Polygenic Effects on the Z Chromosome Underlie Hybrid Incompatibility in *Papilio* and *Heliconius* Butterflies

Tianzhu Xiong[1]*†, Shreeharsha Tarikere[1]‡, Neil Rosser[1], Xueyan Li[2], Masaya Yago[3], James Mallet[1]

[1]Department of Organismic and Evolutionary Biology, Harvard University; Cambridge, MA, 02138, USA.
[2]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences; Kunming, Yunnan, 650223, China.
[3]The University Museum, The University of Tokyo; Bunkyo-ku, Tokyo, 113-0033, Japan.
*Corresponding author. Email: xtz20042328@hotmail.com
†Current address: Department of Molecular Biology and Genetics, Cornell University; Ithaca, NY, 14853, USA.
‡Current address: Department of Earth Sciences, Carleton University; Ottawa, ON, K1S 5B6, Canada.

Abstract: The fitness of animal hybrids follows two empirical rules: hybrids of the heterogametic sex are more unfit (Haldane's Rule), and the sex chromosome is disproportionately involved in incompatibility (the large-X/Z effect). Whether these rules result from genetic mechanisms shared across taxa remains unknown, and existing explanations rarely consider female heterogametic taxa such as butterflies. Here, we investigate hybrid incompatibilities in *Papilio* and *Heliconius* butterflies, and show that defects coincide with unbalanced introgression between the Z chromosome and its genetic background. This polygenic mechanism predicts both rules because introgressed ancestry on the Z chromosome is more skewed in females, and is more variable than on all autosomes. Therefore, the explanation for both rules in butterflies shares little similarity with prevailing theories relying on dominance.

Introduction

Speciation is a complex and stochastic process, yet it obeys empirical rules across taxa with sexual reproduction (*1*). Haldane's Rule states that among the hybrids between different species, the heterogametic sex (the sex with XY or ZW sex chromosomes) tends to have lower fitness (*2*, *3*). A second rule is the so-called large-X effect, which states that the sex chromosome is disproportionately involved in hybrid incompatibility (*1*, *4*). Haldane's Rule is entirely phenomenological, but it holds across many phylogenetically diverse organisms (*5*–*7*). Whether adherence to Haldane's Rule emerges from a common set of genetic mechanisms is an open question (*8*). The large-X effect also appears robust, but without mapped incompatibility factors, the evidence is often indirect and hard to interpret (*9*–*13*).

Several mechanisms can explain Haldane's Rule. First, dominance theory posits that the single X chromosome in males might expose recessive genes that are deleterious in a hybrid genetic background, thus increasing the likelihood of incompatibility (*4*, *14*). Second, the evolution of sex chromosomes may be faster than autosomes, thus Haldane's Rule can be produced via a variety of processes involving hemizygous haploid selection (*15*), sex-specific selection (*16*), and sex chromosome conflict (*17*, *18*). Accelerated sex chromosome evolution also provides a natural explanation for the large-X effect. Empirical evidence for these mechanisms is based mostly on male heterogametic taxa, such as mammals and *Drosophila* (*19*, *20*), but whether they are general explanations for the two rules is unknown. It has also been suggested that spermatogenesis is particularly prone to disruption, which can explain the higher incidence of male sterility in XX/XY systems (*21*), but this is not applicable to Haldane's Rule in female heterogametic taxa.

In Lepidoptera (butterflies and moths), the female is the heterogametic sex (with Z/W sex chromosomes), and hybrid females are more prone to defects than males (*2*, *22*). To date, little is known about the genomic basis of Haldane's Rule in Lepidoptera, except for a few studies using sparse genetic markers (*23*–*25*) and a recent whole-genome Quantitative Trait Locus (QTL) study in *Heliconius* (*26*). Nonetheless, these studies demonstrate that hybrid female sterility is associated with the Z chromosome, consistent with a large-Z effect. Here, we map hybrid incompatibility between two closely related butterflies, *Papilio bianor* and *Papilio dehaanii* (Fig. 1A), in which interspecific crosses produce fertile males but completely sterile females (*27*, *28*). Hybrids also develop abnormal body size, which is frequently observed in *Papilio* (*29*). To test for the genetic basis of the two rules in this system, we carry out QTL studies of body size (pupal weight) and female reproduction (ovary dysgenesis) in backcross hybrids. We then compare *Papilio* with *Heliconius* (*26*) to test whether these genera share a similar genomic basis for the two rules.

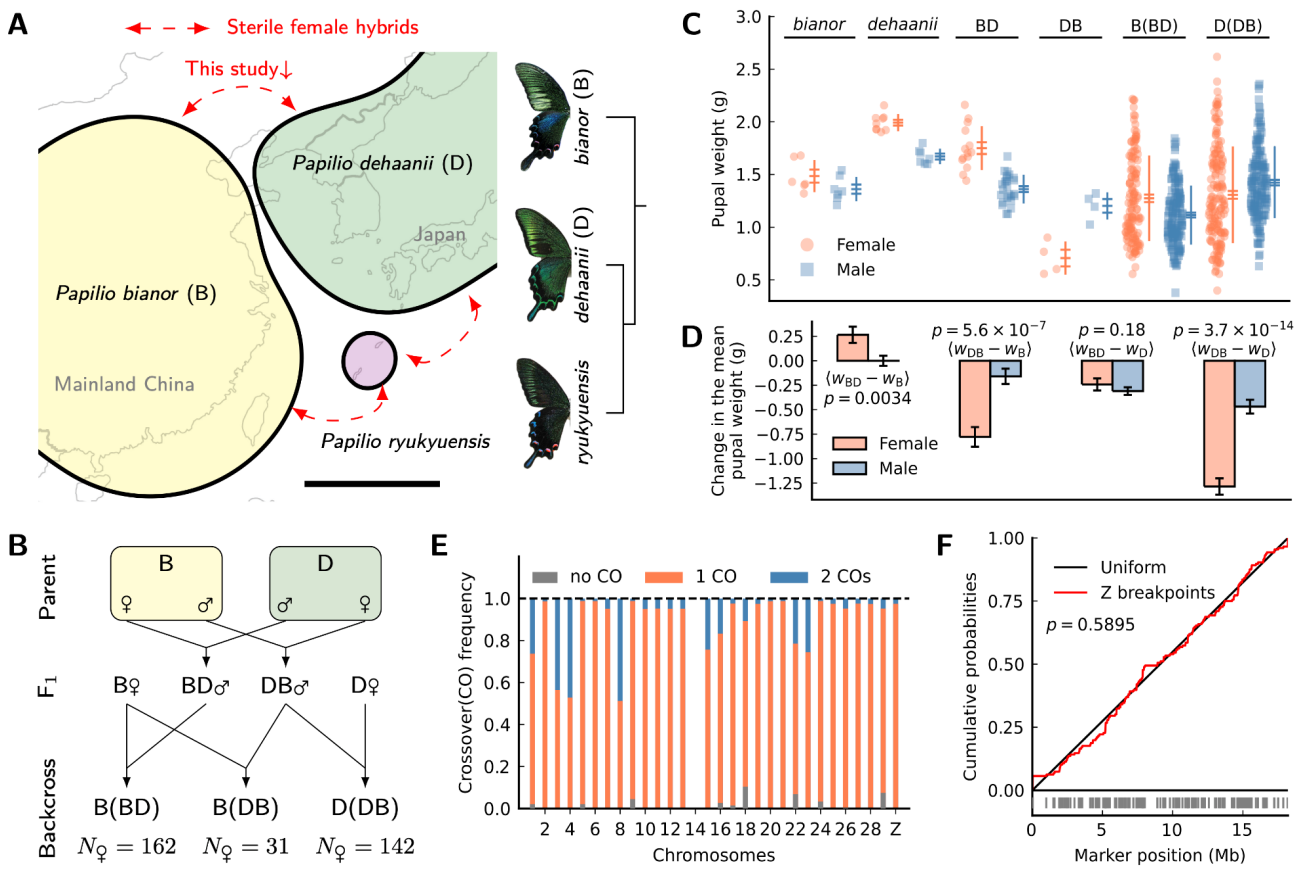

Figure 1. The study design, variation in pupal weight, and patterns of meiotic crossover in males. (A) Geographic distribution and species relationships in the *Papilio bianor* complex. Scale bar=1000km. **(B)** Crossing design. **(C)** Pupal weight variation among parents, F_1s, and backcrosses. Horizontal bars represent the mean +/- standard error. Vertical lines represent the mean +/- standard deviation. **(D)** Changes in mean pupal weight between F_1s and parents in each sex. Error bars are standard errors of the mean changes. The significance of male vs female differences is shown as *p*-values of Z-tests. **(E)** The crossover frequency in F_1 males per chromosome pair per meiosis. Chr14 is excluded due to assembly problems. **(F)** Recombination breakpoints on the Z chromosome are uniformly distributed (*p*-value is from a Kolmogorov-Smirnov test). Vertical bars are inferred breakpoints.

Haldane's Rule between *P. bianor* and *P. dehaanii* involves asymmetrically inherited elements

To investigate hybrid abnormality, we performed reciprocal F_1 crosses and backcrosses between the two species (Fig. 1B). We follow the order (female × male) in notation. For instance, "B(BD)" is equivalent to "*bianor* ♀ × (*bianor* ♀ × *dehaanii* ♂) ♂", where "B" and "D" stand for *bianor* and *dehaanii*, respectively.

Pupal weight (*W*) is treated as a proxy for adult body size. F_1 females with a *dehaanii* mother ("DB") are significantly smaller than females of either parental species, but in the reciprocal cross with a *bianor* mother ("BD") they span the range of parental females (Fig. 1C). For F_1 males, deviation in pupal weight from pure males is less extreme than that of F_1 females (Fig. 1D). We interpret female-biased abnormal size as a defect conforming to Haldane's Rule.

Focusing on hybrid female sterility, we dissected ovaries across the pedigree and determined major ovary phenotypes (Fig. 2). To our surprise, while F_1 females with a *bianor* mother ("BD") have almost empty ovaries (Fig. 2C, 2J), F_1 females in the reciprocal cross ("DB") develop and lay superficially normal eggs (Fig. 2B, 2I). However, eggs laid by mated females never hatch. We score ovary phenotype rather than female fertility *per se*: ovaries with regularly spaced and spherical follicles are subsequently all classified as "Normal". Since deformed ovaries lead to sterility and can be readily scored, this is a sufficient and efficient approach to detect sterility factors.

Ovary and female size phenotypes differ significantly between reciprocal crosses and are therefore examples of "Darwin's corollary" of hybrid incompatibility (*30*). This phenomenon is likely due to asymmetrically inherited genetic elements. For this reason, we separate backcross types according to maternal origin in QTL analysis.

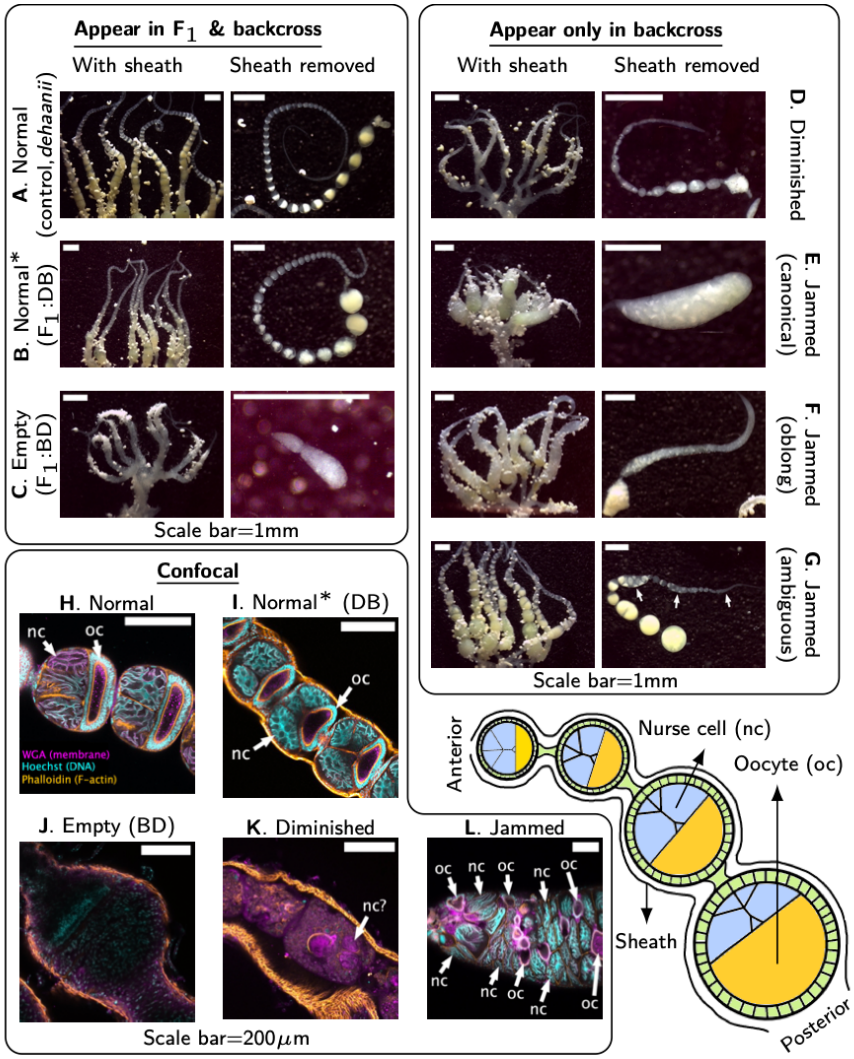

Figure 2. Ovary morphology supports a categorical classification of phenotypic defects in *Papilio*. Monochrome confocal images for each channel are in Fig. S1 and S2 (WGA stains membranes in magenta; Hoechst stains DNA in cyan; Phalloidin stains F-actin in orange). A schematic diagram of butterfly ovarioles is shown at the bottom right. Scale bars for stereoscope images are approximate. **(A, H)** Phenotype Normal in a pure individual. The confocal image shows follicles with sheath removed. **(B, I)** Phenotype Normal in F_1 females of the cross DB. Sheath is retained in the confocal image. **(C, J)** Phenotype Empty in F_1 females of the cross BD. Sheath is retained in the confocal image with little tissue inside. **(D, K)** Phenotype Diminished. Sheath is retained in the confocal image with a substantial amount of interior tissue without discernible follicle structures. **(E-G, L)** Phenotype Jammed. Sheath is removed in the confocal image, and follicle cells fuse into tubes with many nurse cells and oocytes. Subtype canonical: each ovariole collapses into a single tube of oocytes and nurse cells. Subtype oblong: tubes are elongated. Subtype ambiguous: tubes and isolated follicles coexist.

132 **Patterns of meiotic crossover on the Z chromosome in F_1 males**

133 Understanding how recombination determines chromosomal ancestry is crucial to our subsequent
134 analysis. To infer haplotypes and crossover patterns, we carried out whole-genome low-coverage (~1x)
135 sequencing in backcrosses, while F_1s and parents were sequenced to higher depths (>5x and >30x,
136 respectively). Prior to crossover analysis, we used linkage information from all families to correct
137 assembly errors in the reference genome of *P. bianor* (*31*), except for chromosome 14, where errors
138 remained unresolved (Fig. S4). As assembly errors can affect inference of recombination breakpoints,
139 we do not report crossover patterns on chromosome 14.
140
141 We inferred the crossover pattern in F_1 males by counting the estimated recombination breakpoints
142 across all backcross offspring (female meiosis in Lepidoptera lacks crossovers). Most F_1 males had at
143 least one crossover per chromosome pair per meiosis, but the degree of crossover interference varied
144 among chromosomes (Fig. 1E). Double crossovers were frequent on some chromosomes (>40% on
145 chromosome 8) but very rare on most. Importantly, the Z chromosome in the *Papilio* crosses had almost
146 no double crossovers, and its recombination breakpoints were approximately uniformly distributed
147 along the chromosomal axis (Fig. 1F). Recombination on the Z chromosome can thus be approximated
148 by a model where the first crossover occurs uniformly and randomly across the chromosome, while the
149 second crossover is completely suppressed. We apply this model to the study of incompatibilities.

150 **Abnormal pupal weight is determined by multiple additive factors scattered across the Z**
151 **chromosome**

152 Pupal weight in offspring of B(BD) and D(DB) backcrosses has a broad distribution (Fig.1C). Such a
153 distribution could be associated with multiple chromosomes or extreme developmental stochasticity.
154 Yet single-marker QTL scans suggest that the Z chromosome alone controls pupal weight variation in
155 females (Fig. 3A). These scans reveal a major QTL near the center of the Z chromosome in both cross
156 types, explaining over 50% of the phenotypic variance (Fig. S8). Nonetheless, we reason below that this
157 major QTL is likely a statistical artifact caused by multiple factors scattered across the Z chromosome.
158
159 The following evidence supports Z-linked polygenic architecture. First, pupal weight depends almost
160 continuously on the fraction of introgressed ancestry on the Z chromosome (Fig. 3B; from here on the
161 Z-chromosome ancestry fraction is denoted as f_Z, and the autosomal ancestry fraction as f_A). Second,
162 genetic variance of pupal weight is much smaller than that expected for a single locus of major effect,
163 so that a crude estimate of the number of loci (n_{CWZ}) (*32*) using segregating genetic variance is
164 effectively infinite (Table S1). Thus, multiple additive factors may be involved. Third, we compare two
165 extreme genetic architectures: a single-QTL model versus a polygenic additive model in which weight
166 varies linearly with f_Z. Given the polygenic model, the strength of the genotype-phenotype association
167 (R^2) will always be larger for models using f_Z (i.e., $R^2_{1\text{-marker}}/R^2_{Z\text{-ancestry}}<1$ for all markers). Conversely,

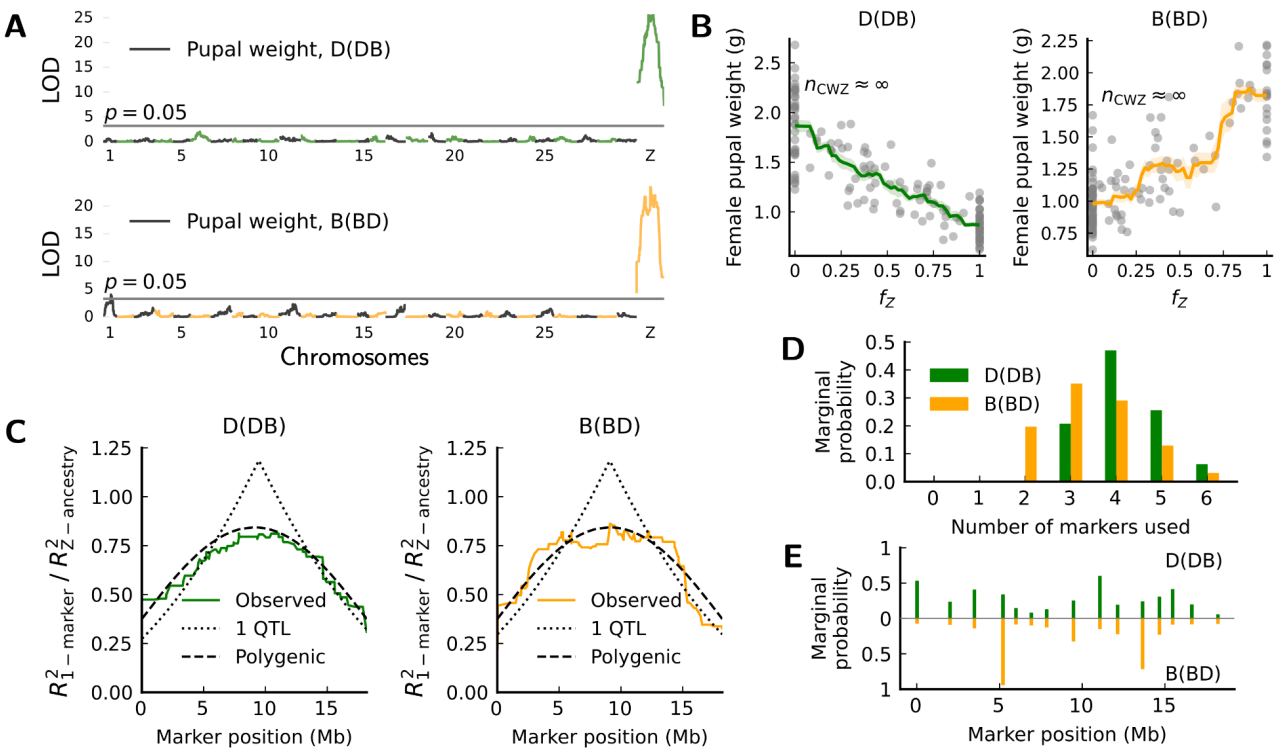

Figure 3. Pupal weight variation in female hybrids maps to many additive Z-linked effects. (A) One-dimensional QTL scan shows that variation in pupal weight is mainly explained by the Z chromosome. LOD scores peak at the Z chromosome center. Alternating colors represent different chromosomes. **(B)** Ancestry fraction on the Z chromosome (f_Z) is highly informative of pupal weight. A crude estimator of the effective number of genes (n_{CWZ}) (*32*) using segregating phenotypic variance approaches infinity. **(C)** Genotype-phenotype regression does not support any 1-QTL architecture for pupal weight. The distribution of regression power is measured by $R^2_{1\text{-marker}}/R^2_{Z\text{-ancestry}}$ (the ratio between R^2 using a single marker versus a "polygenic" model of Z chromosome ancestry). $R^2_{1\text{-QTL}}$ and $R^2_{polygenic}$ are the fit between observed and theoretical curves. The polygenic model provides a good fit for the observed curves. **(D, E)** Bayesian QTL model selection (*33*) based on 15 sparsely spaced markers favors the simultaneous inclusion of multiple markers (≥3) in predicting pupal weight. Markers more favored by models have higher marginal probabilities of being selected. Marginal probabilities equal the partial sum of the posterior probabilities of QTL models.

given the single-QTL model, markers tightly linked to the QTL will surpass f_Z in association strength (i.e., $R^2_{1\text{-marker}}/R^2_{Z\text{-ancestry}}>1$ for some markers). Using the approximate crossover model, we derive $R^2_{1\text{-marker}}/R^2_{Z\text{-ancestry}}$ on the Z chromosome analytically under each architecture (Supplementary Text–2.2), and the observed patterns closely resemble polygenic predictions (Fig. 3C). Fourth, Bayesian QTL model selection (*33*) also favors multiple additive markers in predicting pupal weight (Fig. 3D). Posterior probabilities for markers being selected are more evenly distributed in D(DB) females (Fig. 3E, top), congruent with a near-linear relationship between pupal weight and f_Z (Fig. 3B, left). For B(BD) females, this relationship is less smooth (Fig. 3B, right), and a two-QTL model is a slightly better fit to the observed $R^2_{1\text{-marker}}/R^2_{Z\text{-ancestry}}$ curve (Fig. S9). Nonetheless, the two apparent QTLs flank the chromosome center symmetrically with similar effects, and this is indistinguishable from a polygenic model with abrupt weight increase at two f_Z thresholds (*34*). Finally, the polygenic model offers a simple explanation for the single QTL at the Z chromosome center: after a single crossover, central markers are individually more informative of f_Z than markers near chromosome ends (Supplementary

Text: Theorem 2). Central markers thus provide richer phenotypic information when phenotype varies almost linearly with f_Z, generating a statistical peak in association strength. In summary, while the exact number of Z-linked factors remains unknown, our evidence is highly consistent with a multigenic architecture, rendering the Z-chromosome ancestry fraction informative in predicting phenotype.

Polygenic basis of ovary dysgenesis on the Z chromosome

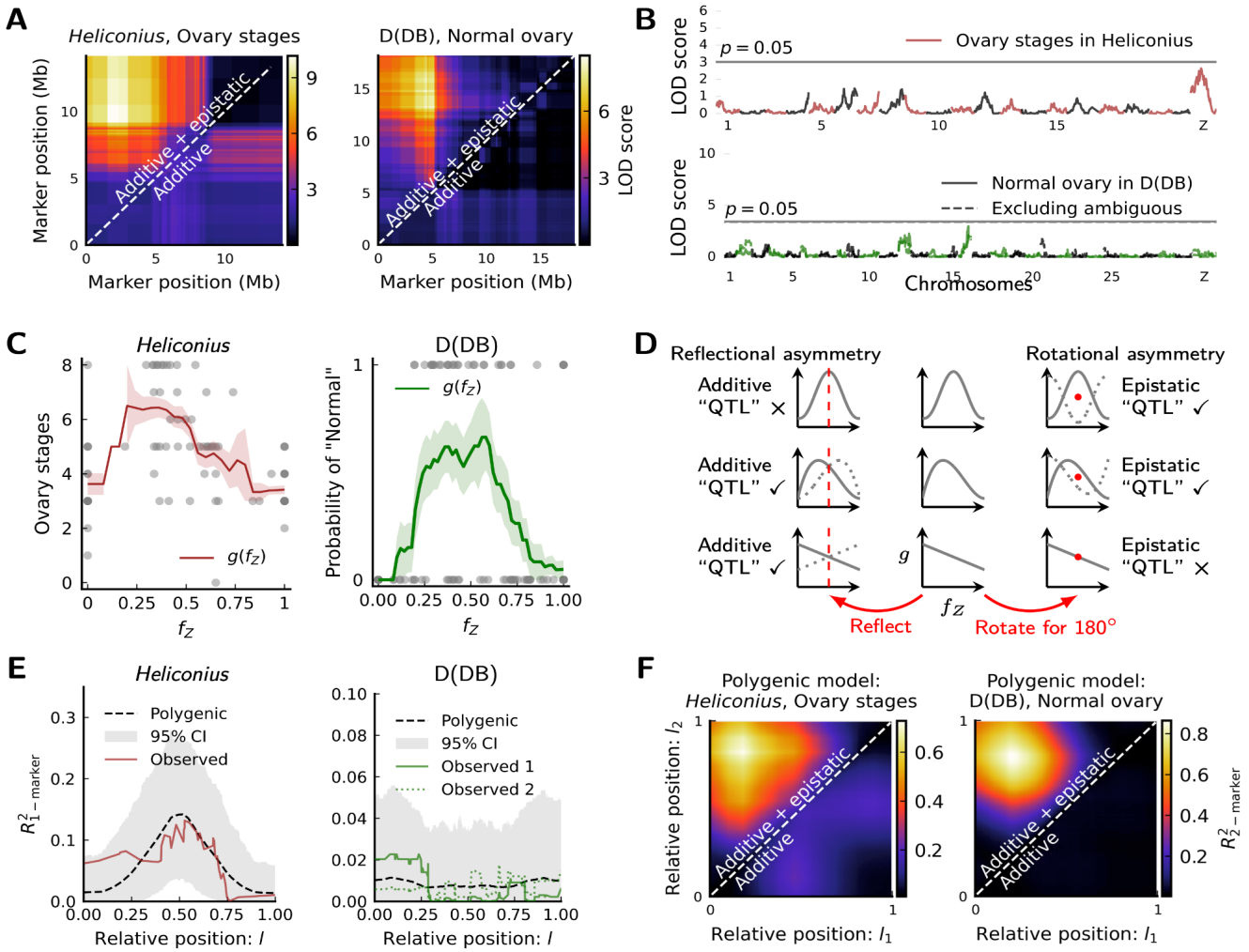

Figure 4. The polygenic basis of ovary dysgenesis in *Heliconius* and D(DB) females. (A) Two-marker scans identify a pair of epistatic QTL located near both ends of the Z chromosome. **(B)** One-marker scans identify a very weak additive QTL on the Z chromosome in *Heliconius* (previously reported to be significant by a different scoring method). No significant additive QTL on the Z chromosome in D(DB) females. **(C)** Ancestry fraction on the Z chromosome (f_Z) is informative of ovary phenotype. The moving average of phenotype scores (window size=0.2) is used as the polygenic model $g(f_Z)$. Shaded areas are means +/- standard errors. Dots in D(DB) represent one round of phenotype assignment due to ambiguous individuals. **(D)** If a trait is fully polygenic on the Z chromosome, QTL analysis via marker-phenotype regression only reveals the asymmetry in $g(f_Z)$. The presence of additive "QTL" is equivalent to a reflectional asymmetry in g, while epistatic "QTL" is equivalent to a rotational asymmetry in g. These conclusions are valid for the inferred Z-chromosome crossover model. **(E)** Polygenic models recover observed results of one-marker scans. Dashed curves and shaded areas represent the means and 95% confidence intervals of $R^2_{1\text{-marker}}$, across 1,000 independent draws of individuals matching real sample sizes ($N_{Heliconius}$=86, $N_{D(DB)}$=142). For D(DB), observed 1 classifies ambiguous individuals as "Normal", while observed 2 classifies them as the opposite. **(F)** Polygenic models recover observed epistasis in two-marker scans. Heatmaps show the means of $R^2_{2\text{-marker}}$ across 100 independent draws of individuals matching real sample sizes.

Ovary dysgenesis leads to hybrid female sterility. The only previous attempt to map hybrid sterility factors genome-wide in Lepidoptera found a pair of epistatic QTLs near each end of the Z chromosome in the backcross from *Heliconius pardalinus sergestus* to *H. p. butleri* (Fig. 4A, left) (*26*). Additionally, that study suggested a weak single-locus QTL near the Z chromosome center (Fig. 4B, top). In *Papilio* D(DB) females, two-marker scans also identified a pair of epistatic QTLs on the Z chromosome for phenotype "Normal" (Fig. 4A, right), although we found no single-locus QTL on the Z chromosome (Fig. 4B, bottom).

Backcross females in *Heliconius* and D(DB) females in *Papilio* develop more normal ovaries when the Z chromosome is recombined, regardless of the precise chromosomal location of introgressed ancestry (Fig. 4C, S11). Since the Z-chromosome ancestry fraction carries so much phenotypic information, we argue that the architecture may again be polygenic on the Z chromosome for both cases, and that the mapped epistatic QTLs on the Z chromosome are likely artifacts. We extend the polygenic model developed previously for pupal weight to ovary dysgenesis. We hypothesize that the expected ovary phenotype is in general a continuous function (g) of the Z-chromosome ancestry fraction (f_Z). In practice, g is the moving average of phenotypic scores with respect to f_Z (Fig. 4C). Again, if phenotypes depend only on ancestry fraction, a significant QTL does not necessarily imply a major effect of the identified locus in development. Rather, these QTLs arise indirectly from the asymmetry in g (Fig. 4D). With the inferred crossover model, we prove for backcrosses that:

i) In one-marker scans, additive QTLs are caused by the reflectional asymmetry in $g(f_Z)$ with respect to its center (Fig. 4D left. See Supplementary Text: Theorems 6-8 and Fig. S12 for details)

ii) In two-marker scans, epistatic QTL pairs are caused by the rotational asymmetry in $g(f_Z)$ with respect to its center (Fig. 4D right. See Supplementary Text: Theorems 9, 10 and Fig. S14 for details)

For both cases of ovary dysgenesis, g has a unimodal form (Fig. 4C). This form is strong in rotational asymmetry but weak in reflectional asymmetry (e.g., the first two rows in Fig. 4D). Consequently, it predicts strongly epistatic QTLs but no or only weak additive QTL, as observed (Fig. 4E, 4F). In contrast, the g for pupal weight is largely linear, which is only reflectionally asymmetric (e.g., the last row in Fig. 4D). This shape predicts a major additive QTL at the chromosome center (Fig. 3C) but no epistatic QTLs (Fig. S10). Thus, not only does the Z-chromosome ancestry fraction carry phenotypic information, but it also predicts the strength of marker-phenotype association across the Z chromosome, and can explain the presence of both additive and epistatic apparent QTLs. This reasoning therefore supports a polygenic basis of the large-Z effect in both systems.

252 **Modulation of incompatibilities by autosomal and maternal backgrounds**

253 Autosomal and maternal genetic backgrounds are also important in modulating hybrid defects. For
254 pupal weight, if Z-linked additive effects were independent of genetic background, F_1 females between
255 a *dehaanii* mother and a *bianor* father should be of similar size as the paternal species. Instead, they are
256 much smaller than either parental female (Fig. 1C), implying epistasis between the *bianor* Z
257 chromosome and the *dehaanii* background. Likewise, pupal weight in B(BD) females is much smaller
258 than both parental species when *bianor*'s Z chromosome contains no introgression (Fig. 3B). This stable
259 decrease in pupal weight can only be explained by the ~25% autosomal introgression from *dehaanii*,
260 implying again that such Z-linked effects could arise via interactions with autosomes, albeit in the
261 opposite maternal background. Similarly, ovaries are defective in *Heliconius* as well as *Papilio* D(DB)
262 backcrosses even when the Z chromosome comes entirely from the maternal species (Fig. 4C),
263 consistent with a significant role of autosomal introgression in defect development. Defects induced via
264 autosome-only introgression develop rather stably (Fig. 3B, S11C, S13A), indicating that multiple
265 autosomal factors linked loosely are necessary to prevent phenotypic segregation. These results suggest
266 that autosomal components of incompatibility are also somewhat polygenic, and may depend on the
267 autosomal ancestry fraction (*56*).

268 **Exceptions to polygenic architecture**

269 The polygenic architecture we have found, however, does not apply to all incompatibilities in *Papilio*.
270 In females with a *bianor* mother, introgression of a small region (~1Mb) on the Z chromosome from
271 *dehaanii* is sufficient to induce the defective ovary phenotype "Empty" (Fig. S15A, Table S2). This
272 phenotype obliterates nearly all follicle tissues in ovarioles (Fig. 2C, 2J), effectively blocking all other
273 ovary defects. In this maternal background, a small region on chromosome 8 also modulates the
274 development of the "Normal" ovary phenotype (Fig. S15B, Table S2). These results suggest that ovary
275 dysgenesis in *Papilio* from this direction of cross is predominantly affected by narrow genomic regions
276 of large effect, in contrast to the polygenic architecture in the opposite maternal background.

277 **Discussion**

278 ***Genetic logic underlying two rules of speciation***

279 Dominance theory requires incompatibility factors to be recessive on the sex chromosome so that the
280 homogametic sex is sheltered by dominance (*7*, *14*). This theory is applicable to Lepidoptera but faces
281 obstacles here. First, polygenicity implies that a large number of Z-linked incompatibility factors
282 should be simultaneously recessive, which is untested in Lepidoptera. Second, intermediate levels of
283 introgression on the Z chromosome rescue ovary defects in backcross females (Fig. 4C). This is
284 unexpected under dominance theory because intermediate levels of introgression would still expose
285 recessive Z-linked factors in females. Faster-Z theory has little support either: the *dN*/*dS* ratio is similar

between the Z chromosome and autosomes in *Heliconius* (*26*), and sequence divergence in *Papilio* is even greater on autosomes (Fig. S16).

Polygenic incompatibilities are, however, congruent with mechanisms based on asymmetric inheritance (*30*). Recall that f_Z and f_A are the introgressed ancestry fractions on the Z chromosome and on all autosomes, respectively. The key insights are:

i) f_Z is more skewed from f_A in F_1 females than in F_1 males;

ii) f_Z is much more variable than f_A in backcrosses;

iii) f_A is reduced in backcrosses compared to that in F_1 hybrids.

Since intermediate f_Z produces more normal phenotypes in backcrosses, a balancing process appears to exist between ancestry on autosomes and on the Z chromosome: phenotypes degrade when f_Z and f_A deviate from optimal balance. This process is ancestry-based and requires a multilocus genetic basis. Additionally, i) implies that F_1 females are more extremely unbalanced, generating Haldane's Rule; ii) implies that variation of imbalance is largely attributable to the Z chromosome, generating a large-Z effect; and iii) implies that the optimal Z-chromosome ancestry fraction might differ between F_1 and backcrosses.

Interpreting QTL results of highly polygenic traits

We find that QTL analysis of highly polygenic traits is shaped strongly by the statistical structure of ancestry tracts. Spurious major effect loci may appear in association mapping using crosses with long ancestry tracts even with large sample sizes. Here, "ghost" QTLs are likely to result from the cumulative influence of all polygenes linked to focal markers, and a peak in association strength does not result from a major single locus effect in development. This problem has been well recognized in genetic association theories despite a lack of universal solutions (*34*–*37*). Still, empirical analyses rarely consider this complexity because common software assumes only one or two QTLs per chromosome (*38*, *39*). The remedy we adopt here is to model explicitly the generating process of ancestry tracts (e.g., a crossover model), and to integrate information across all markers in picking the best-fitting architecture. Polygenic traits are common in humans (*40*), and between-species traits such as hybrid incompatibilities can be more complex due to greater genomic divergence. A stronger test for polygenicity needs to incorporate many more configurations of ancestry blocks than available in our backcrosses.

Polygenic incompatibility and global epistasis

Hybrid incompatibility is usually perceived as negative epistasis between species-specific mutations (*41*, *42*). While our phenomenological explanation invokes epistasis, the interaction may well be between ancestry fractions in different genomic regions. This is similar to global epistasis, where the phenotypic effect of a genetic change is somewhat independent of specific loci underlying the change (*43*). Global epistasis can emerge as a transformation of additive components. In our case, global

epistasis is manifested as *g* (transformation) on ancestry fraction (additive components). Global epistasis has been found previously in polygenic incompatibility. For instance, hybrid male sterility in *Drosophila* is sensitive not to precise genomic regions but to the total introgressed ancestry fraction (*44–46*). The molecular nature of polygenicity is unresolved. In our case, it is tempting to consider epigenetic mechanisms between autosomes and the Z chromosome. For instance, genetic variance of pupal weight in backcross males is much smaller than in females (Table S3). This is consistent, for instance, with dosage compensation in Lepidoptera in which both Z chromosomes in males are partially suppressed (*47*), which will dampen the effects of introgressed alleles. When we embarked on these studies of hybrid incompatibility in Lepidoptera, we hoped to locate key genes causing defects. To our surprise, it seems clear instead that sterility and other incompatibilities are often polygenic, and that this polygenicity can provide simple explanations for some hitherto mysterious rules of hybrid incompatibility.

References

1. J. A. Coyne, Two rules of speciation. *Speciation and its consequences*, 180–207 (1989).

2. J. B. S. Haldane, Sex ratio and unisexual sterility in hybrid animals. *J. Genet.* **12**, 101–109 (1922).

3. H. A. Orr, Haldane's rule. *Annu. Rev. Ecol. Syst.*, 195–218 (1997).

4. M. Turelli, H. A. Orr, Dominance, epistasis and the genetics of postzygotic isolation. *Genetics*. **154**, 1663–1679 (2000).

5. D. C. Presgraves, H. A. Orr, Haldane's rule in taxa lacking a hemizygous X. *Science*. **282**, 952–954 (1998).

6. M. Schilthuizen, M. Giesbers, L. W. Beukeboom, Haldane's rule in the 21st century. *Heredity* . **107**, 95–102 (2011).

7. L. F. Delph, J. P. Demuth, Haldane's rule: genetic bases and their empirical support. *J. Hered.* **107**, 383–391 (2016).

8. H. A. Orr, Haldane's rule has multiple genetic causes. *Nature*. **361**, 532–533 (1993).

9. D. C. Presgraves, Sex chromosomes and speciation in Drosophila. *Trends Genet.* **24**, 336–343 (2008).

10. D. C. Presgraves, Evaluating genomic signatures of "the large X-effect" during complex speciation. *Mol. Ecol.* **27**, 3822–3830 (2018).

11. E. L. Larson, S. Keeble, D. Vanderpool, M. D. Dean, J. M. Good, The composite regulatory basis of the large X-effect in mouse speciation. *Mol. Biol. Evol.* **34**, 282–295 (2017).

12. J. Kitano, J. A. Ross, S. Mori, M. Kume, F. C. Jones, Y. F. Chan, D. M. Absher, J. Grimwood, J. Schmutz, R. M. Myers, Others, A role for a neo-sex chromosome in stickleback speciation. *Nature*. **461**, 1079–1083 (2009).

13. D. E. Irwin, Sex chromosomes and speciation in birds and other ZW systems. *Mol. Ecol.* **27**, 3831–3851 (2018).

14. H. A. Orr, M. Turelli, Dominance and Haldane's rule. *Genetics*. **143**, 613 (1996).

15. B. Charlesworth, J. A. Coyne, N. H. Barton, The relative rates of evolution of sex chromosomes and autosomes. *Am. Nat.* **130**, 113–146 (1987).

16. C.-I. Wu, A. W. Davis, Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic bases. *Am. Nat.* **142**, 187–212 (1993).

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

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

19. B. Davies, E. Hatton, N. Altemose, J. G. Hussin, F. Pratto, G. Zhang, A. G. Hinch, D. Moralli, D. Biggs, R. Diaz, Others, Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. *Nature*. **530**, 171–176 (2016).

20. J. P. Masly, D. C. Presgraves, High-resolution genome-wide dissection of the two rules of speciation in *Drosophila*. *PLoS Biol.* **5**, e243 (2007).

21. Y. Tao, D. L. Hartl, Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. III. Heterogeneous accumulation of hybrid incompatibilities, degree of dominance, and implications for Haldane's rule. *Evolution*. **57**, 2580–2598 (2003).

22. D. C. Presgraves, Patterns of postzygotic isolation in Lepidoptera. *Evolution*. **56**, 1168–1183 (2002).

23. C. D. Jiggins, M. Linares, R. E. Naisbit, C. Salazar, Z. H. Yang, J. Mallet, Sex-linked hybrid sterility in a butterfly. *Evolution*. **55**, 1631–1638 (2001).

24. R. E. Naisbit, C. D. Jiggins, M. Linares, C. Salazar, J. Mallet, Hybrid sterility, Haldane's rule and speciation in *Heliconius cydno* and *H. melpomene*. *Genetics*. **161**, 1517–1526 (2002).

25. S. Kost, D. G. Heckel, A. Yoshido, F. Marec, A. T. Groot, A Z-linked sterility locus causes sexual abstinence in hybrid females and facilitates speciation in *Spodoptera frugiperda*. *Evolution*. **70**, 1418–1427 (2016).

26. N. Rosser, N. B. Edelman, L. M. Queste, M. Nelson, F. Seixas, K. K. Dasmahapatra, J. Mallet, Complex basis of hybrid female sterility and Haldane's rule in *Heliconius* butterflies: Z-linkage and epistasis. *Mol. Ecol.* **31**, 959–977 (2022).

27. S. A. Ae, A study of the *Papilio bianor* Group mainly based on hybridization (Lepidoptera, Papilionidae). *Tyo Ga*. **41**, 13–19 (1990).

28. H. Kitahara, K. Shirai, Crossing experiments with *Papilio okinawensis* Fruhstorfer from Okinawa Island and *P. dehaanii* C. & R. Felder from central Honshu, Japan (Lepidoptera, Papilionidae). *Lepid. Sci.* **69**, 85–91 (2018).

29. S. A. Ae, A study of hybrids between Japanese and Himalayan *Papilio* butterflies. *Spec. Bull. Lep. Soc. Jap* (1966).

30. M. Turelli, L. C. Moyle, Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics*. **176**, 1059–1088 (2007).

31. S. Lu, J. Yang, X. Dai, F. Xie, J. He, Z. Dong, J. Mao, G. Liu, Z. Chang, R. Zhao, Others, Chromosomal-level reference genome of Chinese peacock butterfly (*Papilio bianor*) based on third-generation DNA sequencing and Hi-C analysis. *GigaScience*. **8**, giz128 (2019).

32. Z. B. Zeng, Correcting the bias of Wright's estimates of the number of genes affecting a quantitative character: a further improved method. *Genetics*. **131**, 987–1001 (1992).

33. R. D. Ball, Bayesian methods for quantitative trait loci mapping based on model selection: approximate analysis using the Bayesian information criterion. *Genetics*. **159**, 1351–1364 (2001).

34. X. R. Maside, H. F. Naveira, A polygenic basis of hybrid sterility may give rise to spurious localizations of major sterility factors. *Heredity* . **77 (Pt 5)**, 488–492 (1996).

35. J. Wallin, M. Bogdan, P. A. Szulc, R. W. Doerge, D. O. Siegmund, Ghost QTL and hotspots in experimental crosses: novel approach for modeling polygenic effects. *Genetics*. **217**, iyaa041 (2021).

36. P. M. Visscher, C. S. Haley, Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models. *Theor. Appl. Genet.* **93**, 691–702 (1996).

37. J. Slate, From Beavis to beak color: a simulation study to examine how much QTL mapping can reveal about the genetic architecture of quantitative traits. *Evolution*. **67**, 1251–1262 (2013).

38. K. W. Broman, H. Wu, Ś. Sen, G. A. Churchill, R/qtl: QTL mapping in experimental crosses. *Bioinformatics*. **19**, 889–890 (2003).

39. K. W. Broman, D. M. Gatti, P. Simecek, N. A. Furlotte, P. Prins, Ś. Sen, B. S. Yandell, G. A. Churchill, R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics*. **211**, 495–502 (2019).

40. E. A. Boyle, Y. I. Li, J. K. Pritchard, An expanded view of complex traits: from polygenic to omnigenic. *Cell*. **169**, 1177–1186 (2017).

41. T. Dobzhansky, *Genetics and the Origin of Species* (Columbia University Press, New York, NY, 1937), *The Columbia Classics in Evolution*.

42. H. J. Muller, "Isolating mechanisms, evolution, and temperature" in *Biol. Symp.* (1942), vol. 6, pp. 71–125.

43. J. Diaz-Colunga, A. Skwara, K. Gowda, R. Diaz-Uriarte, M. Tikhonov, D. Bajic, A. Sanchez, Global epistasis on fitness landscapes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **378**, 20220053 (2023).

44. D. C. Presgraves, C. D. Meiklejohn, Hybrid sterility, genetic conflict and complex speciation: lessons from the *Drosophila simulans* clade species. *Front. Genet.* **12**, 669045 (2021).

45. H. F. Naveira, Location of X-linked polygenic effects causing sterility in male hybrids of *Drosophila simulans* and *D. mauritiana*. *Heredity*. **68 (Pt 3)**, 211–217 (1992).

46. M. A. Liénard, L. O. Araripe, D. L. Hartl, Neighboring genes for DNA-binding proteins rescue male sterility in *Drosophila* hybrids. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E4200–7 (2016).

47. L. F. Rosin, D. Chen, Y. Chen, E. P. Lei, Dosage compensation in *Bombyx mori* is achieved by partial repression of both Z chromosomes in males. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2113374119 (2022).

48. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j.* **17**, 10–12 (2011).

49. Picard toolkit. *Broad Institute, GitHub repository* (2019), (available at https://broadinstitute.github.io/picard/).

50. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, Others, Twelve years of SAMtools and BCFtools. *GigaScience*. **10**, giab008 (2021).

51. T. S. Korneliussen, I. Moltke, NgsRelate: a software tool for estimating pairwise relatedness from next-generation sequencing data. *Bioinformatics*. **31**, 4009–4011 (2015).

52. P. Rastas, Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics*. **33**, 3726–3732 (2017).

53. W. E. Castle, An improved method of estimating the number of genetic factors concerned in cases of blending inheritance. *Science*. **54**, 223 (1921).

54. S. Wright, *Evolution and the Genetics of Populations* (University of Chicago Press, 1968).

55. S. P. Otto, C. D. Jones, Detecting the undetected: estimating the total number of loci underlying a quantitative trait. *Genetics*. **156**, 2093–2107 (2000).

56. Between the two systems, f_A is more variable in *Heliconius* backcrosses with fewer autosomes (29 autosomes in *Papilio* and 20 autosomes in *Heliconius*), and there appear to be more detectable autosomal phenotypic effects in *Heliconius* backcross due to a larger variance in f_A (ref. (*26*): their Fig. 6a).

Acknowledgments

We thank Naomi Pierce and Adam Cotton for providing background information on the study system; Janet Sherwood, Shui Xu, Yuchen Zheng, Jinbo Hu, and Anastasios Kougionis for their assistance and knowledge in breeding/sourcing host plants and butterflies; Cassandra Extavour for discussing oogenesis; John Wakeley, Robin Hopkins, Liang Qiao, Sarah Dendy, Nathaniel Edelman, Shuzhe Guan, Fernando Seixas, and Yuttapong Thawornwattana for their intellectual support during the project.

Funding: The Quantitative Biology Initiatives at Harvard University (TX); The National Science Foundation-Simons Center for Mathematical and Statistical Analysis of Biology at Harvard University #1764269 (TX); The Department of Organismic and Evolutionary at Harvard University (TX, ST, JM); Sigma Xi Grant in Aid of Research (TX); Harvard University GSAS Student Council Summer Research Grant (TX)

Author contributions: Conceptualization: TX, JM; Methodology: TX, JM; Software: TX, NR; Formal analysis: TX; Investigation: TX, ST; Resources: XL, MY; Data curation: TX; Visualization: TX, ST; Funding acquisition: TX, JM; Project administration: TX; Supervision: JM; Writing – original draft: TX, ST, JM; Writing – review & editing: TX, ST, NR, XL, MY, JM

Competing interests: The authors declare no competing interests

Data and materials availability: Raw reads are released in the NCBI Sequence Read Archive (BioProject: PRJNA892033). Source data and code for main and supplementary figures will be deposited in Dryad and will be released at the time of publication. Source code is also available from: https://github.com/tzxiong/2022_Papilio_HybridIncompatibilityMapping