ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Complex basis of hybrid female sterility and Haldane's rule in *Heliconius* butterflies: Z-linkage and epistasis

Neil Rosser[1,2] | Nathaniel B. Edelman[1,3,4] | Lucie M. Queste[2] | Michaela Nelson[2] | Fernando Seixas[1] | Kanchon K. Dasmahapatra[2] | James Mallet[1]

[1]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, USA

[2]Department of Biology, University of York, York, UK

[3]Yale Institute for Biospheric Studies, Yale University, New Haven, Connecticut, USA

[4]Yale School for the Environment, Yale University, New Haven, Connecticut, USA

Correspondence
Neil Rosser and Nathaniel B. Edelman, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, USA.
Emails: neil.rosser@york.ac.uk (N.R.); edelmannate@gmail.com (N.B.E.)

Funding information
Natural Environment Research Council, Grant/Award Number: NE/K012886/1

Abstract

Hybrids between species are often sterile or inviable. Hybrid unfitness usually evolves first in the heterogametic sex—a pattern known as Haldane's rule. The genetics of Haldane's rule have been extensively studied in species where the male is the heterogametic (XX/XY) sex, but its basis in taxa where the female is heterogametic (ZW/ZZ), such as Lepidoptera and birds, is largely unknown. Here, we analyse a new case of female hybrid sterility between geographic subspecies of *Heliconius pardalinus*. The two subspecies mate freely in captivity, but female F1 hybrids in both directions of cross are sterile. Sterility is due to arrested development of oocytes after they become differentiated from nurse cells, but before yolk deposition. We backcrossed fertile male F1 hybrids to parental females and mapped quantitative trait loci (QTLs) for female sterility. We also identified genes differentially expressed in the ovary as a function of oocyte development. The Z chromosome has a major effect, similar to the 'large X effect' in *Drosophila*, with strong epistatic interactions between loci at either end of the Z chromosome, and between the Z chromosome and autosomal loci on chromosomes 8 and 20. By intersecting the list of genes within these QTLs with those differentially expressed in sterile and fertile hybrids, we identified three candidate genes with relevant phenotypes. This study is the first to characterize hybrid sterility using genome mapping in the Lepidoptera and shows that it is produced by multiple complex epistatic interactions often involving the sex chromosome, as predicted by the dominance theory of Haldane's rule.

KEYWORDS
Dobzhansky–Muller incompatibilities, Haldane's rule, hybrid sterility, lepidoptera, speciation, ZW sex determination

1 | INTRODUCTION

Hybrids between diverging populations may be sterile or inviable (Darwin, 1859; Presgraves, 2010). Because such postzygotic

incompatibilities are common between species, elucidating their genetic basis is seen as key to understanding speciation (Butlin et al., 2012; Castillo & Barbash, 2017; Coughlan & Matute, 2020; Nosil & Schluter, 2011). Hybrid dysfunction often results from epistatic

interactions among genes known as 'Dobzhansky–Muller incompatibilities' (Coyne & Orr, 2004; Dobzhansky, 1936; Muller, 1942). Under the Dobzhansky–Muller model, diverging populations acquire different alleles at two or more loci. In hybrids, previously untested combinations of alleles at different loci are brought together and interact to reduce fitness (Brideau et al., 2006; Maheshwari & Barbash, 2011; Orr, 1995; Presgraves, 2007; Tang & Presgraves, 2009).

Dobzhansky–Muller incompatibilities (DMIs) may involve a single pair of genes (Sweigart et al., 2006), but they are more likely to be complex, even early in speciation (e.g. Kalirad & Azevedo, 2017; Phadnis, 2011). This is because the expected number of two-locus DMIs is predicted to increase approximately as the square of the number of divergent substitutions between species: the 'snowball effect' (Matute et al., 2010; Orr, 1995; Orr & Turelli, 2001). Furthermore, DMIs involving more than two loci should accumulate even more rapidly, because, as the number of interacting loci increases, so too does the number of potentially negative combinations (Orr, 1995). In keeping with these predictions, widespread DMIs across the genomes of a number of species have been inferred from genetic association data (Good et al., 2008; Schumer et al., 2014). There is also evidence that polymorphic alleles with negative epistatic interactions are common even within (Corbett-Detig et al., 2013).

One of the few generalizations about speciation is Haldane's rule (Haldane, 1922), which states that among hybrids, when one sex is absent, rare or sterile, it is usually the heterogametic sex (males in XX/XY systems and females in ZZ/ZW systems). Greater sterility of the heterogametic sex has been found in 213 of 223 pairs (>95%) of a diverse array of taxa, and has at least 10 phylogenetically independent origins (Delph & Demuth, 2016; Schilthuizen et al., 2011). The ubiquity of Haldane's rule therefore suggests that postzygotic incompatibilities evolve with some predictability across a wide range of taxa (Coyne, 1992). Hybrid sterility of the heterogametic sex also evolves early, typically before hybrid inviability (Coyne & Orr, 1989a, 1997; Presgraves, 2002, 2010). It may therefore have a disproportionate role in reducing gene flow, and as such is of particular interest for understanding speciation (Coughlan & Matute, 2020; Ramsey et al., 2003).

Most explanations for Haldane's rule depend on DMIs. The hypothesis to have received the most support is dominance theory in which hybrid sterility and inviability are produced by interactions between the sex chromosomes and autosomes (Coyne & Orr, 2004). In the homogametic sex of hybrids, sex-linked alleles produce incompatibilities only if dominant, whereas in heterogametic hybrids both dominant and recessive sex-linked alleles can cause incompatibilities. If alleles causing incompatibilities are on average recessive, the heterogametic sex is expected to suffer more than the homogametic sex (Orr, 1997; Turelli & Moyle, 2007; Turelli & Orr, 1995). Nonetheless, male heterogametic species without strongly differentiated sex chromosomes also conform to Haldane's rule (Presgraves & Orr, 1998), suggesting that other forces also contribute, such as 'faster-male' evolution (Wu & Davis, 1993). The genetic and molecular mechanisms of hybrid sterility have been identified in some cases (Bayes & Malik, 2009; Brideau et al., 2006; Mihola et al., 2009; Schartl, 2008; Tang & Presgraves, 2009), but this work has been primarily carried out in organisms with XX/XY sex determination in which male hybrids are sterile or inviable.

Lepidoptera (butterflies and moths) yielded the first example of a sex-linked trait (Doncaster & Raynor, 1906), even before *Drosophila* (Morgan, 1910, 1911). Lepidoptera are also among the groups of taxa Haldane considered when formulating his eponymous rule (Haldane, 1922). They have ZW/ZZ sex determination, where females are the heterogametic sex and, in accordance with Haldane's rule, are more susceptible to hybrid dysfunction (Presgraves, 2002). As such, they are critical in evaluating the relative impact of dominance and faster male evolution in Haldane's rule and have provided evidence that faster-Z evolution may contribute to the phenomenon in female heterogametic systems (Prowell Pashley, 1998; Sackton et al., 2014).

Several examples of Haldane's rule have been reported in *Heliconius* butterflies (Nymphalidae), which comprise about 48 species that occur throughout much of tropical America (Jiggins, 2017). Female hybrid sterility has been observed in crosses between *Heliconius cydno* (*sensu lato*) and *Heliconius melpomene* (Naisbit et al., 2002; Salazar et al., 2005; Sánchez et al., 2015), and also between geographically distant subspecies of *Heliconius melpomene* (Jiggins et al., 2001). Here, we investigate the genetic basis of Haldane's rule in hybrids between two subspecies of *Heliconius pardalinus*: *H. pardalinus butleri* and *H. pardalinus sergestus*. These largely allopatric subspecies are strongly genetically differentiated, with *H. p. butleri* more closely related over most of its genome to its sympatric relative *Heliconius elevatus*, thereby rendering *H. pardalinus* paraphyletic (Rosser et al., 2019). They inhabit different habitats, with *H. p. sergestus* restricted to dry forests in the Huallaga/Mayo valleys of the Andes, and *H. p. butleri* inhabiting lowland rainforest across the adjacent Amazon basin (Figure 1). Although they mate freely in captivity, they rarely co-occur in nature, and F1 hybrid females in both directions of cross are completely sterile (Rosser et al., 2019). Here, we characterize the ovary phenotype in parental populations, F1 hybrids and backcrosses. We use backcrosses to *H. p. butleri* to generate a QTL map and intersect these data with genes differentially expressed between fertile and sterile individuals, to identify candidate genes and epistatic interactions responsible for hybrid sterility.

2 | MATERIALS AND METHODS

2.1 | Butterfly rearing, nucleic acid preservation and ovary dissection

Captive populations of *H. p. sergestus* and *H. p. butleri* were established in insectaries in Tarapoto, Peru, as previously described (Rosser et al., 2019). The majority of the stocks came from near Tarapoto (within the white box in Figure 1), but a small number came from Pucallpa, Ucayali. Female butterflies were collected from insectaries 15 days after eclosion, allowing time for eggs to develop fully (Dunlap-Pianka et al. 1977; Naisbit et al., 2002). During this time, they were isolated from males to ensure all were unmated. Wings were removed and stored in glassine envelopes as vouchers. Thorax and head were removed and stored in NaCl-saturated dimethyl sulfoxide at −20°C for DNA extraction and processing. Approximately

FIGURE 1 Distribution of *H. pardalinus* in Peru. The yellow dots correspond to collection localities of *H. pardalinus sergestus* and the red dots to *H. pardalinus butleri*, which intergrades with other subspecies in central and southern Peru and the Amazon basin. Geographic data are from Rosser, Phillimore, Huertas, Willmott and Mallet (2012); Rosser et al. (2019). Samples used for population genomic analysis were all collected near Tarapoto within the white box. The majority of the stocks used to found captive populations were also collected in this area, but with a small number of individuals from Pucallpa
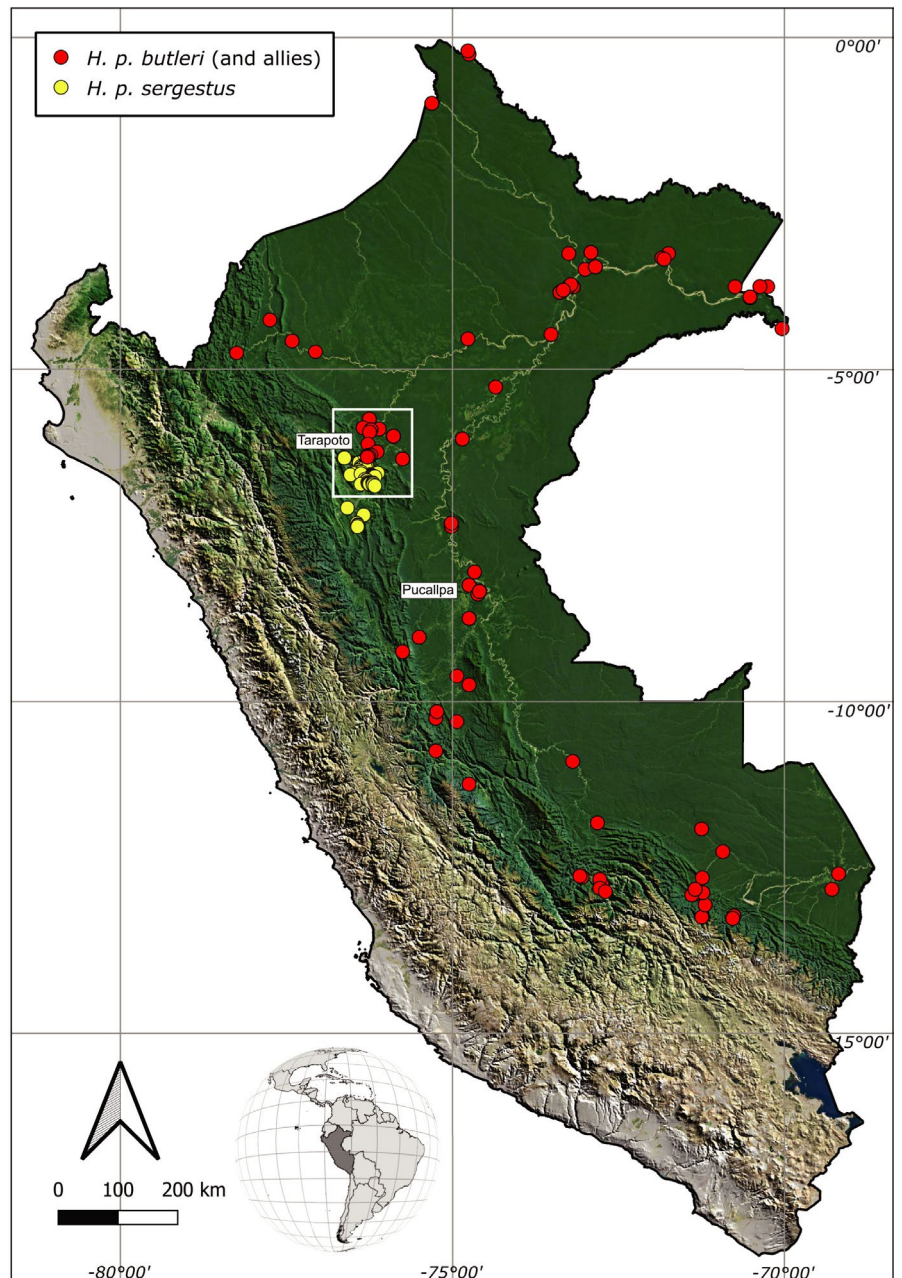
half of the ovaries were dissected immediately, and for the remainder, abdomens were stored in 96% ethanol and transported to the laboratory for fine dissection. In all cases, ovaries were dissected from the abdomen in ice-cold phosphate-buffered saline (PBS) using fine forceps and insect pins. Tracheae and fat bodies were removed manually, and images were taken at 8×, 12.5× and 20× magnification for phenotyping. Of the ovaries dissected in the field, six backcrosses and two pure *H. pardalinus butleri* were stored in RNALater solution for RNAseq (Thermo Fisher AM7020).

2.2 | Ovary staining and phenotyping

For every dissection, we scored the developmental progress of ovaries on a scale of 0 (empty ovaries) to 3 (containing fully developed yolky eggs) based on gross morphology (Figure S1, and see examples

in Figure 2). Three images from each ovary were scored blind by two independent scorers. The resulting six scores per ovary were averaged to yield a single fertility score for each individual.

In a subset of samples, we characterized the earliest arrested developmental stages of oocytes through nuclear staining with DAPI, using the stages described in the silkmoth (*Bombyx mori*) as a reference (Figure 3). Individual ovarioles from alcohol-stored ovaries were removed and rehydrated by 15-min incubations in serial dilutions of ethanol in 0.1% Tween-20 in 1X phosphate-buffered saline (PBT) (Ethanol concentrations: 95%, 90%, 80%, 60%, 40%, 20%, 0%). Once fully rehydrated, ovaries were incubated in acridine orange solution (Thermo Fisher A1301; 5 g/ml in PBT) to visualize cytoplasm. They were then washed in PBT before being stained with DAPI (1 L/ml in PBT), washed once more in PBT and mounted on slides with VectaShield (Vector Labs). Slides with stained ovarioles were scanned with a Zeiss Axio Scan Z1,
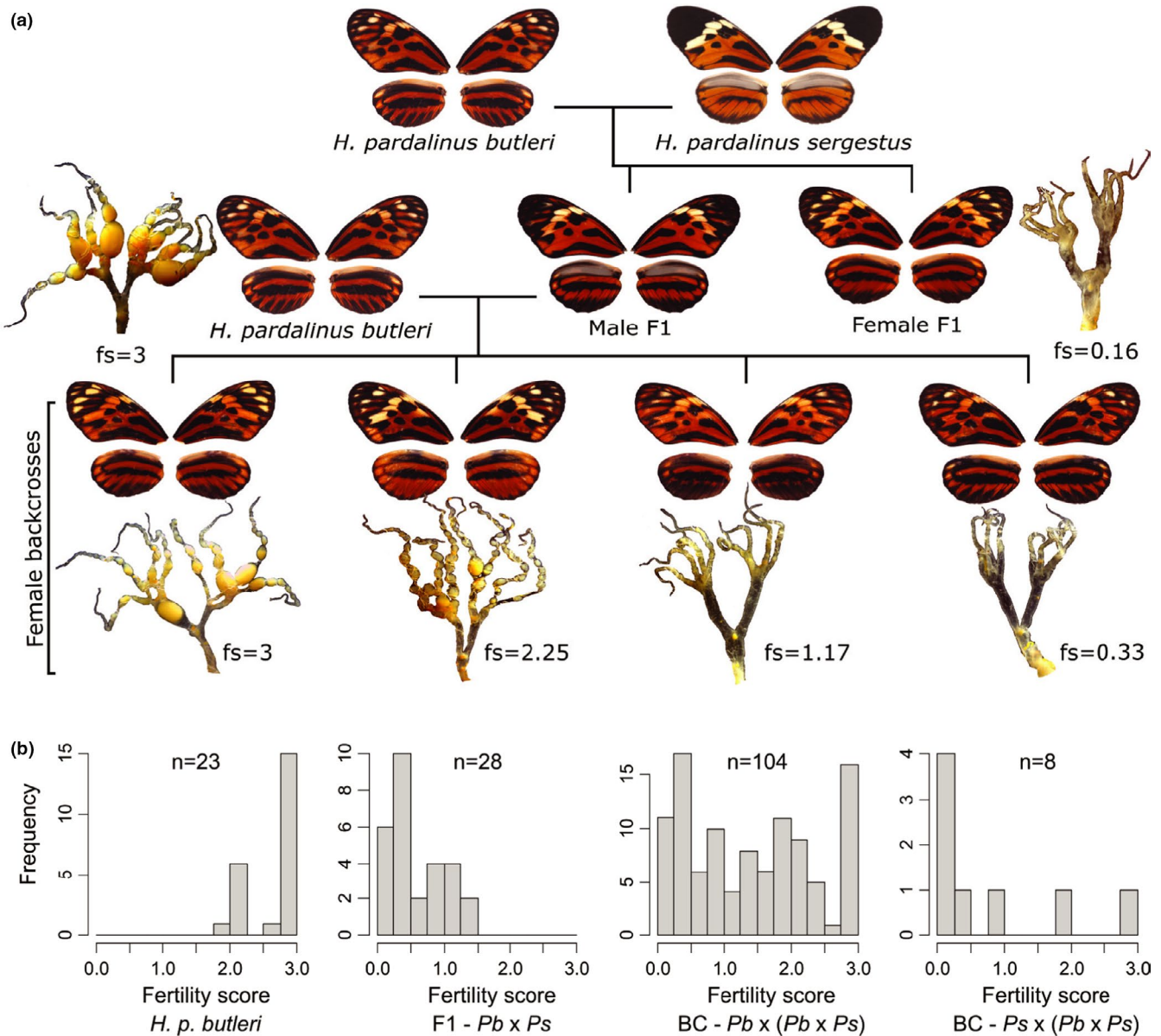
(a)

(b)

FIGURE 2 Crossing scheme and distribution of phenotypes. (a) Crossing *H. p. butleri* with *H. p. sergestus* in either direction produces sterile female F1s, while male F1s are fertile. Backcrossing these males in either direction produces females with variable fertility. Example wing phenotypes and dissected ovaries for backcrosses to *H. p. butleri* are shown, with fertile individuals to the left and sterile individuals to the right; fs = fertility score assigned to the dissected ovary. (b) Histograms of ovary fertility scores for (i) *H. p. butleri* females, (ii) F1s produced by mating female *H. p. butleri* (*Pb*) with male *H. p. sergestus* (*Ps*), (iii) backcrosses produced by mating fertile male F1s (*Pb* × *Ps*) with *H. p. butleri* females (*Pb*), and backcrosses produced by mating fertile male F1s (*Pb* × *Ps*) with *H. p. sergestus* females (*Ps*). Fertility scores for *H. p. sergestus* females (not shown) ranged from 2.5–3 (*n* = 3)

and high-magnification images were taken with a Zeiss LSM 880 upright confocal microscope. The most highly developed follicle in each ovariole was staged through visual comparison to oocyte development stages described in *Bombyx mori* (Yamauchi & Yoshitake, 1984).

2.3 | DNA extraction and sequencing

RNA-free genomic DNA was extracted from individuals used in QTL mapping (see below) using a Qiagen DNeasy Blood and Tissue Kit and following the manufacturer's standard protocol. Restriction site

FIGURE 3 Developing oocytes. (a) Idealized developing follicle stages (Yamauchi & Yoshitake, 1984) (b) Brightfield and confocal images of DAPI-stained ovaries. Each row displays an overview image, as well as individual follicles at indicated stages from the same ovary. Scale bars for ovariole overviews are shown below the relevant column. Scale bar for stages 3–5 is shown below the stage 3 column, except where indicated in image. 'Not observed' represent stages not present in the illustrated ovariole. In ovary images, one ovariole (OV) and the oviduct (OD) are indicated. Individual follicles are encircled by dashed lines. Where visible, one nurse cell nucleus (NC) and the oocyte cell nucleus (OC) in the highlighted follicle are indicated

(a)

(b)
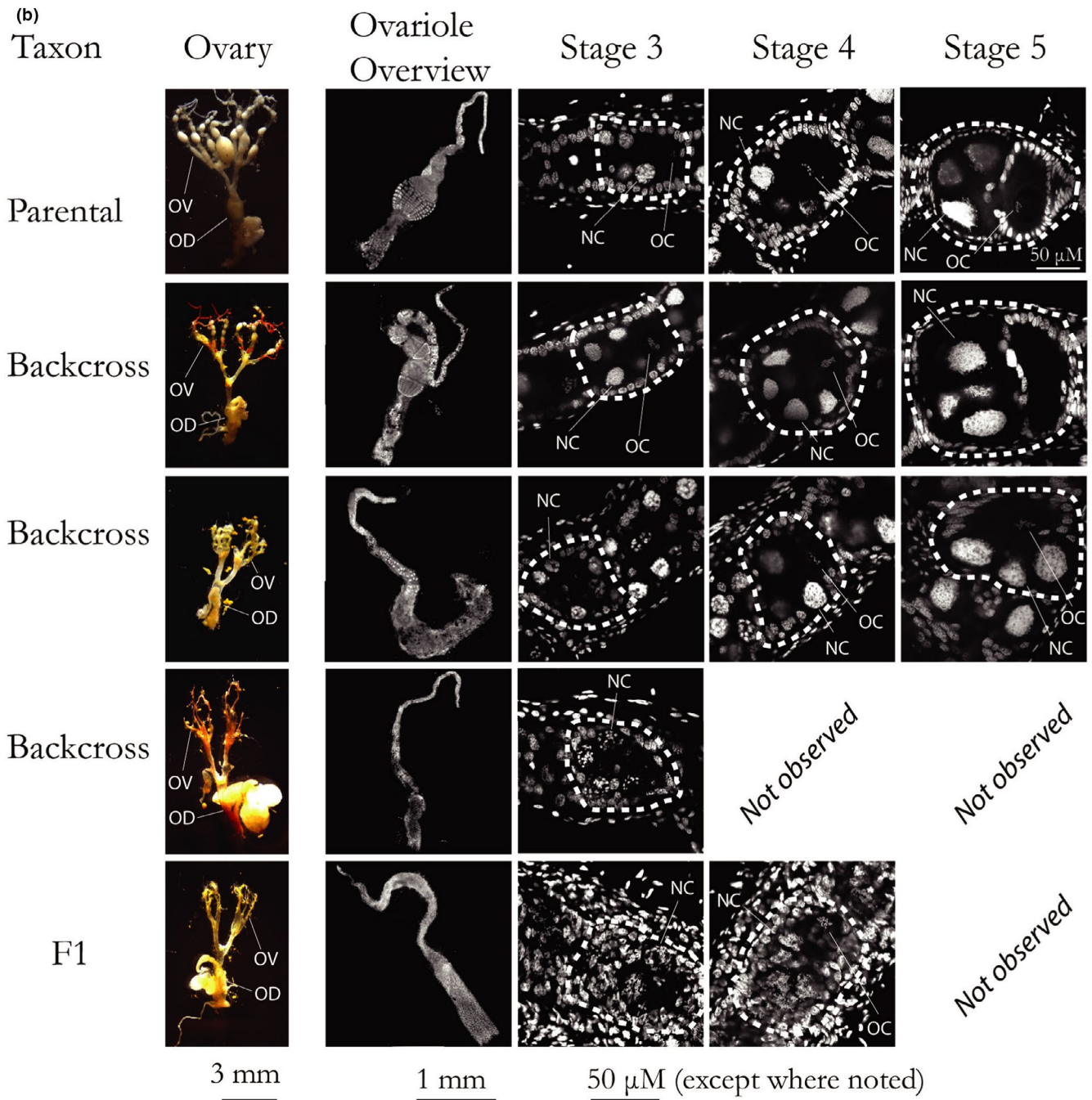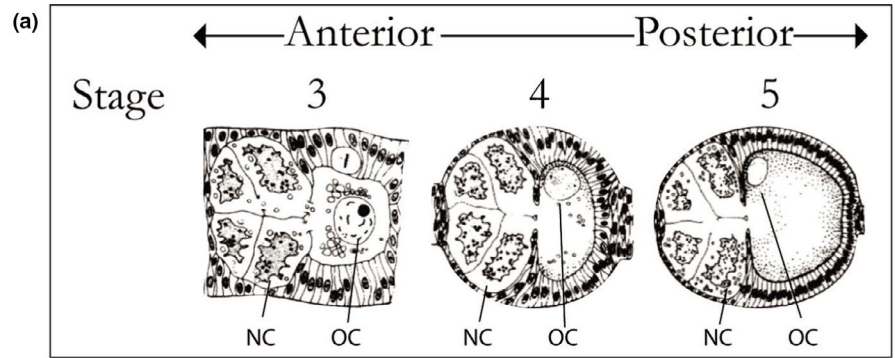
associated DNA (RADSeq) libraries were prepared using a protocol modified from Etter et al. (2011), Hoffman et al. (2014), using a *Pst*I restriction enzyme, sixteen 6 bp P1 barcodes and eight indexes. DNA was Covaris sheared and gel size selected to 300–700 bp. 128 individuals were sequenced per lane, with 125 bp paired-end reads, on an Illumina HiSeq 2500.

2.4 | SNP calling

FastQ files from each RAD library were demultiplexed using process_radtags from Stacks (Catchen et al., 2013), and BWA-MEM (Li, 2013) was used with default parameters to map the reads both to the *H. melpomene* genome (Hmel2.5) (Davey et al., 2017) and to the *H. pardalinus* genome (Hpar) (Seixas et al., 2021). BAM files were then sorted and indexed with SAMtools (Li et al., 2009), and Picard-tools v 1.119 (https://github.com/broadinstitute/picard) was used to add read groups and mark PCR duplicates. To check for incorrectly labelled samples, we estimated the sex of a sample by dividing the mean number of reads per kilobase on the Z chromosome by the mean value for autosomes. This returned a value close to 1 in males and 0.7 in females, which can then be compared with the recorded sex of the sample. To further check for labelling errors, confirm pedigrees and assign samples with unrecorded pedigree to families, we used Plink 1.9 (Chang et al., 2015) to estimate the fraction of the genome that is identical by descent (IBD; $\hat{\pi}$) between all pairwise combinations of samples (siblings and parent–offspring comparisons should yield $\hat{\pi}$ values close to 0.5). In addition, for specimens that were sequenced multiple times in order to improve coverage, we checked that samples were derived from the same individual (with $\hat{\pi}$ values close to one). We then merged these samples, using the MergeSamFiles command from Picard tools, and used Samtools' mpileup command to call single nucleotide polymorphisms (SNPs) for linkage map construction.

2.5 | Linkage map construction

Linkage maps were built using reads aligned to each of the reference genomes using Lep-MAP3 (Rastas, 2017). The ParentCall2 module was used to correct erroneous or missing parental genotypes, and call sex-linked markers using a log-odds difference of >2. We used Filtering2 to remove SNPs showing segregation distortion, specifying a p-value limit of 0.01 (i.e. there is a 1:100 chance that a randomly segregating marker is discarded). Because we genotyped only female offspring, we did not filter sex-linked markers for segregation distortion. We then used SeparateChromosomes2 to cluster markers to linkage groups, specifying zero recombination in females (in Lepidoptera, females meiosis is achiasmatic) and joining pairs of markers with LOD score greater than 14. To obtain recombination distances between markers, we fixed the order of the markers to their order on the Hmel2.5 or Hpar genome assemblies, and then evaluated this order, again using paternally and dual informative markers. Lep-MAP3 outputs fully informative and phased genotypes with no missing data, which can be used for QTL mapping.

2.6 | QTL mapping

Genetic data were analysed as backcrosses (Figure 2) using the paternally inherited allele. We used R/QTL (Broman et al., 2003) to estimate genotype probabilities at 1-cM intervals, using the Haldane mapping function and an assumed genotyping error rate of 0.001. Loci with inferred genotypes were labelled using the chromosome and the centimorgan position. We used Haley–Knott (H-K) regression to test for associations between the estimated genotype probabilities at each marker and fertility score (Haley & Knott, 1992). BB genotypes were coded as 0.5 and BS genotypes were coded as −0.5, where B is the *H. p. butleri* allele and S is the *H. p. sergestus* allele.

We first built a single locus additive QTL model at each position in the genome (H_1; $y = \mu_1 + \beta_1 q_1 + \varepsilon$) and calculated the $\log_{10}$ likelihood ratio (LOD score) comparing (H_1) with the null hypothesis of no QTL (H_0; $y = \mu_1 + \varepsilon$). μ is the intercept and can be interpreted as the average phenotype, β is the slope and can be interpreted as the difference between the phenotypes for the two alternative genotypes, and ε is the error term. To identify loci that act in combination to produce the phenotype, we then estimated LOD scores using all pairwise combinations of typed markers and inferred genotypes at 1-cM intervals across the genome, while allowing interactions between them (H_f; $y = \mu_1 + \beta_1 q_1 + \beta_2 q_2 + \beta_3 q_1 q_2 + \varepsilon$). The difference between LOD values for (H_f) and the corresponding two-locus additive model (H_a; $y = \mu_1 + \beta_1 q_1 + \beta_2 q_2 + \varepsilon$) gives the improvement in fit attributable purely to interactive effects (H_{int}). The difference between LOD_f and the maximum LOD value obtained from single QTL locus models at either marker indicates the presence of a second QTL, allowing for epistasis (H_{fv1}). We also performed these analyses while controlling for kinship. To do this, we used LepMap to estimate $\hat{\pi}$ (IBD) between all individuals. We then created a variance–covariance matrix of genetic relatedness and included this in our models as a random effect. Significance of QTL scans was assessed by permuting the phenotypes relative to the genotypes (10,000 permutations). Because we analysed only female backcrosses, the degrees of freedom for QTL models at sex-linked and autosomal loci are the same, and so we set a single genome-wide significance threshold for each scan.

2.7 | Population genomics

To examine genomic differentiation between the *H. p. sergestus*, *H. p. butleri* and *H. elevatus*, previously published whole-genome re-sequencing data (four individuals each taxon) were used (NCBI accession numbers: ERS977715; ERS235668; ERS977716; ERS977717; SRR3102338; ERS4368506; ERS4368505; ERS4368504; ERS070236; ERS977673; ERS977674; ERS070238). All of these individuals were collected within a small geographic area on either side of the Cordillera Escalera near Tarapoto (see the

white box in Figure 1). Raw reads were filtered for Illumina adapters using cutadapt (Martin, 2011) and mapped to the Hmel2.5 (Davey et al., 2017; Seixas et al., 2021) genomes using BWA-MEM v0.7.15. Duplicate reads were removed using sambamba v0.6.8 (Tarasov et al., 2015), and the Genome Analysis Toolkit (GATK) v3.8 RealignerTargetCreator and IndelRealigner modules (Depristo et al., 2011; McKenna et al., 2010) were used to realign reads around indels. Genotype calling was performed for each taxon separately with bcftools (Li et al., 2009) mpileup and call modules (Li, 2011), using the multiallelic and rare-variant calling option (-m) and requiring a minimum mapping quality and base quality of 20. Genotype calls were required to have a minimum quality score (QUAL) of 20, RMSMappingQuality (MQ) ≥ 20, genotype quality (GQ) ≥ 20 and a minimum individual depth of coverage (DP) ≥ 8 (or DP ≥ 4 for the Z chromosome of females). Genotypes within 5 bp of an indel were recorded as missing data.

Differentiation (F_{ST}), pairwise genetic distances (D_{XY}) and nucleotide diversity (π) between the three taxa studied were estimated along the genome in overlapping 25 kb windows (with 5 kb steps) using the popgenWindows.py script (available from https://github.com/simonhmartin/genomics_general). To test for faster evolution on the Z chromosome, we used the extractCDSAlignments.py python script to get sequence alignments of transcripts from VCF files, based on the *H. melpomene* reference gene annotation. Only alignments with at least 300 bp were considered, sequences containing stop codons were filtered out, and codons with missing data at any base were masked. For each transcript, we calculate the proportion of nonsynonymous (*dn*) and synonymous (*ds*) substitutions between all possible pairwise comparisons of *sergestus-butleri* sequences, following the method of Li (1993), as implemented in the 'seqinr' package in R (Charif & Lobry, 2007). The average of these two values is then used to calculate the *dn/ds* ratio per transcript.

2.8 | RNA extraction and sequencing

Previtellogenic (i.e., before yolk deposition) follicles ($n = 8$) were dissected from ovaries stored in RNALater. These included six backcrosses with scores ranging from 0.6–3 and two pure *H. p. butleri* individuals. Only previtellogenic follicles were retained for RNA sequencing. Tissue was blotted dry with Kimwipes to remove excess RNALater solution, transferred to TRIZOL and homogenized with the PRO200 tissue homogenizer (PRO Scientific). RNA was extracted with the Direct-zol RNA miniprep kit (Zymo R2051). The mRNA libraries were prepared by the Harvard University Bauer Core with the KAPA mRNA HyperPrep kit, with mean fragment insert sizes of 200–300 bp, and were sequenced on a NovaSeq S2, producing an average of 49 million paired-end, 50 bp reads per library (Table S1).

RNASeq reads were mapped to the *H. melpomene* v2.5 transcriptome (Pinharanda et al., 2019) using kallisto (Bray, Pimentel, Melsted & Pachter, 2016). Approximately 70% of reads were mapped to the transcriptome per sample and that value did not differ between the *H. pardalinus butleri* samples and the backcrosses

(Table S2). Aligned reads were normalized to account for sequencing coverage, transcript length and RNA composition using sleuth (Pimentel, Bray, Puente, Melsted & Pachter, 2017). Raw counts were log-transformed, and expression differences were calculated by comparing the likelihood of the model: ln(counts) ~ 1 to the model ln(counts) ~ 1 + binaryscope (Pimentel et al., 2017).

In order to identify conserved genes expressed in butterfly oogenesis, we used BLAST to identify *H. melpomene* transcripts orthologous to genes expressed in the ovarian transcriptome of the Speckled Wood butterfly *Parage aegeria* (Carter et al., 2013). In addition, we used OrthoFinder (Emms & Kelly, 2019) to identify transcripts with orthologous genes in *Drosophila melanogaster* and then filtered this list with the keywords 'oogenesis' OR 'follicle' OR 'nurse' OR 'oocyte' using the phenotypic data on Flybase (http://flybase.org). Finally, for each of our strongest candidate genes, we then tested whether pure *H. butleri* individuals have levels of expression consistent with fertile backcrosses and divergent from sterile backcrosses.

3 | RESULTS

We reared 143 F1 hybrid offspring of *H. p. butleri* and *H. p. sergestus*. 29 of these (17 females, 12 males) had a *H. p. sergestus* mother and a *H. p. butleri* father, and 101 (49 females, 52 males) had a *H. p. butleri* mother and a *H. p. sergestus* father (totals sum to less than 143 due to missing data for some individuals). Female F1s in both directions of cross were sterile (Figure 2b, Rosser et al., 2019). To investigate the genetic basis of hybrid sterility between the two populations, we backcrossed fertile F1 hybrid males to both parental species, rearing 320 offspring. Of these, 282 (150 females, 132 males) had a *H. p. butleri* mother and 26 (10 females, 16 males) had a *H. p. sergestus* mother. The approximately equal sex ratios in both F1 and backcross broods suggests a lack of sex bias in immature stage viability.

Almost all individuals from parental populations contained developing follicles that reached the final stages of vitellogenesis, and most had fully developed eggs ($n = 12/13$, with 9/10 *H. p. butleri* and 3/3 *H. p.sergestus*). However, ovaries of F1 hybrids seemed devoid of developing oocytes (Figure 2). Female backcrosses with *H. p. butleri* mothers yielded an approximately bimodal distribution of ovary phenotypes (Figure 2b), while a small sample of backcross females ($n = 8$) to *H. p. sergestus* exhibited a skewed distribution, with mostly sterile individuals (Figure 2b). Logistical constraints prevented a larger sample size in this direction of backcross.

All F1 and backcross individuals had early-stage follicles, but sterile individuals showed arrested development after oocytes reached approximately stage 3. This stage marks a developmental timepoint after oocyte vs. nurse cell differentiation and follicle formation, but before vitellogenesis (yolk deposition) (Büning, 1994; Yamauchi & Yoshitake, 1984). Using the 42 individuals for which we could confidently assign the latest developmental stage and fertility score, we verified that the two metrics were highly correlated (logistic regression, $p = 2.45 \times 10^{-11}$, Figure S2).

3.1 | QTL mapping

We sequenced 87 females from 7 families produced by backcrossing F1 males to *H. p. butleri* females. Using RADSeq reads aligned to Hmel2.5 reference genome, the linkage map for these individuals comprised 124,456 single nucleotide polymorphism (SNPs) across 21 chromosomes, with a total map length of 1106.95 cM. The linkage map built using reads aligned to the Hpar reference genome contained 29% more SNPs (Appendix S1, Table S1 and Figures S3–S6). However, because the Hpar genome is not annotated and QTL results using Hmel2.5 and Hpar were very similar (Tables 1 and S3), we report the results using Hmel2.5 in the main text. We validated this approach by mapping QTL for a colour pattern phenotype in these individuals, and identified a single, strong signal overlapping the well-known *Heliconius* colour pattern gene *Cortex* (Nadeau, 2016, Figure S7).

Scanning the genome for additive, single QTLs associated with fertility score (H_1) revealed a broad central region on the Z chromosome (Figures 4, S8, Tables 1, S3). The maximum LOD value was observed at 29.2 cM (Figure 4b,c), with mean predicted fertility scores of 1.81 for the *H. p. butleri* allele and 0.93 for the *H. p. sergestus* allele ($R^2 = 0.20$). The *H. p. butleri* allele had higher predicted fertility scores than the *H. p. sergestus* allele all along the Z chromosome, but the difference declined to nearly zero towards the distal end of the chromosome.

When scanning for interacting QTLs, we identified a negative interaction between a pair of loci at opposing ends (~5 cM and ~55 cM) of the sex chromosome, with the full epistatic model explaining 54% of the variance in fertility score (Figure 5, Tables 1 and S3). This pair of loci was highly significant ($p < 0.001$) irrespective of whether we tested the combined additive effects and interaction (H_f), the additive effect of the second locus plus the interaction (H_{fv1}) or the interaction alone (H_{int}), and was robust to family-specific effects (Figure S8) and the reference genome used. Recombinant Z chromosomes (Z_{BS} or Z_{SB}) had higher fitness (i.e. greater average fertility scores) than either nonrecombinant (Z_{BB} or Z_{SS}) (Figure 5d). When testing H_f,

we also identified possible associations between the Z chromosome and chromosomes 4 ($p < 0.05$ using Hpar, $p < 0.1$ using Hmel2.5), 12 ($p < 0.1$) and 15 ($p < 0.01$) (Figure 5a, Tables 1 and S3). The confidence intervals for the Z-linked QTLs in these pairs were wide but all broadly overlapping, and encompassed the single additive QTL at 29.21 cM (see column 10 in Tables 1 and S3). We then tested for the single QTL at 29.2 cM on the sex chromosome while controlling for the epistatically interacting pair of QTLs at either end. It remained significant, but its position shifted slightly to 33.86 cM. Bringing these three QTLs together in a single model ($y = \mu_1 + \beta_1 q_1 + \beta_2 q_2 + \beta_3 q_3 + \beta_4 q_1 q_2 + \varepsilon$) explained 62% of the variance in fertility score.

To understand these results further, we divided individuals into four groups depending on their genotypes at the two interacting loci on the Z chromosome (Z_{BB}, Z_{SS}, Z_{BS}, Z_{SB}). For each of these groups, we then plotted fertility against the fraction of the autosomes homozygous for *H. p. butleri* alleles (*B/B*) (Figure 6a). We hypothesized that if sterility is driven by interactions between the Z chromosome and autosomes, this fraction should be positively correlated with fertility score for those individuals holding a Z_{BB}. As expected, for Z_{BB} individuals, we found a significant positive correlation between the proportion of autosomal markers derived from *H. p. butleri*. Interestingly, we also found a significant negative correlation for Z_{SB} individuals. We then conducted QTL mapping on each of these groups. For individuals with a recombinant Z_{SB} chromosome, we identified a significant interaction ($LOD_{int} = 6.97$, $p < 0.01$, $R^2 = 0.79$) between loci at 9.3 cM on chromosome 8 and 11.9 cM on chromosome 20 (Figure 6b–d). No significant QTLs were detected for the other subgroups (Z_{BB}, Z_{SS} and Z_{BS}).

3.2 | Population genomics of the Z chromosome

Nucleotide diversity (π) in *H. p. sergestus* was low along most of the Z chromosome, but higher in *H. p. butleri* and *H. elevatus*, which were near identical (Figure 7a). Pairwise genetic differentiation (D_{xy}) was very similar between all three taxa, barring a 250 kb region in the

TABLE 1 Summary of significant single-locus (H_1) and two-locus (H_f) QTL models (using reads aligned to Hmel2.5)

LOD_f	chr	QTL1 marker	cM	limits (cM)	limits (physical)	chr	QTL2 marker
4.21**	Z	Hmel221001o:7109812	29.21	10.49–35.02	4,710,269–7,966,755		
6.72+	4	Hmel204001o:9208991	49.03	25.72–49.03	6,102,423–7,036[†]	Z	Hmel221001o:7109812
6.65+	12	c12.loc25	25	17.47–55.93	5,100,627–16,319,705	Z	Hmel221001o:5752071
7.87**	15	Hmel215003o:8041294	38.48	10–44.3	4,230,382–9,907,975	Z	Hmel221001o:5752071
14.52***	Z	Hmel221001o:3045330	4.65	2.33–5.81	2,179,644–4,463,341	Z	Hmel221001o:10565964
7.76*	8	Hmel208001o:1005579	9.3	8.14–10	618,400–1,390,337	20	Hmel220003o:5817143

Note: The first column gives the LOD value of the full model (H_f), with the significance estimated by permutation (+$p < 0.1$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). The next columns are the chromosome and QTL marker (scaffold and median physical position within the peak). The centimorgan limits are the Bayesian credible intervals, and the physical limits are the nearest typed flanking markers of that interval (all physical limits were on the same scaffold as the QTL peak, except for the chromosome 4 interaction † with Z, which was on scaffold Hmel204003 of Hmel2.5). The final five columns give the parameter estimates and R^2 of the model. $\beta_1 q_1$ and $\beta_2 q_2$ are the estimated additive effects for the QTLs, that is, the difference between the average fertility scores for the alternative genotypes, and $\beta_3 q_1 q_2$ is the coefficient for the interaction between the 2 loci. Model coefficients comprise the estimated value, the standard error and the significance (thresholds as above). The significant interaction between chromosome 8 and 20 was detected using individuals holding a Z_{SB} chromosome only.
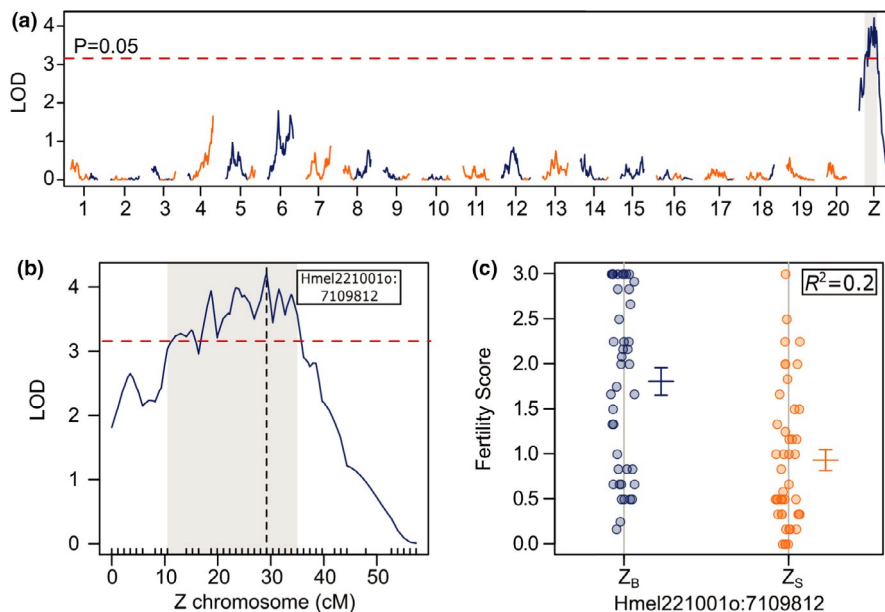
FIGURE 4 Single QTL analysis. (a) LOD values at each marker across the genome, calculated using H-K regression and with reads aligned to Hmel2.5. The red dashed line indicates the genome-wide significance threshold ($p < 0.05$; 10,000 permutations), and the grey shaded area the Bayesian credible intervals for the peak on Z. Lines are coloured depending on whether the *H. p. butleri* allele (blue) or the *H. p. sergestus* allele (yellow) had higher fertility. (b) Enlargement of Z chromosome, with the QTL peak at 29.21 cM indicated by the vertical dashed line (corresponding to physical position Hmel221001o:7109812). (c) Fertility scores at the QTL peak Z markers are hemizygous and coded by a single letter (B = *H. p. butleri* and S = *H. p. sergestus*) and explain 20% of the variance in fertility score. Errors bars are standard errors

centre of the Z chromosome (6.5–6.75 MB) between *H. p. butleri* and *H. elevatus*, where it dropped close to zero (Figure 7b). This region was also characterized by high F_{ST} between *H. p. sergestus* and *H. p. butleri*, which falls in the centre of the additive QTL peak (Figure 7c). F_{ST} was generally elevated at the ends of the Z chromosome as well, possibly due to the two epistatic QTLs; however, these regions also have low rates of recombination (Figure 7c), which can lead to high F_{ST} values even in the absence of selection burri, 2017. Overall, the mean F_{ST} between *H. p. butleri* and *H. p. sergestus* for the Z chromosome was 0.37, making it the most divergent chromosome. The autosomes ranged from 0.23 (chromosome 3) to 0.35 (chromosome 19), with an overall mean of 0.27. There was no correspondence

between the autosomal QTLs and F_{ST} outliers (Figure S9), and we found no significant differences in the distribution of average dn/ds ratios between transcripts on the Z chromosome and those on the autosomes (Figure S10).

3.3 | Differential expression analysis

The dysgenic sterility phenotype is first evident in early stage, pre-vitellogenic oocytes (Figure 3). We focused on this region of ovaries in quantifying RNA expression differences among backcross individuals. We dissected the previtellogenic (approx. stage 3 and

| cM	limits (cM)	limits (physical)	μ_1	$\beta_1 q_1$	$\beta_2 q_2$	$\beta_3 q_1 q_2$	R^2
			1.37 ± 0.09***	0.88 ± 0.19***			0.2
29.21	15–36.19	5201659–7976401	1.32 ± 0.09***	-0.56 ± 0.18**	0.81 ± 0.18***	-0.59 ± 0.37	0.3
18.75	3.49–33.86	2210592–7862353	1.39 ± 0.09***	0.44 ± 0.18*	0.87 ± 0.18***	1 ± 0.36**	0.3
18.75	14.07–30	4795563–7180089	1.34 ± 0.09***	0.29 ± 0.18	0.7 ± 0.18***	1.47 ± 0.36***	0.34
55.08	53.91–56.24	8861371–13311117	1.4 ± 0.07***	0.6 ± 0.15***	-0.03 ± 0.15	-2.53 ± 0.3***	0.54
11.86	6–12	3012264–6466723	1.61 ± 0.08***	0.24 ± 0.16	-0.22 ± 0.16	2.68 ± 0.32***	0.79
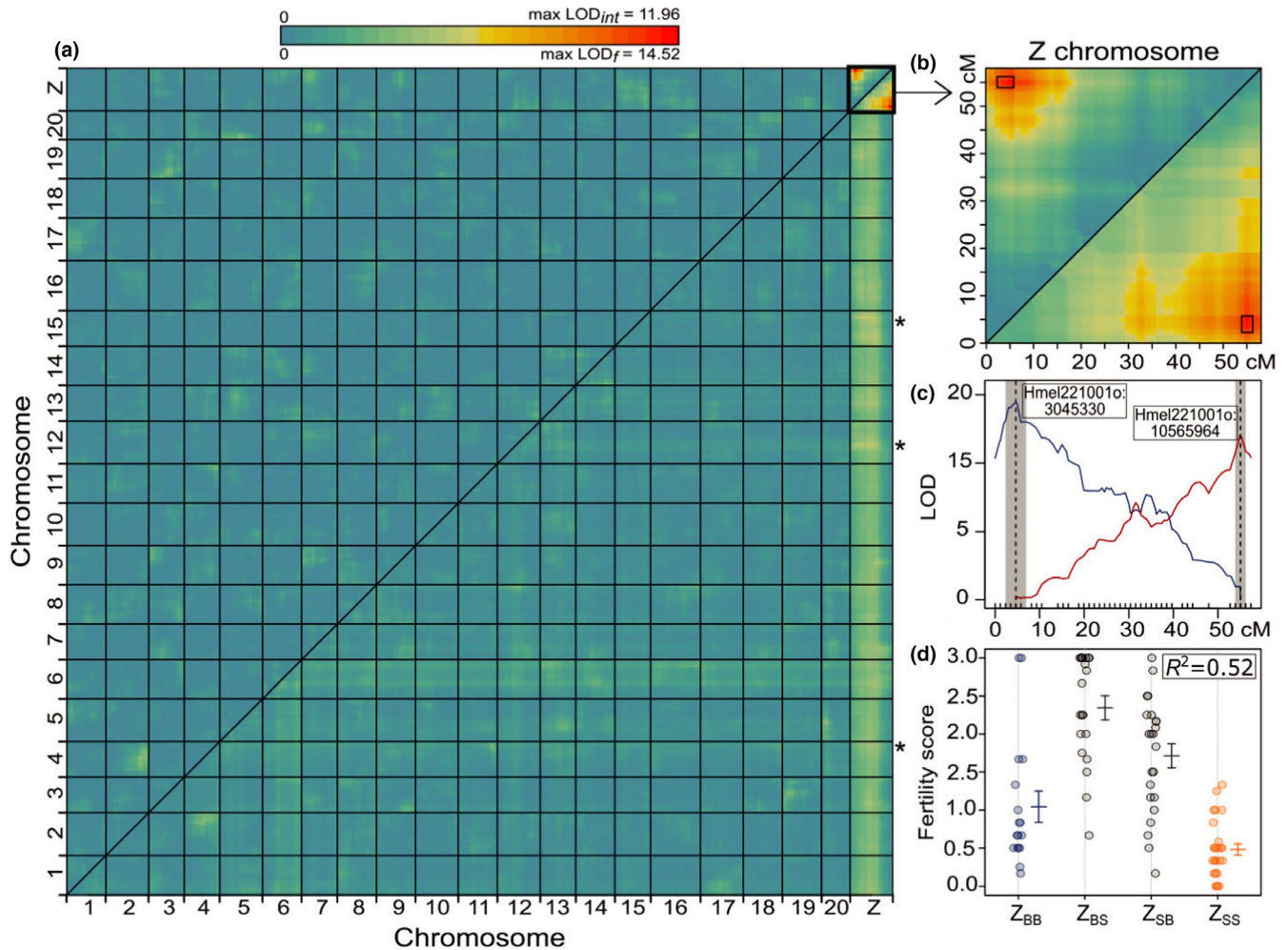
FIGURE 5 Multiple QTL analysis. (a) Heat map for LOD_f values (the full model; lower right triangle) and LOD_{int} scores (the interaction component; upper left triangle) between pairwise combinations of markers across the genome, using H-K regression and reads aligned to Hmel2.5. Blues indicate low scores, reds indicate high scores (maximum observed $LOD_{int} = 11.96$, maximum observed $LOD_f = 14.52$). Statistically significant LOD_f values between the Z chromosome and the autosomes are marked with an asterisk, for p-values see Tables 1 and S32. (b) Enlargement of the Z chromosome, with the Bayesian credible intervals of the significant interaction shown as black boxes. (c) Profile LOD curves for the epistatic QTL on Z chromosome, with the blue line for the proximal QTL and the red line for the distal QTL. The vertical dotted lines give the positions of the QTL peaks, and the grey shaded errors indicate the Bayesian credible intervals. The physical positions of the markers at the QTL peaks are shown in the text boxes. (d) Fertility scores for 87 backcross individuals grouped by their haplotypes at the two interacting markers on the Z chromosome (Hmel221001o:3045330 and Hmel221001:10565964). These four haplotypes explain 52% of the variance in fertility score. Unrecombined pairs of markers inherited from *H. p. butleri* (Z_{BB}) or *H. p. sergestus* (Z_{SS}) are coloured blue and orange, respectively. Errors bars are standard errors

earlier) follicles from six backcross ovaries, two of which were assigned a fertility score of 0–1, two of 1–2 and two of 2–3. We microdissected previtellogenic follicles to investigate the specific phenotype of developmental failure in early-stage oocytes and further classified the phenotypes with a binary scheme: 0 for absence of vitellogenic follicles (fertility score 0–0.99), 1 for the presence of any vitellogenic follicles (fertility score 1–3). In each case, tissue was dissected from all four ovarioles of a single ovary, and we acquired approximately 49 million reads per individual. After filtering our data for sequencing and mapping quality, we quantified expression of 16,774 protein coding genes (Figures 8a and S11).

We then carried out a principal components analysis of these expression data. PC1 explained over 50% of variance and

separated the three fertility score categories in order (Figure 8b). We performed a Wald test to evaluate the effect of change in expression of each transcript to the fertility phenotype in the backcrosses (Chen et al., 2011). After correcting for multiple comparisons, a total of 14% or 2,315 transcripts showed significant effects of expression on binary phenotype ($q < 0.05$). Of these, 941 displayed a positive association with development, meaning that the transcript was expressed at a higher level in more highly developed ovaries. The remaining 1,386 differentially expressed transcripts displayed a negative association with development. We identified 1,771 transcripts in the *H. melpomene* transcriptome that gave strong BLAST hits to genes expressed in *Pararge aegeria* eggs and ovaries. 306 (17%) of these genes were also differentially
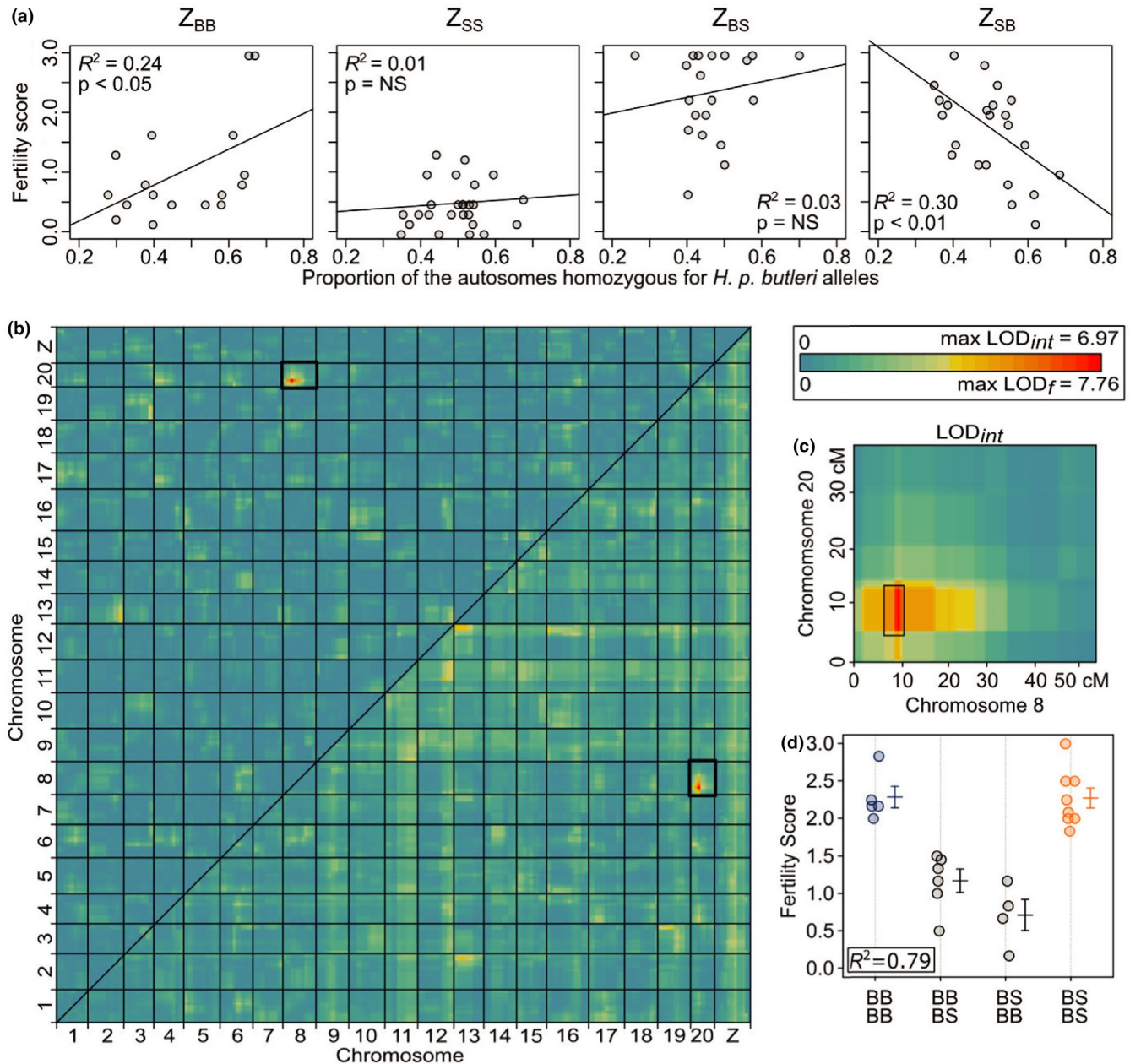
FIGURE 6 Analysis of Z-linked epistatic markers. (a) For each Z chromosome haplotype (Z_{BB}, Z_{SS}, Z_{BS}, Z_{SB}), the proportion of the autosome that is homozygous for *H. p. butleri* alleles was plotted against fertility score. (b) Heat map for two dimensional QTL scan using only Z_{SB} individuals. LOD_f values are shown in the lower right triangle and LOD_{int} values in the upper left triangle. The highlighted box shows the significant associations identified between chromosomes 8 and 20. (c) Enlargement of LOD_{int} between chromosome 8 and chromosome 20, with the Bayesian credible intervals of the QTLs shown as black boxes. (d) Fertility scores for the four autosomal genotypes of Z_{SB} individuals, with the genotype at chromosome 8 (Hmel208001o:1005579) written above, and the genotype at chromosome 20 (Hmel220003o:5817143) written below. These genotypes explain 79% of the variance in fertility score. Errors bars are standard errors

expressed in backcrosses with different developmental phenotypes. None of the chromosomes were enriched for differentially expressed genes; however, the Z chromosome had significantly fewer highly expressed genes than expected, and also significantly fewer *P. aegeria* ovarian genes (Figure S12).

We then searched within the Bayesian credible intervals of the QTLs for differentially expressed transcripts with orthologs implicated in oogenesis in either *D. melanogaster* or *P. aegeria* (Figure 8). Applying this approach to the two interacting QTLs

on the Z chromosome, we identified one gene (*magu*) in the first QTL 4.65 cM, and eight in the second QTL at 55 cM (*Egfr*, *fax*, *Gs2*, *Nedd8*, *parvin*, *Prm*, *sls*, *Syx7*). Within the central additive QTL on the Z chromosome at 29.2 cM, we found two genes (*trol* and *csw*). In the highly divergent region within this QTL (6.5–6.75 Mb, Figure 7c), there are 14 genes, one of which has an orthologue (*ncd*) required for spindle assembly in oocytes in *Drosophila* (Endow & Komma, 1997). However, only three were significantly differentially expressed among fertile and infertile hybrids, and none of
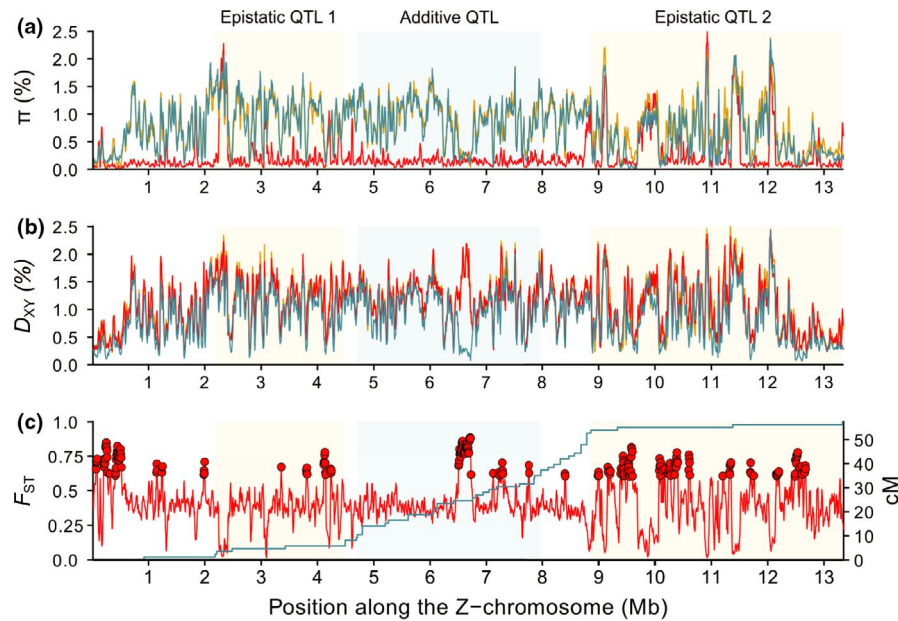
FIGURE 7 Population genetic summary statistics and recombination rate along the Z chromosome. (a) Nucleotide diversity (π) within *H. p. sergestus* (red), *H. p. butleri* (yellow), and *H. elevatus* (blue). (b) Mean pairwise absolute genetic distance (D_{xy}) between *H. p. butleri* and *H. p. sergestus* (red), *H. p. butleri* and *H. elevatus* (blue), and *H. p. sergestus* and *H. elevatus* (yellow). (c) Genetic differentiation (F_{ST}; red line) between *H. p. butleri* and *H. p. sergestus*, with genome-wide F_{ST} outliers as points, based on Z-scores >3. The blue line shows genetic distance (cM) plotted against physical distance (Mb). Shaded areas correspond to the Bayesian credible intervals for the two epistatic QTL at 4.65 and 55 cM, and the single additive QTL at 29.21 cM. D_{xy} and F_{ST} were calculated in sliding windows of 25 kb (with 5 kb increments)

those had orthologs known to be involved in oogenesis. Within the QTL at 11.86 cM on chromosome 20, we identified 11 genes (*baz*, *CG12104*, *CG1572*, *CrebB*, *Ect4*, *Eip75B*, *ine*, *mys*, *Pitslre*, *Ran*, *TpnC73F*). In the QTL at 9.3 cM on chromosome 8, there were only 3 differentially expressed transcripts, only one of which had an orthologue known to be involved in oogenesis (*Art1*).

4 | DISCUSSION

Here, we show that crossing *H. p. butleri* and *H. p. sergestus* in both directions results in F1 hybrid females that are sterile due to disrupted oocyte development, and QTL analysis of backcrosses to *H. p. butleri* shows that sterility is sex-linked. We identify a strong epistatic interaction between loci at opposite ends of the Z chromosome, and a broader, additive QTL towards the centre. In addition, we identify an epistatic interaction involving the Z chromosome and chromosomes 8 and 20, as well as possible associations linking the Z chromosome with chromosomes 4, 12 and 15. By intersecting these QTLs with the results of differential expression analysis, we identify candidate genes with orthologues known to be involved in oogenesis (Figure S13).

4.1 | Genetics of hybrid incompatibility in *Heliconius pardalinus*

To our knowledge, this is the first study of Haldane's rule in Lepidoptera using modern genomic techniques to demonstrate

a complex, epistatic basis of hybrid sterility, as predicted in the Dobzhansky–Muller model. Hybrids between *H. p. butleri* and *H. p. sergestus* are also consistent with the 'two rules of speciation' (Coyne & Orr, 1989b). The first of these is Haldane's rule—the tendency for greater sterility/inviability in the heterogametic sex than in the homogametic sex. There is general consensus that Haldane's rule can be explained in part by dominance theory, which proposes that interactions between recessive X- or Z-linked alleles from one species and a hybrid autosomal genetic background cause incompatibilities in the heterogametic sex (Coyne & Orr, 2004), and our data are broadly consistent with this. Faster evolution of Z-linked genes may also play a role (Charlesworth, Campos & Jackson, 2018; Charlesworth, Coyne & Barton, 1987; Sackton et al., 2014); however, we did not detect any difference between the ratios of nonsynonymous (*dn*) to synonymous (*ds*) substitutions on the autosomes and Z chromosome (Figure S10).

The second rule of speciation is the 'large X effect' on hybrid incompatibility (in Lepidoptera, this is a large effect of the Z chromosome). In hybrids between *Drosophila mauritiana* and *D. sechellia*, the X chromosome has about four times more hybrid male sterility factors than a comparably sized autosomal region (Masly & Presgraves, 2007), and X-linked loci are involved in female, as well as male, hybrid sterility in the *D. virilis* group (Orr & Coyne, 1989). There is, in addition, a large X effect in taxa with undifferentiated sex chromosomes (Dufresnes et al., 2016; Hu & Filatov, 2016), and a large Z effect in birds (Ellegren, 2009). In Lepidoptera, sex-linked hybrid sterility has been shown in *Colias* and *Heliconius* (Grula & Taylor, 1980; Jiggins et al., 2001; Naisbit et al., 2002), and in general, the Z chromosome appears to be
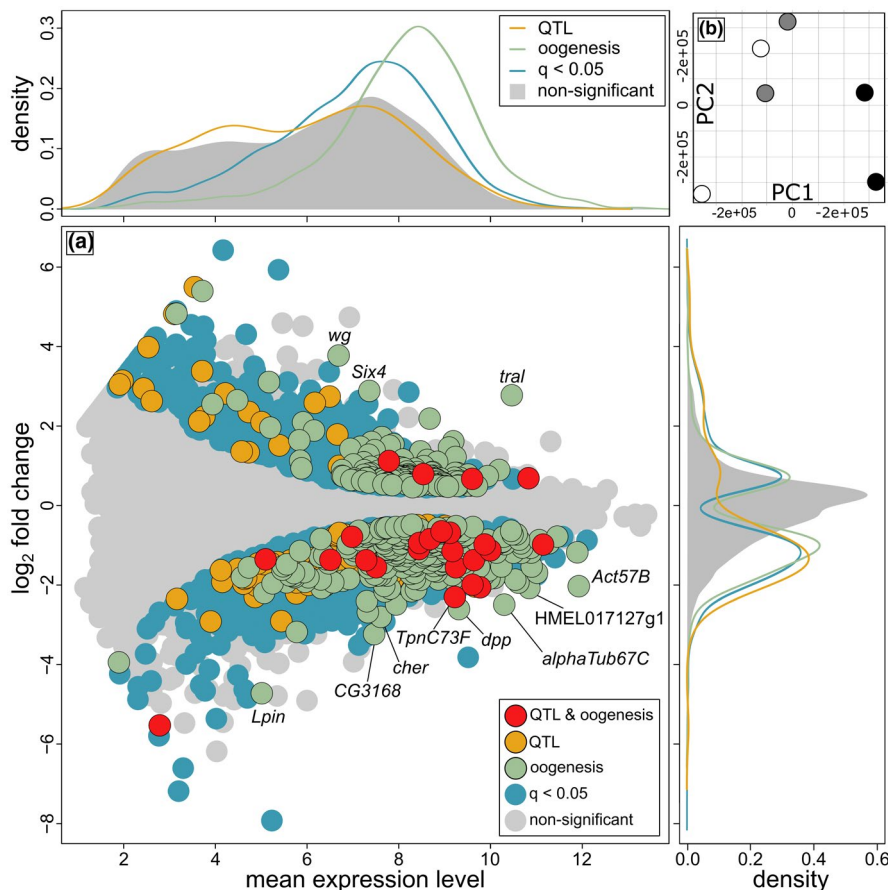
FIGURE 8 Backcross differential expression. (a) Mean expression of transcripts plotted against fold change in expression between fertile (fertility score ≥1) and sterile (fertility score <1) individuals. Positive values of fold change imply higher expression in fertile ovaries. Significantly differentially expressed transcripts are shown in blue of which those with orthologues implicated in oogenesis in either *D. melanogaster* or *P. aegeria* are in green, those within the Bayesian credible intervals of QTLs on chromosomes 8, 20 and Z are shown in orange, and those fitting all these criteria are in red. QTL and/or oogenesis outliers (defined as those falling in the top 1% of transcripts ranked using mean expression × fold change) are labelled. Nonsignificant transcripts are in grey. The shaded grey density plots above and to the right shows the distributions of mean expression levels and fold change values, respectively, with the coloured lines indicating those (i) implicated in oogenesis, (ii) found within QTLs and (iii) significantly differentially expressed in fertile/sterile individuals. Interestingly, most QTL transcripts are overexpressed in the sterile ovaries. (b) Principal component analysis of the six backcross previtellogenic follicles reveals that PC1 segregates individuals based on morphology score. Black: Score = 1, Binary = 0; Grey: Score = 2, Binary = 1; white: Score = 3, Binary = 1

a hotspot for genetic differences between species (Prowell Pashley, 1998; Sperling, 1994). Here, we document three sex-linked QTLs, suggesting a large effect of the Z chromosome on hybrid sterility in *H. pardalinus* (but see Coyne and Orr (1989b) and Hollocher and Wu (1996) for discussion of how observational bias or experimental design can also produce artefactual large sex chromosome effects).

Despite this strong QTL signal, transcripts on the Z chromosome are not more likely to be differentially expressed than those on the autosomes. However, the Z chromosome is significantly depleted in transcripts highly expressed in the ovary and those orthologous to *P. aegeria* genes expressed in the ovary (Figure S12). This pattern has also been seen in the chicken Z chromosome, and similarly, *Drosophila* X chromosomes have a depleted complement of testis-expressed genes, though mouse and human X chromosomes show an increased concentration of genes showing male-specific expression (Ellegren, 2011). Genes involved in oogenesis and those

significantly differentially expressed between ovaries of varying development were skewed towards being overexpressed in ovaries of females with low fertility scores. This pattern was even more extreme among all genes in QTL intervals, regardless of their function (Figure 8a). This could mean that the high expression is due to a general phenomenon such as increased chromatin availability, or derepression of transcriptional regulators.

In *Heliconius melpomene*, crosses between Guiana and Central American populations show hybrid female sterility in only one direction of cross (Jiggins et al., 2001). This kind of asymmetry in hybrid sterility is expected when Dobzhansky–Muller incompatibilities are relatively few, due to recent divergence (Muller, 1942; Turelli & Moyle, 2007). In *H. pardalinus*, crosses in both directions between *H. p. sergestus* and *H. p. butleri* produce sterile hybrid females, suggesting a more complex, multilocus cause of hybrid sterility. Moreover, if hybrid female sterility arises due to epistatic interactions between

the Z chromosome and autosomes, there must be autosomal loci at which *H. p. butleri* alleles are dominant, and others at which *H. p. sergestus* alleles are dominant.

The observation that individuals with unrecombined Z chromosomes (Z_{BB} and Z_{SS}) have low average fertility supports this (Figure 5d). A Z chromosome inherited from *H. p. sergestus* (Z_{SS}) will have deleterious interactions with any autosomal loci where *H. p. butleri* alleles are dominant, and so in a backcross to *H. p. butleri*, individuals carrying such a chromosome should never have full fitness. Similarly, individuals with a Z chromosome inherited from *H. p. butleri* (Z_{BB}) should also have reduced fertility, because of deleterious interactions with autosomal loci with a dominant *H. p. sergestus* allele. However, in a backcross *H. p. butleri*, some fraction of offspring bearing unrecombined *H. p. butleri* Z chromosomes should be fully fertile; those that happen to be homozygous for *H. p. butleri* alleles at all *H. p. sergestus* autosomal dominant loci that interact with the Z. Indeed, two individuals do that these are clearly visible as outliers in Figure 5d, and as predicted, they have the highest proportion of their autosomes homozygous for *H. p. butleri* alleles (*B/B*) (Figure 6a).

It is less easy to explain why individuals holding a recombined Z_{BS} or Z_{SB} chromosome have high average fertility. Male hybrid sterility between Bogotá and US populations of *Drosophila pseudoobscura* is the product of complex epistasis between seven genes, which includes interactions between sex-linked markers (Orr & Irving, 2001; Phadnis, 2011). Subsequent work on *D. pseudoobscura* and *D. persimilis* has shown that epistasis can even modify the dominance of loci causing hybrid male sterility (Chang & Noor, 2010). Given this potential for complexity, a complete explanation of the epistatic interactions in our crosses requires further work. Nonetheless, we note that if *H. p. butleri* and *H. p. sergestus* have differentially fixed derived alleles at opposing ends of the Z chromosome, one of these recombinants could represent the ancestral haplotype. For example, the high fitness of individuals bearing a Z_{BS} chromosome could potentially be explained if it were ancestral, and thus compatible with many derived alleles at autosomal loci.

In contrast, Z_{SB} individuals are notable for high variance in fertility (Figure 5d). They show a negative correlation between fertility and the proportion of their autosomes that is homozygous for *butleri* alleles (Figure 6a), and the variance in their fertility can be explained largely by an interaction between chromosome 8 and chromosome 20 (Table 1, Figure 6b). Females that are either homozygous or heterozygous at both loci are fully fertile, but individuals homozygous at one locus and heterozygous at the other are less fertile (Figure 6d). As such, it is unclear whether this pair of loci have any effect on the sterility of F1 females, even though they clearly have some effect in the backcross we studied here.

4.2 | Candidate genes and comparison with *Drosophila* hybrid incompatibility loci

Oocyte development fails in sterile hybrid females in *H. pardalinus* at Lepidoptera stages 3–4 (Figure 3) (homologous with stages 8–9 of oogenesis in *D. melanogaster*), a period characterized by border follicle cell migration (Yamauchi & Yoshitake, 1984). Within the Bayesian credible intervals of the QTLs on the Z chromosome and chromosomes 8 and 20, we identified 23 transcripts differentially expressed between sterile and fertile females and with orthologs known to be involved in oogenesis in either *D. melanogaster* or the Speckled Wood butterfly (*P. aegeria*). Three of these are known to be associated with border follicle cells, all of which had levels of expression in pure *H. butleri* consistent with fertile backcrosses (Figure S13).

Within the proximal epistatic Z-linked QTL at 4.65 cM, we identified only one gene known to be involved in oogenesis, *magu*, mutants of which are known to cause defective border cell migration in *D. melanogaster* (Raza et al., 2019). Within the distal Z-linked QTL at 55.08 cM, 8 oogenesis genes were identified. One of these, *Epidermal growth factor receptor* (*Egfr*), guides dorsal migration of border cells during *Drosophila* oogenesis stage 9 (Duchek & Rørth, 2001), and is also expressed in the ovarian transcriptome of *P. aegeria* (Carter et al., 2013). We found 11 genes involved in oogenesis within the QTL on chromosome 20. One of these, the multi-PDZ domain protein *bazooka* (*baz*), regulates border cell migration (Pinheiro & Montell, 2004), is expressed in the *P. aegeria* ovarian transcriptome and furthermore is notable for being highly overexpressed in sterile individuals ($\log_2$ fold change = −5.52 for transcript HMEL016161g1.t3, the sixth lowest value in the data set). Within the QTL on chromosome 8, we found no genes known to be involved in border follicle cells. However, HMEL037834g1.t2, with ortholog *Neurexin 1* (*Nrx-1*), stood out as having the third highest $\log_2$ fold change (β = 5.51) in the data set (Figure S11). While not known to be involved in oogenesis, *Neurexin 1* is known to influence expression of *gurken* (*grk*) (Geng & Macdonald, 2007). The asymmetrical localization of *gurken* mRNA is key for its function during oogenesis, to establish anterior–posterior and dorso-ventral axes in the egg and embryo, and *gurken* encodes a TGFα family signalling ligand that activates the intracellular MAP kinase pathway via the product of *Egfr*.

Differentially expressed transcripts located within QTL intervals, such as those discussed above, represent candidate regions for *cis*-acting differences between the two subspecies. Investigation of differential expression on its own, we can also identify putative *trans*-acting effects, or downstream consequences of the QTLs identified here. *Trailer-hitch* (*tral* has strong differential expression, high overall expression in ovaries, and is known to be involved in *Drosophila* oogenesis at stages 8–9 (Figures 8, S14. Figure S13) (Snee & Macdonald, 2009; Wilhelm, Buszczak & Sayles, 2005). Like *Nrx-1*, *tral* is involved in specifying the localization of the dorso-ventral patterning gene *grk*. We also noticed that alternative splices of transcript HMEL015815g1, orthologous to gene *spire* (*spir*) stood out as outliers in Figure S11. Although mapping to chromosome 1 and not in a QTL, HMEL015815g1.t2 was significantly underexpressed in sterile individuals ($\log_2$ fold change = 4.85), and HMEL015815g1.t6 significantly overexpressed ($\log_2$ fold change = −3.94). *spir* is also involved in stages 8–9 of oogenesis in *D. melanogaster*, where it affects the dorsal–ventral and anterior–posterior axes of the egg

(Dahlgaard, Raposo, Niccoli & St Johnston, 2007; Wellington et al., 1999).

The above discussion is based on the hypothesis that differences in expression *per se* are responsible for the sterile phenotype we observe in hybrids. However, although we finely dissected the ovaries to collect comparable regions, it is possible that the populations of cells were different between the sterile and fertile ovaries. This could be the case if cell death in sterile individuals occurred during an intermediate stage of previtellogenic growth. In that case, the differential expression may reflect further development in fertile ovaries rather than misexpression in sterile ones.

Drosophila have long been used as a model to study developmental genetics, including the genetic basis of hybrid sterility. Some classical Dobzhansky–Muller incompatibilities have been identified and characterized in the genus (Bayes & Malik, 2009; Brideau et al., 2006; Tang & Presgraves, 2009). Because *Drosophila* has XY sex determination, in hybrids it is normally males that show sterility (Haldane, 1922). However, hybrid female dysgenesis has been observed in *D. melanogaster* in so-called P-M hybrids in which oogenesis arrests at a very early stage (Bingham, Kidwell & Rubin, 1982; Kidwell, Kidwell & Sved, 1977; Schaefer, Kidwell & Fausto-Sterling, 1979). This phenotype is due to a loss of control of P element transposition, normally suppressed via the *Drosophila* piRNA pathway in P strain flies (Evgen'ev et al., 1997; Kelleher, Edelman & Barbash, 2012). Superficially, the *Heliconius* sterility phenotype described in this study parallels this *Drosophila* case. The hypothesis that transposon silencing through the piRNA pathway is mis-regulated in sterile female hybrids has been explicitly tested in a different *Heliconius* hybrid system, *H. melpomene* and *H. cydno*. A subset of transposable elements were indeed derepressed in F1 hybrids, but there was no evidence that piRNAs themselves or three proteins involved in the piRNA pathway were misexpressed (Pinharanda, 2017). In our case, low fertility *H. pardalinus* female hybrids expressed three proteins in the piRNA pathway (*piwi/aubergine*, *AGO2/3* and *vasa*) at lower levels than in more fertile individuals, although only *vasa* expression differences were significant (Figure S14). In addition, one of our candidate genes, *tral*, forms a complex with piRNA proteins that inhibits P element transposition of a variety of transposons Liu, Qi, Wang and Lin (2011). A *Drosophila*-like transposon derepression mechanism is therefore plausible, but the evidence remains inconclusive at present.

4.3 | Evolution of hybrid incompatibilities

Heliconius p. sergestus is endemic to the dry forests of upper Huallaga valley in the Andes and is separated from *H. p. butleri* in the Amazonian lowlands by the intervening Cordillera Escalera (Figure 1). Nonetheless, the two subspecies are known to come into contact occasionally, and putative wild hybrids exist (Michel Cast pers. comm.; Brown, 1976; Rosser et al., 2019). Theory predicts that in the face of gene flow, DMIs are more likely to be maintained when they are linked to traits involved in divergent

adaptations (Bank, Bürger & Hermisson, 2012), and *Heliconius* provide a possible example of this (Merrill, Van Schooten, Scott & Jiggins, 2011). Divergent selection to different habitats could thus have facilitated the evolution of sterility within *H. pardalinus*, similar to how hybrid inviability has evolved between plant populations as a pleiotropic consequence of adaptations to heavy metals (Macnair & Christie, 1983).

However, an alternative hypothesis is that hybrid sterility arose during an initial split between *H. elevatus* and *H. pardalinus*, only to be lost by hybridization between sympatric populations in the Amazon, but retained in the allopatric subspecies *H. p. sergestus*. Phylogenetic analysis is consistent with this: phylogenies made within Fst islands of divergence support the traditional taxonomic groupings (Kryvokhyzha, 2014), but concatenated whole genome phylogenies render *H. pardalinus* paraphyletic, with *H. p. butleri* more closely related to the widespread Amazonian species *H. elevatus* than to *H. p. sergestus* (Rosser et al., 2019). Moreover, despite strong assortative mating, *H. p. butleri* and *H. elevatus* are known to be fully fertile, while crosses between *H. p. sergestus* and *H. elevatus* are sterile, with phenotypes similar to those found here between *H. p. sergestus* and *H. p. butleri* (Rosser et al., 2019). Intriguingly, in the central 250 kb region of high F_{ST} between *H. p. sergestus* and *H. p. butleri* (Figure 7c), we observed a reduction in D_{xy} between the Amazon taxon *H. p. butleri* and *H. elevatus* (Figure 7b). The notable drop in diversity (π) in this same region in both *H. p. butleri* and *H. elevatus* (Figure 7a) suggests a strong, recent selective sweep that also introgressed between these sympatric populations. Given that this region is in the middle of the main Z chromosome QTL for sterility between *H. p. butleri* and *H. p. sergestus*, introgression of this region is a candidate for explaining the lack of hybrid sterility between *H. elevatus* and *H. p. butleri* in the Amazon.

5 | CONCLUSIONS

The genetics of Haldane's rule and Dobzhansky–Muller incompatibilities have been extensively studied in *Drosophila* and a few other male heterogametic systems, but hitherto there has been little genomic work on female heterogametic systems. Our work with *Heliconius* butterflies represents the first such study in Lepidoptera. We employ thousands of SNPs across the genome to map multiple regions involved in hybrid female sterility and show an especially large effect of the Z chromosome. By intersecting these results with the list of differentially expressed genes among fertile and sterile hybrids, we identify three candidate genes (*magu*, *Egfr*, *baz*) potentially involved in hybrid sterility due to dysfunctional cell migration. Many questions remain unanswered, and functional genetic studies will be required to understand the mechanisms of ovariole development failure in hybrids. Nonetheless, we were able to show that several of the major findings from studies of Haldane's rule in *Drosophila* male sterility (e.g. multilocus effects, epistasis, involvement of the sex chromosome) are replicated in female sterile hybrids in a female heterogametic system. Future work can now address the genetic

basis of sterility, as well as the potential tie-in with selfish genetic elements and with genes that act to defend the genome against their replication.

AUTHOR CONTRIBUTIONS

NR, NE, KKD and JM designed the study. NR, NE, LMQ and MN collected the data. NR, NE and FS analysed the data. All authors contributed to writing the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

RADseq and RNAseq reads (.fastq files), and sterility phenotypes and pedigree (.xlsx) are freely available at https://zenodo.org, https://doi.org/10.5281/zenodo.5628350.

ORCID

Neil Rosser https://orcid.org/0000-0001-7796-2548
Nathaniel B. Edelman https://orcid.org/0000-0002-5142-6760
Lucie M. Queste https://orcid.org/0000-0002-7402-8079

REFERENCES

Bank, C., Bürger, R., & Hermisson, J. (2012). The limits to parapatric speciation: Dobzhansky-Muller incompatibilities in a continent–island model. *Genetics*, 191, 845–863. https://doi.org/10.1534/genetics.111.137513

Bayes, J. J., & Malik, H. S. (2009). Altered heterochromatin binding by a hybrid sterility protein in *Drosophila* sibling species. *Science*, 326, 1538–1541.

Bingham, P. M., Kidwell, M. G., & Rubin, G. M. (1982). The molecular basis of PM hybrid dysgenesis: The role of the P element, a P-strain-specific transposon family. *Cell*, 29, 995–1004. https://doi.org/10.1016/0092-8674(82)90463-9

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-638 optimal probabilistic RNA seq quantification. *Nature Biotechnology*, 34, 525–527.

Brideau, N. J., Flores, H. A., Wang, J., Maheshwari, S., Wang, X., & Barbash, D. A. (2006). Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. *Science*, 314, 1292–1295.

Broman, K. W., Wu, H., Sen, Ś., & Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, 19, 889–890. https://doi.org/10.1093/bioinformatics/btg112

Brown, K. S. (1976). An illustrated key to the silvaniform *Heliconius* (lepidoptera: Nymphalidae) with descriptions of new subspecies. *Transactions of the American Entomological Society*, 102, 373–484.

Büning, J. (1994). *The insect ovary: Ultrastructure, previtellogenic growth and evolution*. Springer Science & Business Media.

Burri, R. (2017). Interpreting differentiation landscapes in the light of long-term linked selection. *Evolution Letters*, 1, 118–131. https://doi.org/10.1002/evl3.14

Butlin, R., Debelle, A., Kerth, C., Snook, R. R., Beukeboom, L. W., Castillo, R. C., Diao, W., Maan, M. E., Paolucci, S., Weissing, F. J., van de Zande, L., Hoikkala, A., Geuverink, E., Jennings, J., Kankare, M., Knott, K. E., Tyukmaeva, V. I., Zoumadakis, C., Ritchie, M. G., ... Macias Garcia, C. (2012). What do we need to know about speciation? *Trends in Ecology & Evolution*, 27, 27–39.

Carter, J. M., Baker, S. C., Pink, R., Carter, D. R., Collins, A., Tomlin, J., Gibbs, M., & Breuker, C. J. (2013). Unscrambling butterfly oogenesis. *BMC Genomics*, 14, 283. https://doi.org/10.1186/1471-2164-14-283

Castillo, D. M., & Barbash, D. A. (2017). Moving speciation genetics forward: Modern techniques build on foundational studies in *Drosophila*. *Genetics*, 207, 825–842.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22, 3124–3140. https://doi.org/10.1111/mec.12354

Chang, A. S., & Noor, M. A. (2010). Epistasis modifies the dominance 661 of loci causing hybrid male sterility in the drosophila pseudoobscura species group. *Evolution: International Journal of Organic Evolution*, 64, 253–260.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience*, 4, s13742–s14015. https://doi.org/10.1186/s13742-015-0047-8

Charif, D., & Lobry, J. (2007). SeqinR 1.0-2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In U. Bastolla, M. Porto, H. Roman, & M. Vendruscolo (Eds.), *Structural approaches to sequence evolution: Molecules, networks, populations. Biological and Medical Physics, Biomedical Engineering* (pp. 207–232). Springer Verlag. ISBN : 978-3-540-35305-8.

Charlesworth, B., Campos, J. L., & Jackson, B. C. (2018). Faster-x evolution: Theory and evidence from drosophila. *Molecular Ecology*, 27, 3753–3771.

Charlesworth, B., Coyne, J. A., & Barton, N. H. (1987). The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*, 130, 113–146. https://doi.org/10.1086/284701

Chen, Z., Liu, J., Ng, H. K. T., Nadarajah, S., Kaufman, H. L., Yang, J. Y., & Deng, Y. (2011). Statistical methods on detecting differentially expressed genes for RNA-seq data. *BMC Systems Biology*, 5, S1. https://doi.org/10.1186/1752-0509-5-S3-S1

Corbett-Detig, R. B., Zhou, J., Clark, A. G., Hartl, D. L., & Ayroles, J. F. (2013). Genetic in compatibilities are widespread within species. *Nature*, 504, 135–137. https://doi.org/10.1038/nature12678

Coughlan, J. M., & Matute, D. R. (2020). The importance of intrinsic postzygotic barriers throughout the speciation process. *Philosophical Transactions of the Royal Society B*, 375, 20190533. https://doi.org/10.1098/rstb.2019.0533

Coyne, J. A. (1992). Genetics and speciation. *Nature*, 355, 511–515. https://doi.org/10.1038/355511a0

Coyne, J. A., & Orr, H. A. (1989a). Patterns of speciation in *Drosophila*. 685. *Evolution*, 43, 362–381. https://doi.org/10.1111/j.1558-5646.1989.tb04233.x

Coyne, J. A., & Orr, H. A. (1989b). Two rules of speciation. In D. Otte, & J. Endler (Eds.), *Speciation and its consequences* (pp. 180–207). Sinauer Associates Inc.

Coyne, J. A., & Orr, H. A. (1997). "Patterns of speciation in *Drosophila*" revisited. *Evolution*, 51, 295–303.

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sinauer Associates.

Dahlgaard, K., Raposo, A. A., Niccoli, T., & St Johnston, D. (2007). Capu and Spire assemble a cytoplasmic actin mesh that maintains microtubule organization in the *Drosophila* oocyte. *Developmental Cell*, 13, 539–553.

Darwin, C. (1859). *On the origin of species by means of natural selection.* John Murray.

Davey, J. W., Barker, S. L., Rastas, P. M., Pinharanda, A., Martin, S. H., Durbin, R., McMillan, W. O., Merrill, R. M., & Jiggins, C. D. (2017). No evidence for maintenance of a sympatric *Heliconius* species barrier by chromosomal inversions. *Evolution Letters*, 1, 138–154.

Delph, L. F., & Demuth, J. P. (2016). Haldane's rule: genetic bases and their empirical support. *Journal of Heredity*, 107, 383–391. https://doi.org/10.1093/jhered/esw026

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation dna sequencing data. *Nature Genetics*, 43, 491. https://doi.org/10.1038/ng.806

Dobzhansky, T. (1936). Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics*, 21, 113. https://doi.org/10.1093/genetics/21.2.113

Doncaster, L., & Raynor, G. H. (1906). Breeding experiments with Lepidoptera. *Proceedings of the Zoological Society of London*, 1, 125–133.

Duchek, P., & Rørth, P. (2001). Guidance of cell migration by EGF receptor signaling during *Drosophila* oogenesis. *Science*, 291, 131–133.

Dufresnes, C., Majtyka, T., Baird, S. J. E., Gerchen, J. F., Borzée, A., Savary, R., Ogielska, M., Perrin, N., & Stöck, M. (2016). Empirical evidence for large x-effects in animals with undifferentiated sex chromosomes. *Scientific Reports*, 6, 1–7. https://doi.org/10.1038/srep21029

Dunlap-Pianka, H., Boggs, C. L., & Gilbert, L. E. (1977). Ovarian dynamics in heliconiine butterflies: Programmed senescence versus eternal youth. *Science*, 197, 487–490. https://doi.org/10.1126/science.197.4302.487

Ellegren, H. (2009). Genomic evidence for a large-z effect. *Proceedings of the Royal Society B: Biological Sciences*, 276, 361–366. https://doi.org/10.1098/rspb.2008.1135

Ellegren, H. (2011). Sex-chromosome evolution: Recent progress and the influence of male and female heterogamety. *Nature Reviews Genetics*, 12, 157–166. https://doi.org/10.1038/nrg2948

Emms, D. M., & Kelly, S. (2019). Orthofinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20, 1–14. https://doi.org/10.1186/s13059-019-1832-y

Endow, S. A., & Komma, D. J. (1997). Spindle dynamics during meiosis in drosophila oocytes. *The Journal of Cell Biology*, 137, 1321–1336. https://doi.org/10.1083/jcb.137.6.1321

Etter, P. D., Preston, J. L., Bassham, S., Cresko, W. A., & Johnson, E. A. (2011). Local *de novo* assembly of RAD paired-end contigs using short sequencing reads. *PLoS One*, 6, e18561. https://doi.org/10.1371/journal.pone.0018561

Evgen'ev, M. B., Zelentsova, H., Shostak, N., Kozitsina, M., Barskyi, V., Lankenau, D.-H., & Corces, V. G. (1997). *Penelope*, a new family of transposable elements and its possible role in hybrid dysgenesis in *Drosophila virilis*. *Proceedings of the National Academy of Sciences*, 94, 196–201. https://doi.org/10.1073/pnas.94.1.196

Geng, C., & Macdonald, P. (2007). Identification of genes that influence *gurken* expression. *Fly*, 1, 259–267.

Good, J. M., Dean, M. D., & Nachman, M. W. (2008). A complex genetic basis to X-linked hybrid male sterility between two species of house mice. *Genetics*, 179, 2213–2228. https://doi.org/10.1534/genetics.107.085340

Grula, J. W., & Taylor, O. R. Jr (1980). Some characteristics of hybrids derived from the sulfur butterflies, *Colias eurytheme* and *C. philodice*: Phenotypic effects of the x chromosome. *Evolution*, 673–687.

Haldane, J. B. S. (1922). Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, 12, 101–109. https://doi.org/10.1007/BF02983075

Haley, C. S., & Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, 69, 315. https://doi.org/10.1038/hdy.1992.131

Hoffman, J. I., Simpson, F., David, P., Rijks, J. M., Kuiken, T., Thorne, M. A., Lacy, R. C., & Dasmahapatra, K. K. (2014). High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences*, 111, 3775–3780. https://doi.org/10.1073/pnas.1318945111

Hollocher, H., & Wu, C. I. (1996). The genetics of reproductive isolation in the drosophila simulans clade: X vs. autosomal effects and male vs. female effects. *Genetics*, 143, 1243–1255.

Hu, X. S., & Filatov, D. A. (2016). The large-x effect in plants: increased species divergence and reduced gene flow on the silene x-chromosome. *Molecular Ecology*, 25, 2609–2619.

Jiggins, C. D. (2017). *The ecology and evolution of heliconius butterflies*. Oxford University Press.

Jiggins, C. D., Linares, M., Naisbit, R. E., Salazar, C., Yang, Z. H., & Mallet, J. (2001). Sex linked hybrid sterility in a butterfly. *Evolution*, 55, 1631–1638. https://doi.org/10.1111/j.0014-3820.2001.tb00682.x

Kalirad, A., & Azevedo, R. B. (2017). Spiraling complexity: A test of the snowball effect in a computational model of RNA folding. *Genetics*, 206, 377–388.

Kelleher, E. S., Edelman, N. B., & Barbash, D. A. (2012). *Drosophila* interspecific hybrids phenocopy piRNA-pathway mutants. *PLoS Biology*, 10, e1001428. https://doi.org/10.1371/journal.pbio.1001428

Kidwell, M. G., Kidwell, J. F., & Sved, J. A. (1977). Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics*, 86, 813–833.

Kryvokhyzha, D. (2014). *Whole genome resequencing of Heliconius butterflies revolutionizes our view of the level of admixture between species*. Uppsala Universitët. Master's thesis.

Li, H. (2011). A statistical framework for snp calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27, 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H. (2013). *Aligning sequence reads, clone sequences and assembly contigs with BWA MEM*. arXiv preprint arXiv:1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, W. H. (1993). Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *Journal of Molecular Evolution*, 36, 96–99. https://doi.org/10.1007/BF02407308

Liu, L., Qi, H., Wang, J., & Lin, H. (2011). PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition. *Development*, 138, 1863–1873. https://doi.org/10.1242/dev.059287

Macnair, M., & Christie, P. (1983). Reproductive isolation as a pleiotropic effect of copper tolerance in mimulus guttatus? *Heredity*, 50, 295–302.

Maheshwari, S., & Barbash, D. A. (2011). The genetics of hybrid incompatibilities. *Annual Review of Genetics*, 45, 331–355. https://doi.org/10.1146/annurev-genet-110410-132514

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

Masly, J. P., & Presgraves, D. C. (2007). High-resolution genome-wide dissection of the two rules of speciation in *Drosophila*. *PLoS Biology*, 5, e243. https://doi.org/10.1371/journal.pbio.0050243

Matute, D. R., Butler, I. A., Turissini, D. A., & Coyne, J. A. (2010). A test of the snowball theory for the rate of evolution of hybrid incompatibilities. *Science*, *329*, 1518–1521. https://doi.org/10.1126/science.1193440

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*, 1297–1303. https://doi.org/10.1101/gr.107524.110

Merrill, R. M., Van Schooten, B., Scott, J. A., & Jiggins, C. D. (2011). Pervasive genetic associations between traits causing reproductive isolation in *Heliconius* butterflies. *Proceedings of the Royal Society B: Biological Sciences*, *278*, 511–518.

Mihola, O., Trachtulec, Z., Vlcek, C., Schimenti, J. C., & Forejt, J. (2009). A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science*, *323*, 373–375. https://doi.org/10.1126/science.1163601

Morgan, T. H. (1910). Sex-limited inheritance in *Drosophila*. *Science N.S*, *32*, 125–133.

Morgan, T. H. (1911). Random segregation versus coupling in Mendelian inheritance. *Science N.S*, *34*, 384.

Muller, H. (1942). Isolating mechanisms, evolution, and temperature. In *Biological symposia, temperature and evolution. Isolating mechanisms* (Vol. 6, pp. 71–125). Genetic Control of Embryonic Development.

Nadeau, N. J. (2016). Genes controlling mimetic colour pattern variation in butterflies. *Current Opinion in Insect Science*, *17*, 24–31. https://doi.org/10.1016/j.cois.2016.05.013

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

Nosil, P., & Schluter, D. (2011). The genes underlying the process of speciation. *Trends in Ecology & Evolution*, *26*, 160–167. https://doi.org/10.1016/j.tree.2011.01.001

Orr, H. A. (1995). The population genetics of speciation: The evolution of hybrid incompatibilities. *Genetics*, *139*, 1805–1813. https://doi.org/10.1093/genetics/139.4.1805

Orr, H. A. (1997). Haldane's rule. *Annual Review of Ecology and Systematics*, *28*, 195–218. https://doi.org/10.1146/annurev.ecolsys.28.1.195

Orr, H. A., & Coyne, J. A. (1989). The genetics of postzygotic isolation in the *Drosophila virilis* group. *Genetics*, *121*, 527–537. https://doi.org/10.1093/genetics/121.3.527

Orr, H. A., & Irving, S. (2001). Complex epistasis and the genetic basis of hybrid sterility in the drosophila pseudoobscura bogota-usa hybridization. *Genetics*, *158*, 1089–1100.

Orr, H. A., & Turelli, M. (2001). The evolution of postzygotic isolation: Accumulating Dobzhansky-Muller incompatibilities. *Evolution*, *55*, 1085–1094. https://doi.org/10.1111/j.0014-3820.2001.tb00628.x

Phadnis, N. (2011). Genetic architecture of male sterility and segregation distortion in *Drosophila pseudoobscura* Bogotá–USA hybrids. *Genetics*, *189*, 1001–1009.

Pimentel, H., Bray, N. L., Puente, S., Melsted, P., & Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods*, *14*, 687. https://doi.org/10.1038/nmeth.4324

Pinharanda, A. L. P. (2017). *The genomic basis of species barriers in Heliconius butterflies*. University of Cambridge. Ph.D. thesis.

Pinharanda, A., Rousselle, M., Martin, S. H., Hanly, J. J., Davey, J. W., Kumar, S., Galtier, N., & Jiggins, C. D. (2019). Sexually dimorphic gene expression and transcriptome evolution provide mixed evidence for a fast-Z effect in *Heliconius*. *Journal of Evolutionary Biology*, *32*, 194–204.

Pinheiro, E. M., & Montell, D. J. (2004). Requirement for *Par-6* and *Bazooka* in *Drosophila* border cell migration. *Development*, *131*, 5243–5251.

Presgraves, D. C. (2002). Patterns of postzygotic isolation in Lepidoptera. *Evolution*, *56*, 1168–1183. https://doi.org/10.1111/j.0014-3820.2002.tb01430.x

Presgraves, D. C. (2007). Speciation genetics: epistasis, conflict and the origin of species. *Current Biology*, *17*, R125–R127. https://doi.org/10.1016/j.cub.2006.12.030

Presgraves, D. C. (2010). Darwin and the origin of interspecific genetic incompatibilities. *The American Naturalist*, *176*, S45–S60. https://doi.org/10.1086/657058

Presgraves, D. C., & Orr, H. A. (1998). Haldane's rule in taxa lacking a hemizygous X. *Science*, *282*, 952–954. https://doi.org/10.1126/science.282.5390.952

Prowell Pashley, D. (1998). Sex linkage and speciation in Lepidoptera. In D. J. Howard, & S. H. Berlocher (Eds.), *Endless Forms. Species and Speciation* (pp. 309–319). Oxford University Press.

Ramsey, J., Bradshaw, H., Jr., & Schemske, D. W. (2003). Components of reproductive Isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution*, *57*, 1520–1534. https://doi.org/10.1111/j.0014-3820.2003.tb00360.x

Rastas, P. (2017). Lep-MAP3: Robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics*, *33*, 3726–3732. https://doi.org/10.1093/bioinformatics/btx494

Raza, Q., Choi, J. Y., Li, Y., O'Dowd, R. M., Watkins, S. C., Chikina, M., Hong, Y., Clark, N. L., & Kwiatkowski, A. V. (2019). Evolutionary rate covariation analysis of *E-cadherin* identifies *Raskol* as a regulator of cell adhesion and actin dynamics in *Drosophila*. *PLoS Genetics*, *15*, e1007720. https://doi.org/10.1371/journal.pgen.1007720

Rosser, N., Phillimore, A. B., Huertas, B., Willmott, K. R., & Mallet, J. (2012). Testing historical explanations for gradients in species richness in heliconiine butterflies of tropical America. *Biological Journal of the Linnean Society*, *105*, 479–497. https://doi.org/10.1111/j.1095-8312.2011.01814.x

Rosser, N., Queste, L. M., Cama, B., Edelman, N. B., Mann, F., Mori Pezo, R., Morris, J., Segami, C., Velado, P., Schulz, S., Mallet, J. L. B., & Dasmahapatra, K. K. (2019). Geographic contrasts between pre- and postzygotic barriers are consistent with reinforcement in *Heliconius* butterflies. *Evolution*, *73*, 1821–1838.

Sackton, T. B., Corbett-Detig, R. B., Nagaraju, J., Vaishna, L., Arunkumar, K. P., & Hartl, D. L. (2014). Positive selection drives faster-Z evolution in silkmoths. *Evolution*, *68*, 2331–2342. https://doi.org/10.1111/evo.12449

Salazar, C. A., Jiggins, C. D., Arias, C., Tobler, A., Bermingham, E., & Linares, M. (2005). Hybrid incompatibility is consistent with a hybrid origin of *Heliconius heurippa* Hewitson from its close relatives, *Heliconius cydno* Doubleday and *Heliconius melpomene* Linnaeus. *Journal of Evolutionary Biology*, *18*, 247–256. https://doi.org/10.1111/j.1420-9101.2004.00839.x

Sánchez, A. P., Pardo-Diaz, C., Enciso-Romero, J., Muñoz, A., Jiggins, C. D., Salazar, C., & Linares, M. (2015). An introgressed wing pattern acts as a mating cue. *Evolution*, *69*, 1619–1629. https://doi.org/10.1111/evo.12679

Schaefer, R. E., Kidwell, M. G., & Fausto-Sterling, A. (1979). Hybrid dysgenesis in *Drosophila melanogaster*: morphological and cytological studies of ovarian dysgenesis. *Genetics*, *92*, 1141–1152.

Schartl, M. (2008). Evolution of *Xmrk*: an oncogene, but also a 871 speciation gene? *BioEssays*, *30*, 822–832.

Schilthuizen, M., Giesbers, M., & Beukeboom, L. (2011). Haldane's rule in the 21st century. *Heredity*, *107*, 95–102. https://doi.org/10.1038/hdy.2010.170

Schumer, M., Cui, R., Powell, D. L., Dresner, R., Rosenthal, G. G., & Andolfatto, P. (2014). High-resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species. *eLife*, *3*, e02535. https://doi.org/10.7554/eLife.02535

Seixas, F. A., Edelman, N. B., & Mallet, J. (2021). Synteny-based genome assembly for 16 species of *Heliconius* butterflies, and an assessment

of structural variation across the genus. *Genome Biology and Evolution*, *13*(7), evab069. https://doi.org/10.1093/gbe/evab069

Snee, M. J., & Macdonald, P. M. (2009). *Bicaudal C* and *trailer hitch* have similar roles in *gurken* mRNA localization and cytoskeletal organization. *Developmental Biology*, *328*, 434–444. https://doi.org/10.1016/j.ydbio.2009.02.003

Sperling, F. A. (1994). Sex-linked genes and species differences in lepidoptera. *The Canadian Entomologist*, *126*(3), 807–818. https://doi.org/10.4039/Ent126807-3

Sweigart, A. L., Fishman, L., & Willis, J. H. (2006). A simple genetic incompatibility causes hybrid male sterility in *Mimulus*. *Genetics*, *172*, 2465–2479. https://doi.org/10.1534/genetics.105.053686

Tang, S., & Presgraves, D. C. (2009). Evolution of the *Drosophila* nuclear pore complex results in multiple hybrid incompatibilities. *Science*, *323*, 779–782.

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., & Prins, P. (2015). Sambamba: Fast processing of NGS alignment formats. *Bioinformatics*, *31*, 2032–2034. https://doi.org/10.1093/bioinformatics/btv098

Turelli, M., & Moyle, L. C. (2007). Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics*, *176*, 1059–1088. https://doi.org/10.1534/genetics.106.065979

Turelli, M., & Orr, H. A. (1995). The dominance theory of Haldane's Rule. *Genetics*, *140*, 389–402. https://doi.org/10.1093/genetics/140.1.389

Wellington, A., Emmons, S., James, B., Calley, J., Grover, M., Tolias, P., & Manseau, L. (1999). Spire contains actin binding domains and is related to ascidian posterior end mark-5. *Development*, *126*, 5267–5274. https://doi.org/10.1242/dev.126.23.5267

Wilhelm, J. E., Buszczak, M., & Sayles, S. (2005). Efficient protein trafficking requires *trailer hitch*, a component of a ribonucleoprotein complex localized to the ER in *Drosophila*. *Developmental Cell*, *9*, 675–685. https://doi.org/10.1016/j.devcel.2005.09.015

Wu, C. I., & Davis, A. W. (1993). Evolution of postmating reproductive isolation: The composite nature of Haldane's Rule and its genetic bases. *The American Naturalist*, *142*, 187–212. https://doi.org/10.1086/285534

Yamauchi, H., & Yoshitake, N. (1984). Developmental stages of ovarian follicles of the silk worm, *Bombyx mori* L. *Journal of Morphology*, *179*, 21–31. https://doi.org/10.1002/jmor.1051790104

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.