distance. *Theor Appl Genet* 136:186.
Xiong T, Mallet J. 2022. On the impermanence of species: The collapse of genetic incompatibilities in hybridizing populations. *Evolution* 76:2498–2512.
Yu SB et al. 1997. Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proceedings of the National Academy of Sciences* 94:9226–9231.

1093 Zhang Z et al. 2020. Recombining your way out of trouble: the genetic architecture of hybrid fitness under
1094 environmental stress. *Molecular Biology and Evolution* 37:167–182.
1095