Title

Genetic incompatibilities are widespread within species

Journal

Nature, 504(7478)

ISSN

0028-0836

Authors

Corbett-Detig, Russell B
Zhou, Jun
Clark, Andrew G
et al.

Publication Date

2013-12-05

DOI

10.1038/nature12678

Peer reviewed

Published in final edited form as:

Nature. 2013 December 5; 504(7478): 135–137. doi:10.1038/nature12678.

Genetic Incompatibilities are Widespread Within Species

Russell B. Corbett-Detig[1], **Jun Zhou**[1], **Andrew G. Clark**[2,3], **Daniel L. Hartl**[1], and **Julien F. Ayroles**[1,2,4,†]

[1]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA

[2]Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14853, USA

[3]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA

[4]Harvard Society of Fellows, Harvard University, Cambridge, MA 02138, USA

Abstract

The importance of epistasis — non-additive interactions between alleles — in shaping population fitness has long been a controversial topic, hampered in part by lack of empirical evidence[1,2,3,4]. Traditionally, epistasis is inferred based on non-independence of genotypic values between loci for a given trait. However epistasis for fitness should also have a genomic footprint[5,6,7]. To capture this signal, we have developed a simple approach that relies on detecting genotype ratio distortion (GRD) as a signal for epistasis, and we confirm experimentally that instances of GRD represent loci with epistatic fitness effects. In applying this method to a large panel of *Drosophila melanogaster* recombinant inbred lines[8, 9], we conservatively estimate that any two haploid genomes in this study are expected to harbor 1.15 pairs of incompatible alleles. This observation has important implications for speciation genetics, as it indicates that the raw material to drive reproductive isolation is segregating contemporaneously within species and does not necessarily require, as proposed by the Dobzhansky–Muller model, the emergence of incompatible mutations independently derived and fixed in allopatry. The relevance of our result extends beyond speciation, as it demonstrates that epistasis is widespread but that it may often go undetected due to lack of statistical power or lack of genome-wide scope of the experiments.

The role of epistasis in shaping genetic variation and contributing to observable differences within and between populations has been the focus of much debate[1,2,3]. In complex trait genetics, the additive paradigm used in genome-wide association (GWA) studies[10] has recently been challenged by mounting evidence highlighting the importance of non-additive

†Corresponding Author: ayroles@fas.harvard.edu.

Authors contribution: R.B.C.D., J.F.A., conducted bioinformatics and statistical analyses, R.B.C.D., J.F.A., J.Z., performed experiments, J.Z. carried out molecular work; A.G.C., D.L.H., gave analytical and conceptual advice throughout the project; J.F.A conceived of the project.

Authors declare no competing interests.

All code used and generated for this study is available upon request.

interactions between alleles[4]. While the debate has been centered on the relative contribution of epistasis to the genetic variance, we still have a poor grasp of the extent to which epistatis affects the mean genotypic values of traits, an important step towards understanding the genetic basis of complex trait and the organization of molecular pathways[5]. Although epistasis is widely accepted to underlie the genetic basis of speciation, many details of this phenomenon remain poorly understood[2,3,5]. In particular, the evolutionary origins of the alleles that cause reproductive isolation are largely unidentified. Therefore, the importance of epistasis in shaping fitness within and between populations remains an important question in evolutionary biology.

Our understanding of the contribution of epistasis and the molecular details underlying non-additive genetic interactions is limited largely by the scarcity of available data. Although the idea that populations may harbor alleles with epistatic fitness effects has existed in the literature for some time, very few examples have been dissected at the genetic level (except for individual cases[6,11]). Furthermore, as yet, no systematic surveys have been conducted in diploid out-crossing species that are sufficiently powered to detect small fitness effects or to finely map interacting loci.

The traditional approach used to detect epistasis by statistical means relies on the observation of non-additivity of genotypic values between loci for a given phenotype. However, epistasis for fitness should have a genomic signature, regardless of our ability to measure a given phenotype[5,6,7]. In particular, one expects that unfavorable allelic combinations will be under-represented, and this should precipitate a deviation from Mendelian proportions among unlinked incompatible alleles (detected by performing a screen for statistical association between alleles at loci that are not physically linked; Supplemental Methods). Hereafter we refer to such deviations as Genotype-Ratio-Distortion (GRD). In natural populations an exhaustive search for GRD is computationally intractable, statistically underpowered, or both[6]. By contrast, model organisms allow us to create experimental populations, in which the amount of genetic variation and recombination can be controlled, thereby amplifying the signature of epistasis in a background of reduced dimensionality.

Here we apply tests of epistasis to the *Drosophila* Synthetic Population Resource (DSPR)[8, 9] (Extended Data Figure 1). To create the DSPR, two sets of eight highly inbred strains of diverse geographic origins were independently crossed in a round-robin design. Each set was duplicated and maintained for 50 generations in large freely-mating population cages (generating 4 panels A-1,2 and B-1,2). Subsequently, approximately 400 recombinant inbred lines (RILs) in each of four independent panels were created through 20 generations of sib-mating. After inbreeding, each RIL was genotyped at densely spaced markers, allowing a description of each RIL's genome as a genetic mosaic of the eight founding lines originally crossed (Extended Data Figure 1). The 50 generations of recombination and the large number of RILs within a panel provides replication over random allelic permutations. This replication is essential to attain statistical power for the detection of small effect epistasis.

We first excluded the possibility that residual population structure within the DSPR created association among alleles in the absence of epistasis by performing principle component

analysis (Extended Data Figure 2, Supplemental Methods). Subsequently, we identified 22 pairs of epistatically interacting alleles in the DSPR (Figure 1, Extended Data Table 1, Extended Data Figure 3). Importantly, of the 44 incompatible alleles, 27 appear to be shared between two or more strains (Extended Data Table 1). This indicates that incompatible alleles are segregating at polymorphic frequencies in natural populations, and are not a result of inbreeding or long-term maintenance at small population size. Based on the frequencies in the founder strains, we estimate that any pairwise combination of founders has, on average, 1.15 pairs of epistatically interacting alleles. This is probably an underestimate, both because our statistical approach is conservative and because selectively disfavored allelic combinations may be purged by selection during the free-recombination phase of the DSPR.

We next sought to confirm the predicted effect on reproductive fitness and to identify the underlying phenotype of two pairs of incompatible haplotypes (Figure 2-a,c). Using the original founder strains that contributed the putatively interacting alleles, we performed experimental crosses, and in both cases, we discovered that the negative interaction is caused by the minor alleles at each locus (Figure 2b,d; Extended Data Figure 1). Specifically, in the case of one incompatibility between chromosomes 2 and 3, males that are homozygous for both incompatible alleles produce on average 74% fewer offspring compared to all other allelic combinations (P = 5.51121E-09 LRT, Figure 2b, Extended Data Figure 4). No significant effect was detected in females for any combination of genotypes. Using the same approach we validated a second instance of GRD, selected in the low range of effect size, between a haplotype on chromosome X and 3 (Extended Data Figure 5). We again observe a significant decrease (22%) in F_2 male fertility (P = 8.25e-5 LRT, Figure 2b, Extended Data Figure 4), suggesting that GRD is a reliable signature of epistasis. The ‘faster-males’ theory [2,12] and subsequent experimental confirmations [reviewed in [13]] predicts that male infertility will evolve more rapidly than other forms of post-zygotic reproductive isolation. Although we only have phenotypic data for our confirmed examples, the fact that both implicate male fertility as the underlying phenotype suggests that this effect may extend to within-species fitness epistasis.

The DSPR was intercrossed for sufficiently many generations (50+) that little linkage disequilibrium remains; hence this approach allows us to narrow down likely candidate genes associated with epistatic interaction for male fecundity. In total, there are three genes within the haplotype on chromosome arm 2R (~40 kb). *notopleural* (*np*) is at the peak of this region, a gene expressed in mature sperm[14] with alleles that are known to affect viability and sterility [15]. Notably, the human orthologue of *np* is associated with sperm-dysfunction in humans[16]. The interacting haplotype on chromosome arm 3R contains only two genes. In the center of this region is *Cyp12e1*, a P450-cytochrome associated with electron transport in the mitochondria[17]. Interestingly, *Cyp12e1* harbors a non-synonymous mutation in a highly conserved protein domain. Mitochondrial dysfunction is commonly associated with male sterility in humans, plants, and *D. melanogaster*[18], and therefore seems a plausible candidate phenotype.

To confirm that these observations were not specific to the Drosophila DSPR, we used the same method to screen for GRD in two additional RIL panels: The MAGIC panel in

Arabidopsis[19] and the NAM panel in maize[20]. We found 7 instances of GRDs in Arabidopsis and 5 in maize (Table S2). Although we have not validated these results, they suggest that GRD is present in other species as well.

Although the contribution of epistasis to variation in fitness is controversial in some fields[21], the Dobzhansky-Muller incompatibility (DMI) model[22, 2] is a widely accepted guiding principle for biologists studying of the genetic basis of intrinsic, postzygotic reproductive isolation. Largely motivated by this model, which predicts that alleles causing hybrid incompatibility are derived and fixed after population divergence, much empirical work in speciation genetics has been dedicated to mapping DMIs between species that diverged relatively long ago on an evolutionarily time scale[2, 1] (Extended Data Figure 5). However, it is unclear if these known examples of so-called 'speciation genes'[1, 2, 23] are an accurate representation of the earliest events in speciation, which have the greatest biological significance[2]. Even species that have diverged for only ~250,000 years have evolved complete male sterility an estimated 15 times over[24]. A reasonable interpretation of this evidence may concede that known 'speciation genes' are unlikely to be the same as those that initially contributed to reproductive isolation, but that these examples are instructive as to the properties of those genes[2] –logic that closely mirrors our own.

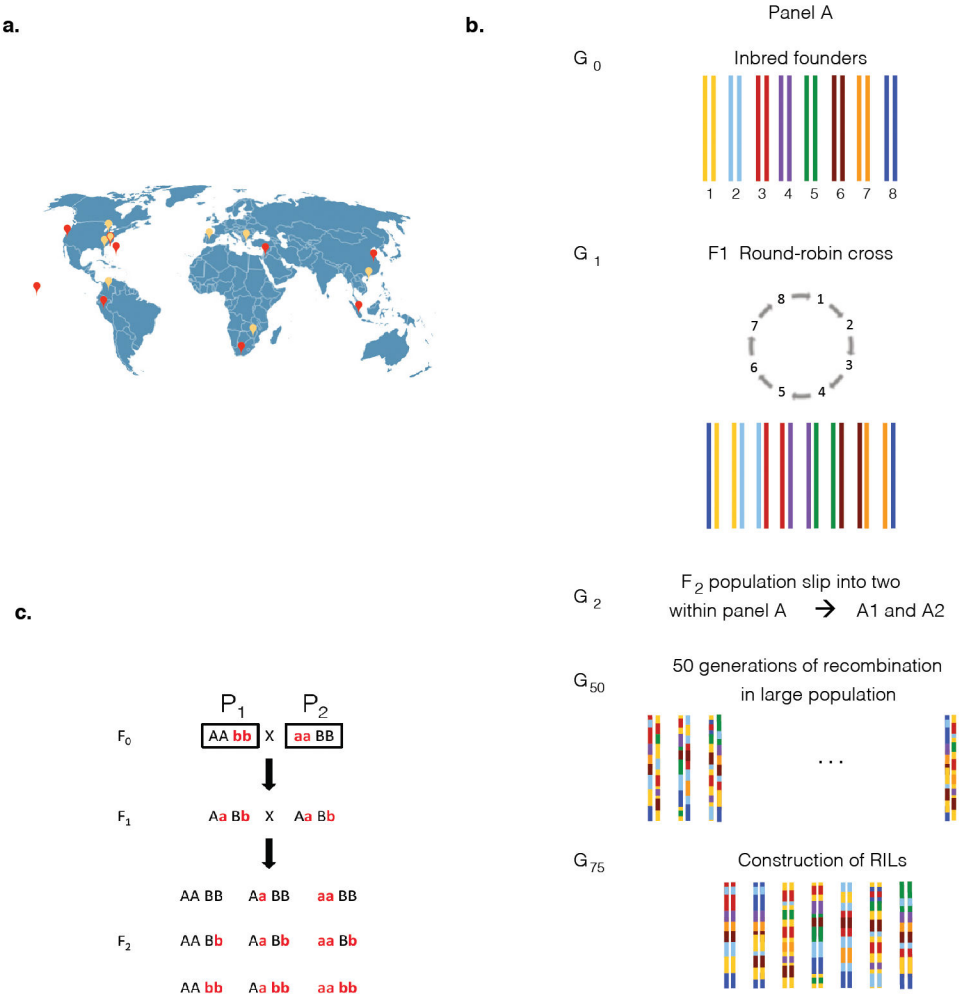
Our central finding, that fitness epistasis is widespread within natural populations indicates that the raw material to drive reproductive isolation is segregating contemporaneous within species and does not necessarily require, as proposed by the DMI model[22], the emergence of genetically incompatible mutations independently derived and fixed in allopatric lineages[23]. We therefore need to explore the possibility that reproductive isolation could be achieved through divergence in frequencies of numerous preexisting, polymorphic, small-effect incompatibilities[26, 27, 28]. The implications of these results go beyond understanding the role of intra-specific incompatibility in the context of speciation. Our work shows that epistasis for fitness related traits has a detectable genomic footprint and supports the idea that latent incompatibilities often exists between segregating variation within populations, only to be released when divergent lineages hybridize. This discovery highlights the importance of understanding the contribution of epistasis to observable phenotypic differences within and between populations.

Methods Summary

We genotyped the RILs of the DSPR by requiring that each putative variant be supported by a minimum of five reads. All sites wherein two or more alleles are supported by five reads were discarded. We confirmed that the RIL panels were free of cryptic population structure by performing principal component analysis (Extended Data Figure 2). We next excluded sites wherein fewer than 150 individuals have a supported genotype, where the minor allele was present in fewer than 10 individuals, or where more than 15% percent of individuals with data had heterozygous genotypes. Following this, we assessed statistical significance for non-independence between pairwise combinations of alleles using a χ^2 test, and applied a 5% false discovery rate to correct for multiple testing. To reduce type 1 error, we restricted our search to inter-chromosomal comparisons and required that each putative instance of GRD be consistent with signal from adjacent variants (see Supplemental Methods).

To confirm the predictions of the GRD scan, we first crossed the two DSPR founder strains that contributed the predicted interacting alleles. We then intercrossed the F_1 progeny to produce F_2 offspring. Virgin F_2 females were then individually and randomly mated to a single F_2 male. After mating for 4 days, the F_2 pairs were individually genotyped at known variable sites near the interacting alleles. We recorded the number of progeny of each pair to assay productivity. We used TaqMan kits to perform qPCR on the F_2 parents, and performed numerous statistical analyses[5,29] to quantify epistatic effects as a product of genotypes at the two sites (see Supplemental Methods).

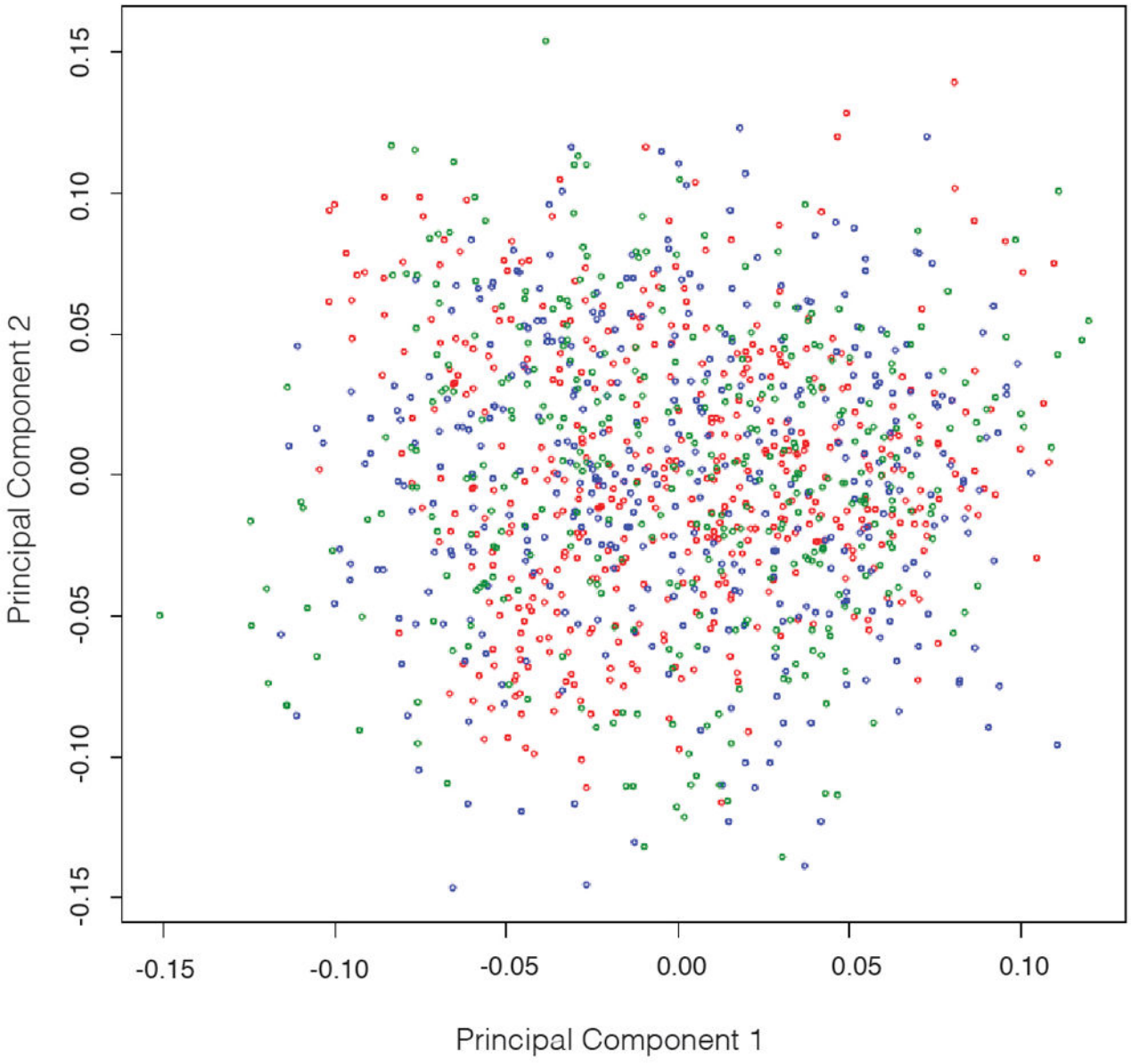
Extended Data



Extended data Figure 1. Description of the DSPR and validation scheme

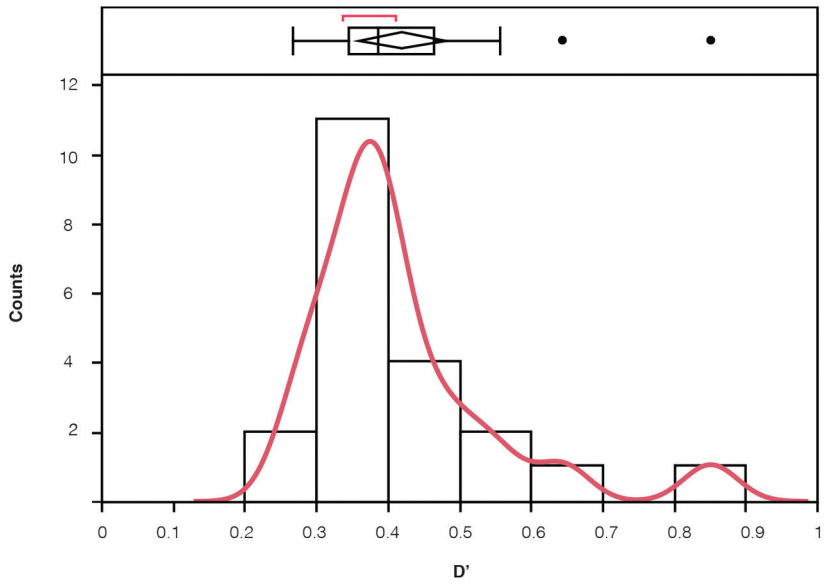
a. Geographic distribution of the DSPR founding strains (in orange panel A and in red panel B). **b.** Construction of the recombinant inbred lines. For each panel each founder strains were crossed in a round-robin design (Line 1 ♀ × Line 2 ♂, Line 2 ♀ × Line 3 ♂,…, Line 8 ♀ × Line 1 ♂) to produce F_1s, the F_1 were then allowed to mate free to produce an F_2 population. In each panel A and B, these F_2 population were split into two independent

population to create panels A1, A2 and B1, B2. Each was allowed to recombine freely for 50 generations, in very large population. After 50 generation, for each replicate panel, about 400 isofemale lines were inbred for 25 generations to create the 4 panels of RIL used in this study. **c.** Crossing scheme used to validate epistatic effects. A pair of founder segregating incompatible allele was selected and crossed to produce F_1's, we then intercrossed the F1 progeny to produce a large F2 population, segregating all possible allelic combinations between alleles at loci 1 and 2. We then counted the progeny each pair produced by intercrossing a large number of F_2's which were later genotyped at sites near to the predicted interacting loci.

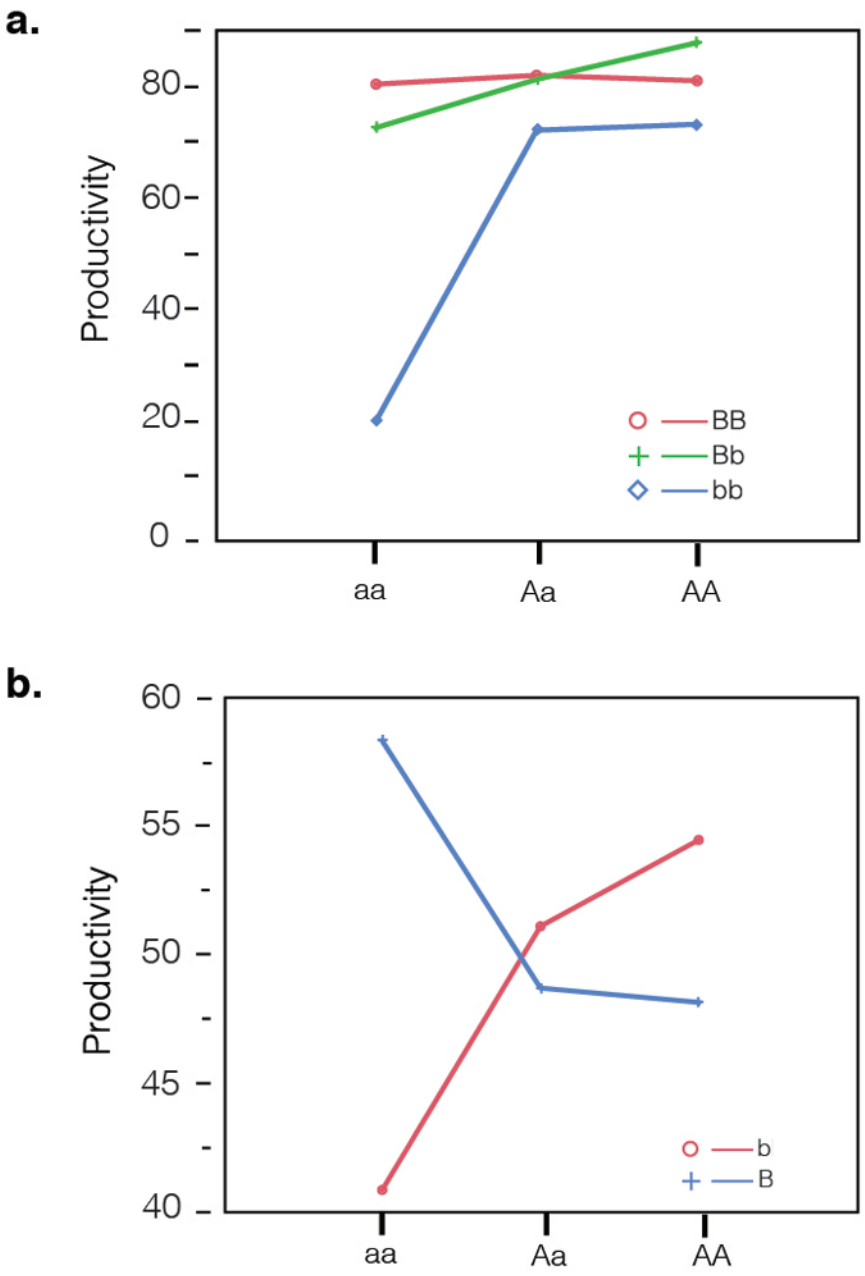


Extended data Figure 2. Principal component analysis of each three DSPR RILs panel

In Green panel A-2, blue panel B-1 and red panel B-2. Showing no evidence of population structure.

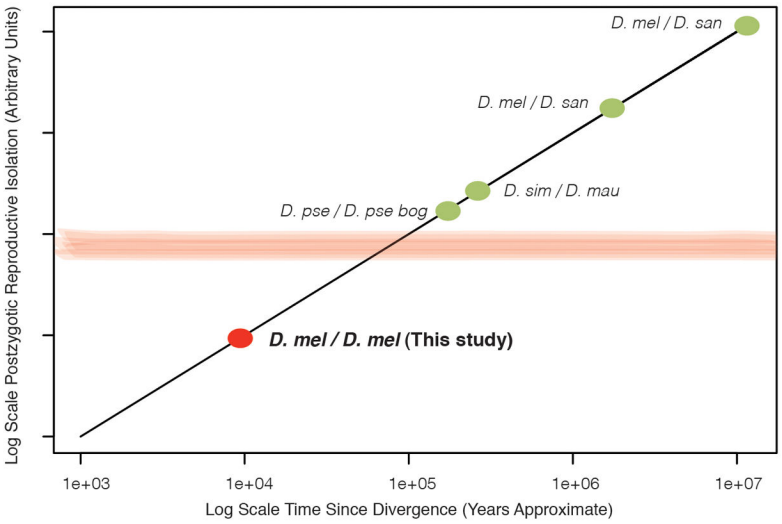


Extended data Figure 3. D′ distribution for significant GRD
(plotted across DSPR panels).



Extended data Figure 4. Epistasis plot for each validated instance of GRD

a. GRD between chromosomes 2R and 3R (tagged by SNPs 2R:4806926, on the X axis and 3R:5870973, colored lines) shows strong negative epistasis due to the low fitness of the aa;bb genotype. The additive-by-additive genetic effect is equal to −13.75 (sensu Phillips et al[5] and Cheverud[29]). **b**. GRD between chromosomes 3L and X (tagged by SNPs 3L: 11510853, on the X axis and X: 16483812, colored lines) also shows negative epistasis. Here the additive-by-additive genetic effect equals −5.94.



Extended data Figure 5. The accumulation of post-zygotic reproductive isolation through time (log-scaled X-axis)

Approximate divergence times of commonly studied *Drosophila* species are indicated by blue stars, and the red star indicates a reasonable expectation for divergence times of stocks used to found the DSPR (~10,000 years). The red-line indicates a very approximate "speciation threshold", and indicates that many species pairs that are commonly studied exceed this threshold significantly.

Extended data Table 1

List of all significant inter-chromosomal GRD identified in the DSPR

Panel	Chromosome 1	Position 1	Chromosome 2	Position 2	Number of RILs counted	1st Major Allele	1st Minor Allele	2dn Major Allele	2
B-2	2L	2767815	3R	4492436	391	C	A	A	
B-2	2L	8027605	X	13053319	418	C	T	C	
B-2	2L	10869984	3R	10633352	177	C	G	T	
B-2	2L	21657908	3R	5870973	444	A	C	G	
B-2	2R	4806926	3R	5870973	443	G	A	G	
B-2	2R	8464341	X	5753834	457	A	G	T	
B-2	2R	20512785	X	19647595	436	T	G	A	
B-2	3L	9627942	X	13622563	263	G	T	A	
B-2	3R	20437352	X	19127400	385	C	T	A	
B-1	2L	66907	3L	11787066	256	A	G	C	
B-1	2L	22131178	3R	5598460	375	A	G	A	
B-1	2L	2896002	X	11823233	274	A	G	T	
B-1	2L	10065007	3R	11588046	283	C	A	T	
B-1	2L	4140219	X	11514794	361	G	A	T	

Panel	Chromosome 1	Position 1	Chromosome 2	Position 2	Number of RILs counted	1st Major Allele	1st Minor Allele	2dn Major Allele	2
B-1	2R	3232234	3R	5598051	356	G	A	G	
B-1	2R	14543771	3R	22691609	355	A	G	T	
B-1	3R	13807981	X	8763898	151	T	C	C	
B-1	3R	18284739	X	14686047	378	G	T	G	
A-2	2L	19531958	X	12584624	326	G	T	C	
A-2	3L	11510853	X	16483812	354	A	T	G	
A-2	3R	23793328	X	14472525	64	A	T	C	
A-2	2L	16549805	3L	10566820	236	C	T	A	

Panel	Chromosome 1	Position 1	Chromosome 2	Position 2	Number of RILs counted	1st Major Allele	1st Minor Allele	2dn Major Allele	HM 2
B-2	2L	2767815	3R	4492436	391	C	A	A	
B-2	2L	8027605	X	13053319	418	C	T	C	
B-2	2L	10869984	3R	10633352	177	C	G	T	
B-2	2L	21657908	3R	5870973	444	A	C	G	
B-2	2R	4806926	3R	5870973	443	G	A	G	
B-2	2R	8464341	X	5753834	457	A	G	T	
B-2	2R	20512785	X	19647595	436	T	G	A	
B-2	3L	9627942	X	13622563	263	G	T	A	
B-2	3R	20437352	X	19127400	385	C	T	A	
B-1	2L	66907	3L	11787066	256	A	G	C	
B-1	2L	22131178	3R	5598460	375	A	G	A	
B-1	2L	2896002	X	11823233	274	A	G	T	
B-1	2L	10065007	3R	11588046	283	C	A	T	
B-1	2L	4140219	X	11514794	361	G	A	T	
B-1	2R	3232234	3R	5598051	356	G	A	G	
B-1	2R	14543771	3R	22691609	355	A	G	T	
B-1	3R	13807981	X	8763898	151	T	C	C	
B-1	3R	18284739	X	14686047	378	G	T	G	
A-2	2L	19531958	X	12584624	326	G	T	C	
A-2	3L	11510853	X	16483812	354	A	T	G	
A-2	3R	23793328	X	14472525	64	A	T	C	
A-2	2L	16549805	3L	10566820	236	C	T	A	

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to S Macdonald and E King for creating the DSPR and sharing the RIL and founder strains with us. We thank C Jones, B de Bivort, T Sackton, SD Kocher, J Grenier, and NE Soltis and three anonymous reviewers for helpful comments and discussions. We thank Xi Shi for technical assistance. This work was supported by grants: NIH GM065169 and GM084236 to DLH, HD059060 to AGC, Harvard Society of Fellows Fellowship and Harvard Milton Funds to JFA. RBCD is supported by a Harvard Prize Fellowship.

Abbreviations

DMI	Dobzhansky-Muller incompatibility
DSPR	*Drosophila* synthetic population resource
GRD	Genotype-Ratio-Distortion
RADseq	restriction-site-associated DNA sequencing
RIL	recombinant inbred line

References

1. Presgraves DC. The molecular evolutionary basis of species formation. Nature Rev Genet. 2010; 11(3):175–180. [PubMed: 20051985]
2. Coyne, JA.; Orr, HA. Speciation. Sunderland, MA: Sinauer Associates; 2004.
3. Cutter AD. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. TREE. 2011; 27(4):209–218. [PubMed: 22154508]
4. Carlborg O, Haley CS. Epistasis: too often neglected in complex trait studies? Nature Rev Genet. 2004; 5:618–625. [PubMed: 15266344]
5. Phillips PC. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nature Rev Genet. 2008; 9(11):855–867. [PubMed: 18852697]
6. Bomblies K, et al. Autoimmune response as a mechanism for a Dobzhansky-Muller- type incompatibility syndrome in plants. PLoS biol. 2007; 5(9):e236. [PubMed: 17803357]
7. Payseur BA, Hoekstra HE. Signatures of reproductive isolation in patterns of single nucleotide diversity across inbred strains of house mice. Genetics. 2005; 171:1905– 1916. [PubMed: 16143616]
8. King EG, Macdonald SJ, Long AD. Properties and power of the Drosophila Synthetic Population Resource for the routine dissection of complex traits. Genetics. 2012; 191(3):935–949. [PubMed: 22505626]
9. King EG, et al. Genetic dissection of a model complex trait using the Drosophila Synthetic Population Resource. Gen Resch. 2012; 22(8):1558–1566.
10. Zuk, Or, et al. The mystery of missing heritability: genetic interactions create phantom heritability. PNAS. 2012; 109(4):1193–1198. [PubMed: 22223662]
11. Bikard D, et al. Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. Science. 2009; 323(5914):623–626. [PubMed: 19179528]
12. Palopoli MF, Wu CI. Genetics of hybrid male sterility between Drosophila sibling species: a complex web of epistasis is revealed in interspecific studies. Genetics. 1994; 138(2):329–341. [PubMed: 7828817]
13. Wu CI, Johnson NA, Palopoli MF. Haldane's rule its legacy: Why are there so many sterile males? TREE. 1996; 11(7):281–284. [PubMed: 21237844]
14. WasBrough RW, et al. The Drosophila melanogaster sperm proteome-II (DmSP-II). J Proteomics. 2010; 73:2171–2185. [PubMed: 20833280]
15. Dockendorff TC, Robertson SE, Faulkner DL, Jongens TA. Genetic characterization of the 44D-45B region of the Drosophila melanogaster genome based on an F2 lethal screen. Mol Gen Genet. 2000; 263(1):137–143. [PubMed: 10732682]

16. Netzel-Arnett S, et al. The glycosylphosphatidylinositol-anchored serine protease PRSS21 (testisin) imparts murine epididymal sperm cell maturation and fertilizing ability. Biol reprod. 2009; 81(5):921–932. [PubMed: 19571264]
17. Kasai S, Tomita T. Male specific expression of a cytochrome P450 (Cyp312a1) in Drosophila melanogaster. Biochem Biophys Res Commun. 2003; 300(4):894–900. [PubMed: 12559957]
18. Meiklejohn CD, Montooth KL, Rand DM. Positive and negative selection on the mitochondrial genome. TIG. 2007; 23(6):259–263. [PubMed: 17418445]
19. Kover PX, et al. multiparent advanced generation inter-cross to fine-map quantitative traits in Arabidopsis thaliana. PLoS gen. 2009; 5(7):e1000551.
20. McMullen MD, et al. Genetic properties of the maize nested association mapping population. Science. 2009; 325(5941):737–740. [PubMed: 19661427]
21. Hill WG, Goddard ME, Visscher PM. Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. PLoS Genet. 2008; 4(2):e1000008. [PubMed: 18454194]
22. Dobzhansky, T. Genetics and the Origin of Species. Columbia University Press; New York: 1937.
23. Orr HA, Turelli M. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution. 2001; 55(6):1085–1094. [PubMed: 11475044]
24. Presgraves DC, Stephan W. Pervasive adaptive evolution among interactors of the Drosophila hybrid inviability gene, Nup96. MBE. 2007; 24(1):306–314.
25. Tao, Yun, et al. Genetic dissection of hybrid incompatibilities between *Drosophila simulans and D. mauritiana*. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. Genetics. 2003; 164(4):1383–1398. [PubMed: 12930747]
26. Fitzpatrick BM. Hybrid dysfunction: Population genetic and quantitative genetic perspectives. Am Nat. 2008; 171(4):491–498. [PubMed: 20374137]
27. Demuth JP, Wade MJ. On the theoretical and empirical framework for studying genetic interactions within and among species. Am Nat. 2005; 165(5):524–536. [PubMed: 15795850]
28. Reed LK, Markow TA. Early events in speciation: polymorphism for hybrid male sterility in Drosophila. PNAS. 2004; 101(24):9009–9012. [PubMed: 15184657]
29. Cheverud JM, Routman EJ. Epistasis and its contribution to genetic variance components. Genetics. 1995; 139(3):1455–1461. [PubMed: 7768453]

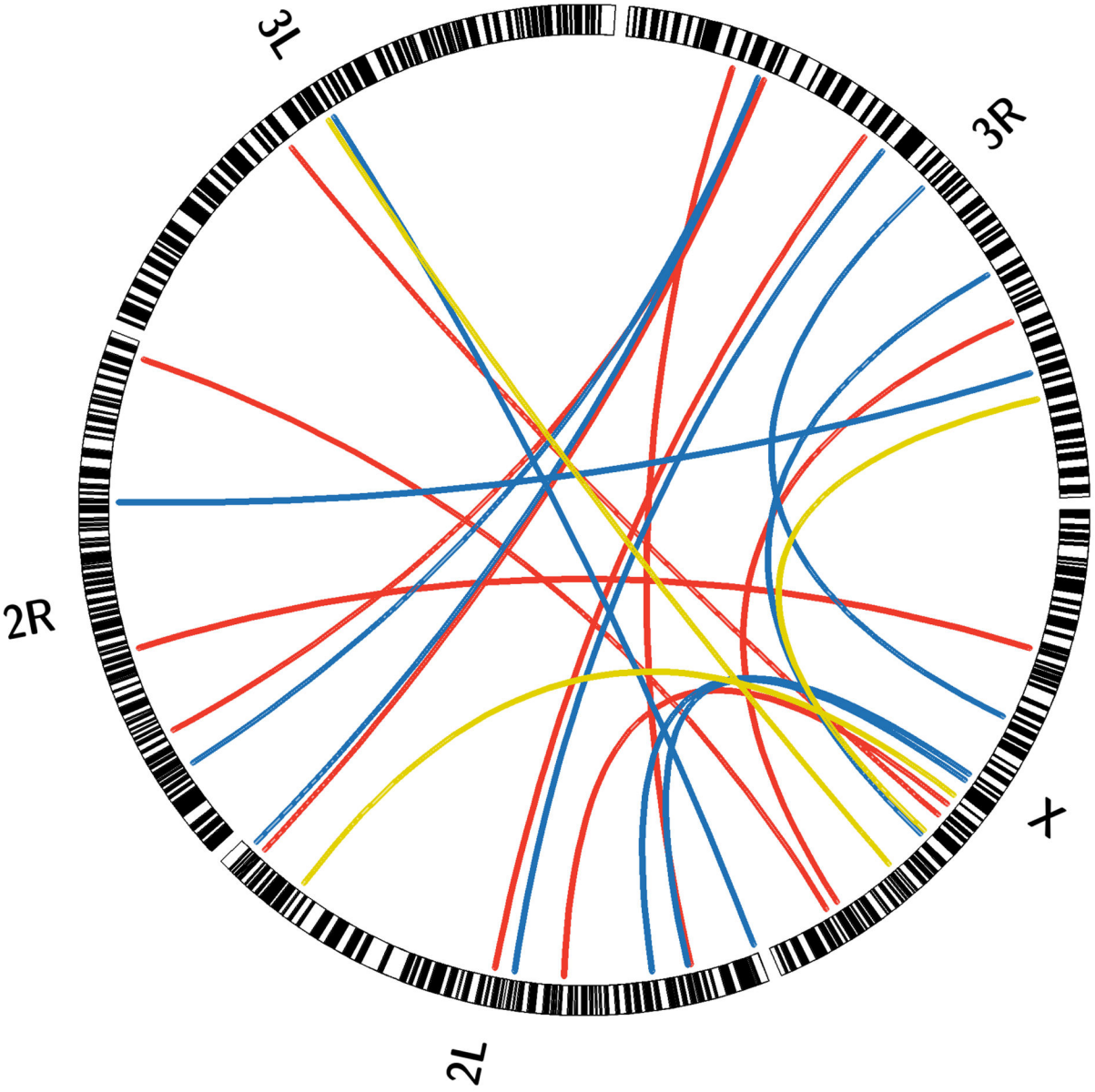


Figure 1. Locus pairs showing significant Genotype Ratio Distortion across the DSPR lines of Drosophila

The outer circle represent each chromosome arm. Each link represents a locus pair showing significant two-locus GRD. Yellow, blue and red links correspond respectively to RIL panel A-2, B-1 and B-2 (5% FDR corrected $P < 0.05$).

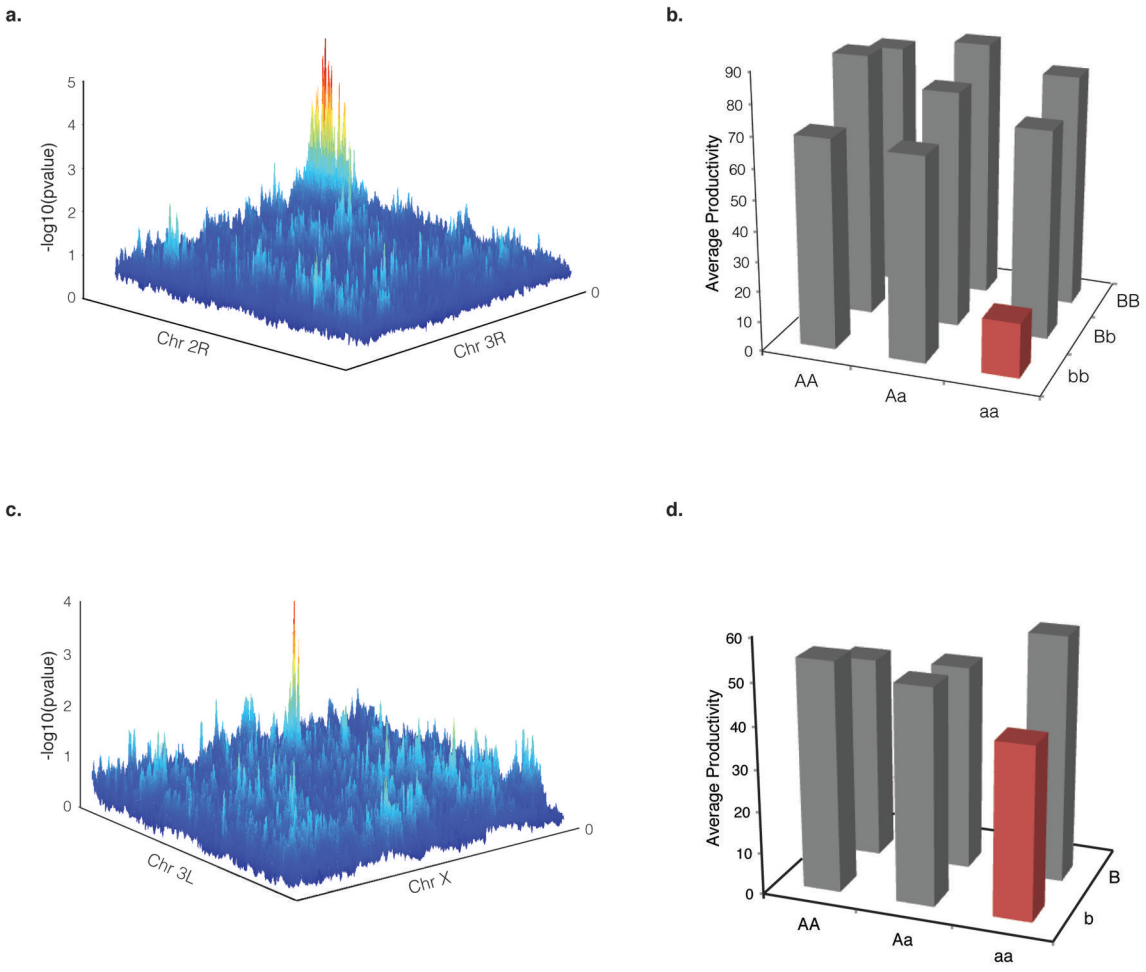


Figure 2. From missing genotypes to epistasis

a. GRD signature between all genotyped loci on chromosomes 2R and 3R in RIL panel B-2. **b.** Average productivity of each genotypic class recovered from 318 F_2 single-pair mating (progeny counts are F_3). As predicted from the GRD signal (in a.), haplotypes tagged by SNPs 2R:4806926 and 3R:5870973 show strong negative epistasis for the *aa;bb* genotypes, $P = 5.51121E-09$ LRT (indicated by the red bar) **c.** GRD between loci on chromosomes 3L and X in RIL panel A-2. **d.** Average productivity of each genotypic class recovered from 401 F_2 single pair mating. Haplotypes tagged by SNPs 3L: 11510853 and X: 16483812 show strong negative epistasis for the minor alleles on each haplotype *aa;bb,* $P = 8.25e-5$ LRT (indicated by the red bar).

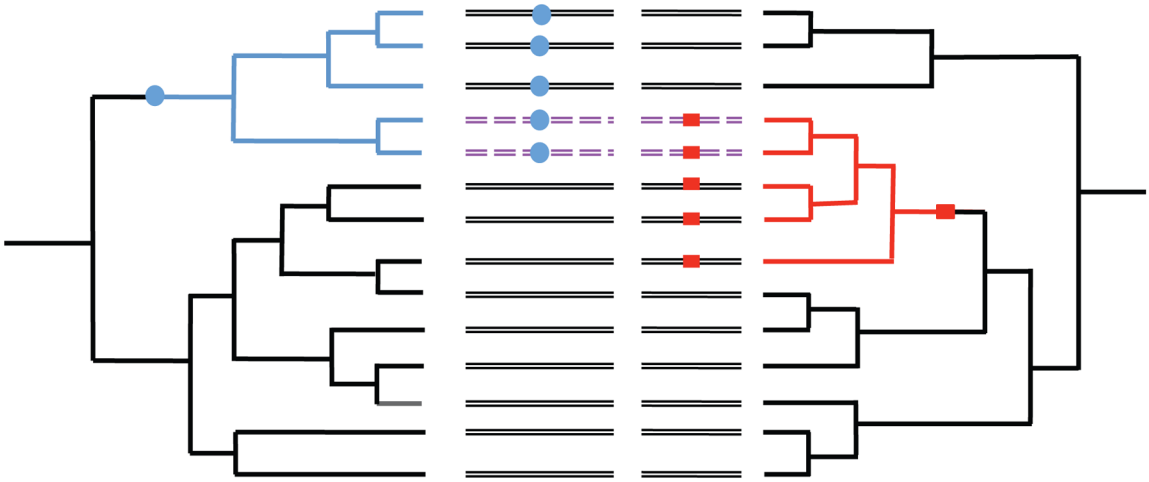


Figure 3. Model for unlinked loci with segregating pairs of incompatible alleles

The dendrograms on the left and the right represent the genealogies of two haplotypes segregating within a species. The blue dot and the red rectangle indicate the origins of incompatible mutations on each respective genealogy. On the left, derived blue alleles are incompatible with derived red alleles on the right. These genealogies yield the individuals shown in the center; wherein each line segment corresponds to a chromosome and each colored square indicates the derived incompatible allele. Importantly, these incompatible allele pairs are polymorphic in this sample of individuals, thus individuals who inherit both incompatible alleles have lower fitness than those with either none or only a single incompatibility.