Natural variation at a single gene generates sexual antagonism across fitness components in Drosophila

Highlights

- Drosophila use cuticular hydrocarbons as sexual signals and to avoid abiotic stress
- A fatty-acyl CoA reductase, DsFAR2-B, produces polymorphic cuticular hydrocarbons
- The polymorphism is sexually antagonistic, harming males but benefiting females
- DsFAR2-B shows a population genetic signal of balancing selection in nature

Authors

Bosco B. Rusuwa, Henry Chung, Scott L. Allen, Francesca D. Frentiu, Stephen F. Chenoweth

Correspondence

s.chenoweth@uq.edu.au

In brief

Rusuwa et al. study the diverse chemical sexual signals of Drosophila. In tropical Australia, a fatty-acyl CoA reductase gene DsFAR2-B creates polymorphic cuticular hydrocarbons with sexually antagonistic fitness effects. Population genomics shows balancing selection at DsFAR2-B, shedding light on how genetic variation is maintained in nature.
Natural variation at a single gene generates sexual antagonism across fitness components in *Drosophila*

Bosco B. Rusuwa,1,2,6 Henry Chung,3,6 Scott L. Allen,1,6 Francesca D. Frentiu,1,4 and Stephen F. Chenoweth1,5,7,*

1School of Biological Sciences, The University of Queensland, St Lucia, Australia
2Department of Biological Sciences, Chancellor College, University of Malawi, Zomba, Malawi
3Department of Entomology and Ecology, Evolution, and Behavior Program, Michigan State University, East Lansing, MI, USA
4School of Biomedical Sciences, Queensland University of Technology, Brisbane, QLD 4001, Australia
5Twitter: @steve_chenoweth
6These authors contributed equally
7Lead contact
*Correspondence: s.chenoweth@uq.edu.au
https://doi.org/10.1016/j.cub.2022.05.038

SUMMARY

Mutations with conflicting fitness effects in males and females accumulate in sexual populations, reducing their adaptive capacity.1,2 Although quantitative genetic studies indicate that sexually antagonistic polymorphisms are common,3–5 their molecular basis and population genetic properties remain poorly understood.6,7 Here, we show in fruit flies how natural variation at a single gene generates sexual antagonism through phenotypic effects on cuticular hydrocarbon (CHC) traits that function as both mate signals and protectors against abiotic stress8 across a latitudinal gradient. Tropical populations of *Drosophila serrata* have polymorphic CHCs producing sexual antagonism through opposing but sex-limited effects on these two fitness-related functions. We dissected this polymorphism to a single fatty-acyl CoA reductase gene, *DsFAR2-B*, that is expressed in oenocyte cells where CHCs are synthesized. RNAi-mediated disruption of the *DsFAR2-B* ortholog in *D. melanogaster* oenocytes affected CHCs in a similar way to that seen in *D. serrata*. Population genomic analysis revealed that balancing selection likely operates at the *DsFAR2-B* locus in the wild. Our study provides insights into the genetic basis of sexual antagonism in nature and connects sexually varying antagonistic selection on phenotypes with balancing selection on genotypes that maintains molecular variation.

RESULTS AND DISCUSSION

The contrasting means by which males and females maximize fitness favors the evolution of sexually dimorphic traits. Because the sexes share a genome, mutations usually affect traits in males and females similarly, which constrains the evolution of sexual dimorphism and produces sexually antagonistic effects on fitness.1 Although quantitative genetic studies have revealed that sexually antagonistic genetic variation is common in sexual populations,3,4,9–13 how sexual antagonism manifests at the genome level remains poorly understood.5,14,15 Little is known of the selective contexts across which alleles produce sexually antagonistic fitness effects or the population genetic processes shaping the DNA sequences that encode them.7 Required are integrative studies that can both characterize sexually antagonistic selection on shared traits and dissect the underlying sexually antagonistic loci.16 However, such investigations remain rare, and those available have detected resolved, rather than ongoing, antagonism.16–18

Here, we focus on a shared trait that is subject to both natural and sexual selection in fruit flies: namely, cuticular hydrocarbons (CHCs). Many insects use CHCs as contact pheromones for mate choice9,12 and also as waterproof agents to prevent desiccation.21,22 These dual functions simultaneously expose them to the combined evolutionary forces of natural and sexual selection, which can differ significantly between the sexes. CHCs are non-volatile long-chain hydrocarbons and exhibit parallel latitudinal clines across different continents in several fruit fly species, which is a strong indication that they experience spatially varying selection in nature.23–26 Notably, in populations closer to the equator, the relative abundance of short-chain CHCs decreases while the abundance of long-chain CHCs tends to increase.23,25 Experimental evolution26 and association studies25 implicate factors such as temperature and humidity as agents of spatially varying selection affecting these clines, which is consistent with the hypothesis that long-chain CHCs confer better protection against evaporative water loss than short-chain homologs.29

The CHC profiles of the Australian fly *Drosophila serrata* exhibit such a cline along the eastern Australian coastline. Importantly, this cline is genetically based and known to be shaped by natural selection (Figure 1A).26 Northernmost populations express
significantly lower amounts of short-chain CHCs (25 or fewer carbon atoms), while long-chain CHCs, such as (Z,Z)-5,9-heptacosadiene (5,9-C27), are expressed in higher amounts (Figure 1A). Examination of individual variation within populations revealed that some flies from northern populations express only trace amounts of all three shortest-chain CHCs: (Z,Z)-5,9-tetraicosadiene (5,9-C24), (Z,Z)-5,9-pentacosadiene (5,9-C25), and (Z)-9-pentacosene (9-C25) (Figure 1B). This can be considered an extreme phenotype because 5,9-C25, which is virtually non-existent in some northern flies, is usually the most abundant CHC in common phenotype individuals, comprising over 55% of the total CHC blend. Like the CHC cline itself, the frequency of these northern phenotype flies is also clinal, increasing in populations near the equator (Figure 1C). Furthermore, a principal components analysis clearly displays two distinct groups that are divided by the second principal component (PC2), one that is exclusive to the northern population of Cooktown and Cardwell and another that is found in all populations. (D) The geographic distribution of common (blue) and northern (orange) phenotype flies along the eastern Australian coastline. The proportion of northern phenotype flies increases towards the equator.

Figure 1. Clinal and polymorphic cuticular hydrocarbons

(A) Latitudinal clines in the D. serrata short-chain (<26 carbon) and long-chain (>25 carbon) dienes along the eastern Australian coastline show a reciprocal pattern. Shown are population means (+/− 1 s.e.) (after Chenoweth and Blows26).

(B) Two CHC phenotypes occur in far northern populations of flies: one where the short-chain CHCs 5,9-C24, 5,9-C25, and 9-C25 are expressed only in trace amounts (northern phenotype) and another, more and widespread, phenotype where these compounds are expressed at normal levels (common phenotype).

(C) Principal components analysis of the clinal CHC data displays two distinct phenotypes that are divided by the second principal component (PC2), one that is exclusive to the northern population of Cooktown and Cardwell and another that is found in all populations.

(D) The geographic distribution of common (blue) and northern (orange) phenotype flies along the eastern Australian coastline. The proportion of northern phenotype flies increases towards the equator.
To investigate potential fitness effects of this CHC variation, we tested for differences in mating success and abiotic stress resistance between northern and common phenotype flies from Cooktown, Australia, where the northern phenotype occurs with highest frequency. Because the northern phenotype appears to be recessive and we needed to assay fitness in an otherwise outbred random genetic background, we created multiple F1 crosses between wild-derived inbred lines that were fixed for the same two inbred lines (Figure S1). We took a bulked segregant approach to fine mapping, applying Illumina whole-genome resequencing to DNA pools of flies displaying either the common or northern CHC profiles. The fine mapping revealed a single narrow QTL peak, again on 3R (Figure 3C; Table S2).

Because our initial mapping cross was between quite distant geographical populations, results might have been complicated by population structure. To circumvent this complication and replicate the association on a within-population scale, we sequenced the genomes of 10 wild-derived inbred lines collected from Cooktown, where the northern phenotype is most common. Four lines were fixed for the common phenotype while six were fixed for the northern phenotype. Window-based summation of the fixed SNPs between the two CHC phenotypes across the genome revealed a peak of maximal differentiation at the same location on 3R identified by the bulked segregant analysis (Figure 3D).

The highest 30 kb interval of the QTL peak identified via bulked segregant and the genomic window of maximal differentiation in the ten Cooktown genomes contained five clustered genes with predicted fatty acyl-CoA reductase (FAR) activity (Figure 4A). FARs play a role in CHC synthesis by converting long-chain fatty acyl-CoAs to long-chain fatty alcohols before converting them to hydrocarbons via a P450 decarbonylase.31 FAR genes have been shown to be important for pheromone biosynthesis in other insects,32 for example moths, where changes in a single gene can produce very different pheromone blends.33 Moreover, a recent genome-wide association study (GWAS) implicated FAR genes as part of the genetic basis of CHC variation in D. melanogaster.34

Genomic clustering of FAR genes is common in Drosophila, and they are thought to evolve via duplication and subsequent gene loss.35 In the homologous genome region underlying the QTL peak, orthologs of the DsFAR genes in D. melanogaster are tightly

Figure 2. Sexually antagonistic fitness effects of CHC phenotypes
(A–C) Male and female mating success (A), desiccation resistance (B), and heat shock resistance (C). Points represent genotypic means, and bars are standard errors. Shown are overall CHC phenotype mean trait values +/- 1 s.e. Colored dots show individual F1 genotype mean trait values.

To investigate potential fitness effects of this CHC variation, we tested for differences in mating success and abiotic stress resistance between northern and common phenotype flies from Cooktown, Australia, where the northern phenotype occurs with highest frequency. Because the northern phenotype appears to be recessive and we needed to assay fitness in an otherwise outbred random genetic background, we created multiple F1 crosses between wild-derived inbred lines that were fixed for the same two inbred lines (Figure S1). We took a bulked segregant approach to fine mapping, applying Illumina whole-genome resequencing to DNA pools of flies displaying either the common or northern CHC profiles. The fine mapping revealed a single narrow QTL peak, again on 3R (Figure 3C; Table S2).

Because our initial mapping cross was between quite distant geographical populations, results might have been complicated by population structure. To circumvent this complication and replicate the association on a within-population scale, we sequenced the genomes of 10 wild-derived inbred lines collected from Cooktown, where the northern phenotype is most common. Four lines were fixed for the common phenotype while six were fixed for the northern phenotype. Window-based summation of the fixed SNPs between the two CHC phenotypes across the genome revealed a peak of maximal differentiation at the same location on 3R identified by the bulked segregant analysis (Figure 3D).

The highest 30 kb interval of the QTL peak identified via bulked segregant and the genomic window of maximal differentiation in the ten Cooktown genomes contained five clustered genes with predicted fatty acyl-CoA reductase (FAR) activity (Figure 4A). FARs play a role in CHC synthesis by converting long-chain fatty acyl-CoAs to long-chain fatty alcohols before converting them to hydrocarbons via a P450 decarbonylase.31 FAR genes have been shown to be important for pheromone biosynthesis in other insects,32 for example moths, where changes in a single gene can produce very different pheromone blends.33 Moreover, a recent genome-wide association study (GWAS) implicated FAR genes as part of the genetic basis of CHC variation in D. melanogaster.34

Genomic clustering of FAR genes is common in Drosophila, and they are thought to evolve via duplication and subsequent gene loss.35 In the homologous genome region underlying the QTL peak, orthologs of the DsFAR genes in D. melanogaster are tightly
clustered in a similar manner. However, D. melanogaster only harbors three consecutive FAR genes in this region (CG17562, CG17560, and CG14893). In D. serrata, two of these FAR genes appear as duplicates in the reference genome (Figure S2), which was generated by long-read sequencing of an inbred line displaying the common CHC phenotype. Interestingly, while all common genomes from Cooktown contained at least one copy of the duplication and therefore all five DsFAR genes, all northern genomes from Cooktown lacked the duplication and possessed only DsFAR1, DsFAR2-B, and DsFAR3-B (Figure 4A). However, this association between the duplication and CHC phenotype was not present in ten inbred lines from a different population; several lines from the D. serrata genomic reference panel (DsGRP) that display the common CHC phenotype also lack the duplication and only possess DsFAR1, DsFAR2-B, and DsFAR3-B. This observation excluded DsFAR2-A and DsFAR3-A as candidate genes.

To further narrow our search, we turned to in situ hybridization of the DsFAR candidates. CHC synthesis occurs in specialized cells called oenocytes. In situ hybridization of DsFAR1, DsFAR2-A/B, and DsFAR3-A/B in D. serrata showed that only DsFAR2-A/B is expressed in oenocytes of both common and northern adults (Figure 4B), narrowing the search to a single candidate gene. To verify the function of DsFAR2 in CHC synthesis, we silenced expression of DsFAR2’s ortholog (CG17560) in D. melanogaster oenocytes by RNA interference. Silencing CG17560 in D. melanogaster produced a strikingly similar phenotypic effect to that seen in D. serrata. Production of short-chain CHCs was dramatically reduced, while many long-chain compounds significantly increased in production (Figures 4C, S3, and S4).

Both the northern and common versions of DsFAR2-B appeared to be functional. The distinction between the northern and common versions of DsFAR2-B was a large number of non-synonymous substitutions. Of the 33 fixed differences between the northern and common alleles, 27 (~82%) were non-synonymous. Furthermore, the majority of the non-synonymous fixed differences (24 of 27, ~89%) were concentrated in the second half of the coding sequence, especially exon 4. This suggests that, as seen for pheromones in moths, evolutionary changes within a single FAR gene are a simple mechanism capable of generating diverse CHC blends in flies that may, in turn, have varied fitness effects through mate choice and stress resistance. Taken together, our results strongly implicate DsFAR2-B in generating a sexually antagonistic CHC polymorphism in D. serrata and provide much needed knowledge about how sexually antagonistic selection manifests at the genome level.

Sexually antagonistic selection can sometimes lead to a form of balancing selection that may be detectable using population genetic analyses. We estimated Tajima’s D for all five DsFAR2-B exons in two natural populations: Cooktown, where the polymorphism is most frequent and balancing selection could be occurring, and a population approximately 2000 km further south in Brisbane, which displays only the common CHC phenotype. In Cooktown, where the polymorphism is most frequent and balancing selection could be occurring, and a population approximately 2000 km further south in Brisbane, which displays only the common CHC phenotype, where balancing selection is less likely. Tajima’s D was not consistent with balancing selection at any DsFAR2-B locus in the Cooktown population, a pattern consistent with balancing selection. In contrast, Tajima’s D was not consistent with balancing selection at any DsFAR2-B exon in the Brisbane population (all p > 0.769), with
Figure 4. A fatty acyl-CoA reductase affects CHC phenotype

(A) Comparison of the orthologous region of *D. melanogaster* where the major QTL was mapped in *D. serrata*. The *D. melanogaster* reference genome contains three fatty acyl-coA reductase genes (CG17562, CG17560, and CG14893), which are orthologous to *D. serrata* DsFAR1, DsFAR2, and DsFAR3, respectively. The

(legend continued on next page)
all D values negative and ranging from −0.59 to −1.16. Positive Tajima’s D for DsFAR2-B in Cooktown strongly suggests balancing selection at this locus but only in populations where the CHC polymorphism is present.

For phenotypic traits with different fitness optima in males and females, sexual antagonism arises when mutations affect traits in similar ways in the two sexes.2,12 An interesting question is how the polymorphism, which affects the CHCs of both sexes, generates the sex-specific fitness effects observed. In terms of mating success, there is evidence for both divergent fitness optima and sex differences in the strength of sexual selection on CHCs in D. serrata. While CHCs are phenotypically correlated with male and female mating success in D. serrata, the associations are always far stronger in males than in females.38-41 Recent experiments suggest phenotypic CHC-mating success associations may be causal only in males.42 Together, these results may help to explain why the DsFAR2-B polymorphism affects male, but not female mating success. With respect to abiotic stress, the increase in long-chain CHCs that accompanies the short-chain CHC reduction in the northern phenotype may account for increased desiccation resistance,43 but it does not account for the effect being restricted to females. Interestingly, responses of males and females to desiccation and thermal stress are often very different in Drosophila flies. For example, short-term exposure of D. melanogaster to low relative humidity conditions leads to a plastic response that increases female, but not male, desiccation resistance.44 Moreover, sex-dependent responses of CHCs have been detected during experimental adaptation to low relative humidity environments in D. melanogaster27 and to higher temperatures in D. simulans.28 These results point to sexually dimorphic physiology in D. serrata and suggest that CHCs may differ in their importance for abiotic stress resistance in males and females. Future physiological studies will be required to understand how such sex differences arise.

The long-chain versus short-chain phenotypic effect of the DsFAR2-B polymorphism closely aligns with the D. serrata latitudinal cline, which suggests that expression of low amounts of short-chain but high amounts of long-chain CHCs is indeed beneficial in warmer tropical populations. In the far north, where temperatures are higher, this benefit to female stress resistance may outweigh the cost to decreased male sexual attractiveness, setting up a scenario of sexual antagonism. In contrast, the cost-benefit relationship likely changes moving south to cooler populations, where northern phenotype individuals have not been observed. It therefore appears that climatic variation is key to determining whether sexual antagonism may or may not arise in a given population. Another potentially influential factor affecting maintenance of the polymorphism is ongoing migration. It is possible that recurrent gene flow of common alleles into Cooktown also helps to prevent fixation of the northern allele. Developing a better understanding of the relative contributions of sexual antagonism and migration will require further investigation but will be important to understand how this polymorphism has been maintained.

There is considerable interest in understanding how sexually antagonistic selection manifests at the genome level.5,15 However, to date, very few investigations have provided gene-level resolution of experimentally confirmed sexually antagonistic loci.16 Our detailed analysis, which spans from phenotypic variation in the wild to identifying the underlying genetic mechanism, confirms the existence of large-effect sexually antagonistic loci in nature. Further, it highlights how a variant with similar phenotypic effects on shared traits can generate a sexual antagonism that plays out across multiple components of fitness, involving both natural and sexual selection. The interplay between clinal variation and the presence of the polymorphism suggests that population differences in sex-specific natural and sexual selection may be important determinants of the sexually antagonistic loci segregating in any given population. The present study provides us with a platform upon which to reconstruct the evolutionary history of a sexually antagonistic locus and investigate how these sex-specific fitness effects arise.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Fitness assays
  - QTL mapping
  - Bulked segregant analysis
  - Genome resequencing
  - Genotyping the DsFAR2 duplication using short-read data
  - Population genetic analysis
  - In situ hybridisation to adult D. serrata oenocytes
  - D. melanogaster RNAi and CHC analyses
- **METHOD DETAILS**
  - Clinal patterns
  - Desiccation resistance
  - Heat shock stress survival
  - Competitive mating success
  - QTL mapping

D. serrata reference genome contains a duplication of DsFAR2 and DsFAR3 that is not present in some sequenced D. serrata genomes, which contain an alternate haplotype in this region. This duplication is not associated with the CHC polymorphism (see results and discussion) (see also Figure S2).

(B) In situ hybridization indicating expression of DsFAR2-A/B in D. serrata oenocytes in common and northern males and females (indicated by the ribbon-like bands) and lack of expression for DsFAR1 and DsFAR3-A/B.

(C) RNAi knockdown of the DsFAR2 ortholog (CG17560) in D. melanogaster oenocytes indicates a major effect on CHCs. Shown are mean +/- 95% confidence interval for male CHCs expressed at a minimum of 20 ng/fly. Asterisks indicate significant differences in mean CHC expression (p < 0.05) between the control and RNAi treatment within each reciprocal cross (t test).

See Figures S3 and S4 for all CHCs in both sexes.
AUTHOR CONTRIBUTIONS

B.B.R. and S.F.C. designed the study. B.B.R. and F.D.F. performed the QTL analysis. H.C. performed the RNAi experiments, the competitive mating success analysis. S.L.A. and B.B.R. conducted molecular population genetic analyses. B.B.R. and S.F.C. drafted the manuscript. We thank L. Rowe, T. Connallon, N. Smith, and J.G. Millar for comments on the manuscript and discussion. We also thank Prof. Sean B. Carroll for financial support (part of this work was done by H.C. as a postdoc in the Carroll lab at the University of Wisconsin-Madison), and J.G. Millar (University of California, Riverside) for analysis, identification, and quantitation of the D. melanogaster CHC profiles in the RNAi experiments. H.C. is supported by a National Science Foundation grant (2054773). Funding for this project was supported by The University of Queensland and the Australian Research Council.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES


## KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deposited data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. serrata reference genome (NCBI)</td>
<td>36</td>
<td>GCA_002093755.1</td>
</tr>
<tr>
<td>D. serrata Hi-C scaffolded genome</td>
<td>This paper</td>
<td><a href="https://doi.org/10.5061/dryad.0p2ngf245">https://doi.org/10.5061/dryad.0p2ngf245</a></td>
</tr>
<tr>
<td>D. serrata Annotation Release 100 (NCBI)</td>
<td>N/A</td>
<td>GCF_002093755.1</td>
</tr>
<tr>
<td>Bulked segregant sequence data (NCBI)</td>
<td>This paper</td>
<td>PRJNA603628</td>
</tr>
<tr>
<td>Cooktown inbred lines sequence data (NCBI)</td>
<td>This paper</td>
<td>PRJNA604036</td>
</tr>
<tr>
<td>Brisbane inbred lines sequence data (NCBI)</td>
<td>37</td>
<td>PRJNA419238</td>
</tr>
<tr>
<td>Cline CHC data</td>
<td>26</td>
<td><a href="https://doi.org/10.5061/dryad.0p2ngf245">https://doi.org/10.5061/dryad.0p2ngf245</a></td>
</tr>
<tr>
<td>Mating success</td>
<td>This paper</td>
<td><a href="https://doi.org/10.5061/dryad.0p2ngf245">https://doi.org/10.5061/dryad.0p2ngf245</a></td>
</tr>
<tr>
<td>Heat shock</td>
<td>This paper</td>
<td><a href="https://doi.org/10.5061/dryad.0p2ngf245">https://doi.org/10.5061/dryad.0p2ngf245</a></td>
</tr>
<tr>
<td>Desiccation</td>
<td>This paper</td>
<td><a href="https://doi.org/10.5061/dryad.0p2ngf245">https://doi.org/10.5061/dryad.0p2ngf245</a></td>
</tr>
<tr>
<td>QTL data</td>
<td>This paper</td>
<td><a href="https://doi.org/10.5061/dryad.0p2ngf245">https://doi.org/10.5061/dryad.0p2ngf245</a></td>
</tr>
<tr>
<td>Aligned DsFAR2B orthologs</td>
<td>This paper</td>
<td><a href="https://doi.org/10.5061/dryad.0p2ngf245">https://doi.org/10.5061/dryad.0p2ngf245</a></td>
</tr>
<tr>
<td>RNAi CHC data</td>
<td>This paper</td>
<td><a href="https://doi.org/10.5061/dryad.0p2ngf245">https://doi.org/10.5061/dryad.0p2ngf245</a></td>
</tr>
</tbody>
</table>

**Experimental models: Organisms/strains**

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. serrata common Inbred line from Forster (FORS4)</td>
<td>This paper</td>
<td>FORS4</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Cooktown (CTN44)</td>
<td>This paper</td>
<td>CTN44</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Cooktown (CTN21)</td>
<td>This paper</td>
<td>CTN21</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Cooktown (134)</td>
<td>This paper</td>
<td>134</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Cooktown (145)</td>
<td>This paper</td>
<td>145</td>
</tr>
<tr>
<td>D. serrata northern Inbred line from Cooktown (CTN42)</td>
<td>This paper</td>
<td>CTN42</td>
</tr>
<tr>
<td>D. serrata northern Inbred line from Cooktown (CTN10)</td>
<td>This paper</td>
<td>CTN10</td>
</tr>
<tr>
<td>D. serrata northern Inbred line from Cooktown (CTN18)</td>
<td>This paper</td>
<td>CTN18</td>
</tr>
<tr>
<td>D. serrata northern Inbred line from Cooktown (CTN32)</td>
<td>This paper</td>
<td>CTN32</td>
</tr>
<tr>
<td>D. serrata northern Inbred line from Cooktown (CTN34)</td>
<td>This paper</td>
<td>CTN34</td>
</tr>
<tr>
<td>D. serrata northern Inbred line from Cooktown (180)</td>
<td>This paper</td>
<td>180</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Brisbane (12)</td>
<td>37</td>
<td>DsGRP12</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Brisbane (28)</td>
<td>37</td>
<td>DsGRP28</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Brisbane (49)</td>
<td>37</td>
<td>DsGRP49</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Brisbane (63)</td>
<td>37</td>
<td>DsGRP63</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Brisbane (133)</td>
<td>37</td>
<td>DsGRP133</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Brisbane (165)</td>
<td>37</td>
<td>DsGRP165</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Brisbane (171)</td>
<td>37</td>
<td>DsGRP171</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Brisbane (193)</td>
<td>37</td>
<td>DsGRP193</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Brisbane (195)</td>
<td>37</td>
<td>DsGRP195</td>
</tr>
<tr>
<td>D. serrata common Inbred line from Brisbane (228)</td>
<td>37</td>
<td>DsGRP228</td>
</tr>
<tr>
<td>D. melanogaster oenocyte GAL4 driver line (oenoGAL4)</td>
<td>45</td>
<td>oenoGAL4</td>
</tr>
<tr>
<td>D. melanogaster RNAi line (UAS-CG17560RNAi-KK)</td>
<td>VDRC Stock Center</td>
<td>104756</td>
</tr>
<tr>
<td>D. melanogaster y,w[1118] with attP landing site</td>
<td>Bloomington Drosophila Stock Center</td>
<td></td>
</tr>
<tr>
<td>D. serrata used for in situ hybridisation</td>
<td>Drosophila Species Stock Center</td>
<td>14028-0681.05</td>
</tr>
</tbody>
</table>

**Oligonucleotides**

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>DsFAR2-A/B (CG17560) Forward in-situ hybridization probe SerCG17560-probeF (S’-TGGCCTGTGCTGGCACACGGG-3’)</td>
<td>This paper</td>
<td>N/A</td>
</tr>
</tbody>
</table>

(Continued on next page)
RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Stephen F. Chenoweth (s.chenoweth@uq.edu.au).

Materials availability
This study did not generate new unique reagents.

Data and code availability

- Raw gas chromatography traces of CHCs used for the clinal analysis of phenotypes have been deposited at Dryad and are publicly available as of the date of publication. DOIs are listed in the key resources table. Mating and abiotic stress trait data used to assay fitness have been deposited at Dryad and are publicly available as of the date of publication. DOIs are listed in the key resources table. D. serrata reference genome: Genome (Dser1.0) has been deposited at GenBank and are publicly available

---

### Reagent or Resource Table

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>DsFAR2-A/B (CG17560) Reverse in-situ hybridization probe SerCG17560-probeR (5'-GGGTATTGTGGAAAAATTAAGGC-3')</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>DsFAR1 (CG17562) Forward in-situ hybridization probe SerCG17562-probeF (5'-ATGGGAGAAACTCGCTGAAGTGCC-3')</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>DsFAR1 (CG17562) Reverse in-situ hybridization probe SerCG17562-probeR (5'-ATGGGAGAAACTCGCTGAAGTGCC-3')</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>DsFAR3-A/B (CG14893) Forward in-situ hybridization probe SerCG14893-probeF (5'-TGCGCAGAGCTCGAAAAACGGC-3')</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>DsFAR3-A/B (CG14893) Reverse in-situ hybridization probe SerCG14893-probeR (5'-GGTTATTGTGTGATAAAAGAAGGC-3')</td>
<td>This paper</td>
<td>N/A</td>
</tr>
</tbody>
</table>

### Software and algorithms

<table>
<thead>
<tr>
<th>Software and algorithms</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>fastQC</td>
<td>46</td>
<td><a href="https://www.bioinformatics.babraham.ac.uk/projects/fastqc/">https://www.bioinformatics.babraham.ac.uk/projects/fastqc/</a></td>
</tr>
<tr>
<td>Samtools</td>
<td>48</td>
<td><a href="https://sourceforge.net/projects/samtools/files/samtools/">https://sourceforge.net/projects/samtools/files/samtools/</a></td>
</tr>
<tr>
<td>PoPoolation2</td>
<td>49</td>
<td><a href="https://sourceforge.net/p/popolation2/wiki/Main/">https://sourceforge.net/p/popolation2/wiki/Main/</a></td>
</tr>
<tr>
<td>Genome Analysis Tool Kit</td>
<td>50</td>
<td><a href="https://github.com/broadinstitute/gatk/releases">https://github.com/broadinstitute/gatk/releases</a></td>
</tr>
<tr>
<td>Integrative Genomics Viewer (IGV)</td>
<td>51</td>
<td><a href="https://software.broadinstitute.org/software/igv/download">https://software.broadinstitute.org/software/igv/download</a></td>
</tr>
<tr>
<td>Molecular Evolutionary Genetic Analysis (MEGA X)</td>
<td>52</td>
<td><a href="https://www.megasoftware.net/download_form">https://www.megasoftware.net/download_form</a></td>
</tr>
<tr>
<td>The R Project for Statistical Computing (R)</td>
<td>R Core Team, 2013</td>
<td><a href="https://www.r-project.org/">https://www.r-project.org/</a></td>
</tr>
<tr>
<td>R/qtl</td>
<td>53</td>
<td><a href="https://rqlt.org/">https://rqlt.org/</a></td>
</tr>
<tr>
<td>PopGenome</td>
<td>54</td>
<td><a href="https://cran.r-project.org/web/packages/PopGenome/index.html">https://cran.r-project.org/web/packages/PopGenome/index.html</a></td>
</tr>
</tbody>
</table>
available as of the date of publication. Accession numbers are listed in the key resources table. Hi-C based scaffolds have been deposited at Dryad and are publicly available as of the date of publication. DOIs are listed in the key resources table. Genotypic and phenotypic data used for the F2 binary QTL analysis have been deposited at Dryad and are publicly available as of the date of publication. DOIs are listed in the key resources table. Bulked segregant sequence data have been deposited at SRA and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Cooktown population sequence data have been deposited at SRA and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Brisbane population sequence data have been deposited at SRA and are publicly available as of the date of publication. DOIs are listed in the key resources table.

This paper does not report original code.

Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

**Fitness assays**

We examined the fitness effects of the CHC polymorphism by assaying mating success and two stress traits; desiccation resistance and heat shock resistance on male and female flies. To create flies that displayed either the northern or common phenotype (the northern form is recessive), but at the same time avoid the effects of inbreeding, we created a series of F1 genotypes formed from crosses within two panels of wild-derived inbred lines from Cooktown (15°28′11.62″S, 145°14′55.27″E) that were fixed for the different CHC phenotypes. This approach allowed us to generate a large number of flies for a set of 12 northern CHC expressing genotypes and a set of 12 common CHC expressing genotypes.

Parental flies were sexed as virgins under light CO2 anaesthesia and then held individually in vials until they were 4 days old. Mating vials were established by pairing 3 males and 3 females from different lines in vials containing 10mL of food media sprinkled with live yeast. Flies were allowed to mate for 4 days after which the vials were cleared. The F1 offspring flies were sexed as virgins and held at a density of five flies per vial before the mating, desiccation and heat tolerance trials commenced. All flies were reared at a constant temperature of 25°C on a 12h:12h light:dark cycle. Flies were maintained on the same standard fly diet comprising 36g agar, 108g raw sugar and 74g yeast, cooked in 2L of water combined with 24mL propionic acid and 12mL nipagin.

**QTL mapping**

To create a mapping population, an F2 intercross was made between two highly inbred lines (20 generations of full-sib mating) of *D. serrata* that were derived from two natural populations at opposing ends of the species’ eastern Australian distribution (Cooktown: CTN42, Forster: FORS4). CTN42 expresses the northern CHC phenotype where the production 5,9-C24, 5,9-C25 and 9-C25 is massively reduced relative to the FORS4 line which expresses the common CHC phenotype. A number of chromosomal inversions have been documented in *D. serrata* and so to minimise their potential effects in the QTL mapping, the founder lines were first confirmed to be homosequential using established protocols for polytene chromosome squashes. Newly emerged virgin F2 flies were sexed (180 males, 203 females) and held singly in fresh vials for five days before being assayed for their CHC profiles.

**Bulked segregant analysis**

The same parental lines used to establish the F2 intercross (FORS4 common phenotype, CTN42 northern phenotype) were used to establish an advanced mass-bred population. We again performed reciprocal crosses to generate a very large population. The population was maintained in 32 glass bottles (500mL) for 60 non-overlapping generations resulting in a census size of approximately 6500 flies. We favoured the “large outbred” approach over randomly crossing pairs of flies every generation because a larger population size could be maintained with this setup, as it does not require sexing of flies at each generation. Flies were maintained on standard laboratory yeast medium, at 25°C with a 12h:12h light:dark photoperiod. We sampled flies for CHC phenotyping and subsequent bulked segregant analysis at generation 60. Newly emerged virgin flies were sexed (n = 600 female flies) under light CO2 anaesthesia and held singly in food vials for five days before being assayed for their CHC profile through gas chromatography as described above. All of the flies phenotyped were individually preserved in 95% ethanol and stored at -80°C for subsequent DNA extraction.

**Genome resequencing**

As our mapping crosses were geographically broad, we attempted to replicate the association at *DsFAR2* by sequencing genomes from a sample of ten wild-derived inbred lines (17 generations of full-sib mating) from a single population in Cooktown where the northern phenotype has been found previously with highest frequency. CHC phenotyping after inbreeding confirmed that the northern phenotype also occurred at high frequency in this sample. Six lines (CTN10, CTN18, CTN32, CTN34, and 180) expressed the northern phenotype whereas the common phenotype occurred in four lines (CTN44, CTN21, 134, 145).
Genotyping the DsFAR2 duplication using short-read data
The inbred lines used for genome resequencing detailed above were used to genotype the DsFAR2 locus.

Population genetic analysis
Population genetic analyses for the Cooktown population where the northern phenotype occurs at highest frequency were performed using the genome resequencing data from 10 wild-derived inbred lines detailed above. We also performed the same analyses in a random sample of ten lines sampled from the D. serrata Genomic Reference Panel (DsGRP)37 (lines 12, 28, 49, 63, 133, 171, 193, 195, and 228), which is a sample from a single endemic population in Brisbane where the northern phenotype has not been observed.

In situ hybridisation to adult D. serrata oenocytes
In situ hybridisation were performed on both FORS4 and CTN42 males and females that are detailed above in the QTL mapping section.

D. melanogaster RNAi and CHC analyses
Oenocyte specific RNAi of CG17560 (the DsFAR2 ortholog, see Figure S2) in D. melanogaster was performed at 25°C by crossing an oenocyte GAL4 driver line, oenoGAL4 (PromE(800) line 2M)45 and a RNAi line, UAS-CG17560RNAi-KK (obtained from the Vienna Drosophila RNAi center). Controls were performed by crossing oenoGAL4 to y,w[1118] with attP landing site VIE-260B (line which the RNAi line was generated from). Reciprocal crosses were performed for each condition and flies are collected and separated by sex within eight hours of eclosion.

METHOD DETAILS
Clinal patterns
Figures 1A, 1C, and 1D were drawn after reanalysis of raw gas chromatography traces from flies analysed in the multi-population dataset in Chenoweth and Blows.26 The principal component analysis of log contrast CHC expression was performed via the prcomp function in R with scaling to unit variance set to TRUE.56,57

Desiccation resistance
A modified version of published procedures58,59, was followed in this experiment. Desiccation resistance was measured by enclosing 20 flies of each sex from each F1 genotype in a standard 30ml Drosophila vial without any medium. Approximately 3g of fresh Drierite desiccant was suspended on the top end of the vial on cotton gauze. The vial was then sealed off with parafilm to sustain low humidity. Trial runs indicated that this set up reduced humidity in the vial to 10% in 30 minutes. Flies were observed at hourly intervals for death, as indicated by failure to right themselves or to move their legs when their vials were tapped or inverted. Because temperature parameters alone can account for up to 97% of variability in desiccation tolerance,60 all tests were carried out at 25°C.

Heat shock stress survival
Heat tolerance was measured as survival after a high (potentially lethal) temperature stress61 at 4-6 days post-eclosion. Twelve replicates of twenty flies for each sex per line cross were enclosed in an empty stoppered vial and placed in a constant temperature cabinet set at 38°C for 30 minutes. Flies were then transferred to fresh media vials and left to recover at 25°C for 24 hours before being scored for survival.

Competitive mating success
Disparities in male and female mating success between phenotypes were assessed by determining whether a focal northern or common fly preferred to mate with a northern or common suitor when potential mates of both types were availed to them. Briefly, individuals were collected as virgins soon after eclosion and held separately by sex in yeasted food vials for four days. On the fourth day, a focal fly (northern/common) was presented with two flies of the opposite sex, one from its own phenotype (not from the same F1 genotype as the focal fly) and another from a contrasting phenotype. Once copulation commenced, the unsuccessful fly was taken out of the vial and identified. To facilitate identification of the two potential suitors, one of either fly phenotype was slightly wing-clipped. The number of wing-clipped competing flies (left-wing clip and right-wing clip) was kept even between the two phenotypes to balance out any effects of wing clipping on courtship success.

QTL mapping
CHCs were extracted from individual flies by washing each fly in 100mL of hexane in a microvial insert for three minutes and then vortexing for one minute.62 The CHC samples were run on an Agilent Technologies 6890N gas Chromatograph (Wilmington, Delaware, United States). Individual fly CHC profiles were derived by integrating the area under the following nine peaks using the Agilent ChemStation Software (version Rev B.04.02), in the order of their retention times: 5,9-C24, 5,9-C25, 9-C24, 9-C26, 2-MeC26, 5,9-C27 heptacosadiene, 2-MeC28, 5,9-C29 and 2-MeC30.63 All the flies phenotyped were then individually preserved in 95% ethanol and kept in a -80°C freezer pending DNA extraction.
SNP discovery and genotyping for the F2 cross has been previously described. Briefly, SNP genotyping of individual flies on 61 SNPs was carried out on a SEQUENOM® MASS ARRAY PLATFORM at the Australian Genome Research Facility (AGRF) using two multiplexes of 30 and 31 SNPs, with approximately 10ng of genomic DNA for each multiplex assay. Out of the 61 markers, three were discarded (s10, s36, and s37) because no difference in genotype calls was detected between the parental lines.

Bulked segregant analysis
A total of 85 northern phenotype flies and 85 common flies were selected for DNA extraction. Two DNA bulks were made, one containing DNA from northern flies and the other DNA from common flies, effectively resulting in two types of pooled DNA sample that were expected to differ genetically at the underlying QTL, but were theoretically undifferentiated for all other regions. The sequencing of these DNA pools was technically replicated using two independent sequencing libraries. Phenol-chloroform extraction was used to obtain fly genomic DNA.

Sequencing of the DNA bulks was performed using an Illumina HiSeq 2000 sequencing machine which produced 90 base pair paired-end reads with a median insert size of ~ 490 base pairs. Quality control of the DNA sequence data was performed using FastQC, and no problems were detected. The D. serrata reference genome is based on one of the parental lines used in this experiment, FORS4 Allen et al. 2017, so we did not need to re-sequence this line. We sequenced the CTN42 line so that the parental line could be determined. The parental CNT42 line was sequenced at the Australian Genome Research Facility (AGRF), whereas the F60 samples were sequenced at The Beijing Genomics Institute (Hong Kong, China). Table S1 provides details of the sequencing libraries, numbers of reads and raw coverage for each pool.

Common and northern pool sequence reads from this experiment were aligned to a version of the D. serrata reference available on NCBI (BioProject: PRJNA355616) that we have scaffolded using DovetailTM Hi-C technology (Dovetail genomics). The scaffolding greatly increased the contiguity of the assembly from an N50 of just under 1 Mbp to an N50 of 30.3 Mbp. To place the scaffolds onto chromosomes, we used 78 physical and linkage markers for D. serrata that have known chromosomal locations. Mapping was performed using the bwa-mem algorithm of the Burrows Wheeler Alignment (BWA) software (version 0.7.12). Local realignment was then performed around indels using the Genome Analysis Tool Kit (GATK, version 2.5-2-gf57256b) following best practices. Average depth of coverage across the 10 genomes was 22x. SNPs relative to the reference genome were called for each of the 10 inbred lines using samtools and bcftools.

Genotyping the DsFAR2 duplication using short-read data
We were able to determine whether a line contained the DsFAR duplication using only short-read data through visual inspection of reads that were mapped to the D. serrata reference genome for genome resequencing, which has both copies of the gene using IGV. It was apparent that some lines were missing the duplication containing DsFAR2-A and DsFAR3-A. For these lines, mapping details for all other lines are as in the genome resequencing section above. We also performed the same analyses in a random sample of ten lines sampled from the Genomic Reference Panel (DsGRP) which is a sample from a single endemic population in Brisbane where the northern phenotype has not been observed.

Population genetic analysis
We estimated Tajima’s D to gain insights into the population genetic processes shaping variation around the DsFAR2-B locus. The six inbred lines with the northern phenotype were remapped to a version of the D. serrata genome that had the ~8 kbp duplicated region masked to improve mapping quality. Mapping details for all other lines are in the genome resequencing section above. We also performed the same analyses in a random sample of ten lines sampled from the D. serrata Genomic Reference Panel (DsGRP) which is a sample from a single endemic population in Brisbane where the northern phenotype has not been observed.

In situ hybridisation to adult D. serrata oenocytes
In situ hybridisation to four- to five-day old adult oenocytes were performed with RNA probes as described previously. Probes for in situ hybridisation were synthesised from D. serrata (UCSD Stock no. 14028-0681.05) cDNA using primers SerCG17560-probeF (5'-TGGCCCTGCTGCTGCCACCGG-3') and SerCG17560-probeR (5'-GGGATGGGTGGAAAATTAGGC-3') for DsFAR2-A/B (CG17560), SerCG17562-probeF (5'-ATGGGAGAAACTCGCTGGG-3') and SerCG17562-probeR (5'-GGGATGGTGTGGAAAATTAAGGC-3') for DsFAR1 (CG17562), and SerCG14893-probeF (5'-TGCCAGCATCTGGAAAACGGC-3') and SerCG14893-probeR (5'-GGGATGGTGTGGAAAATTAAGGC-3') for DsFAR3-A/B (CG14893).
D. melanogaster RNAi and CHC analyses
For quantitation of CHCs, five 5-day old flies were placed in 100 μL of hexane spiked with 50 μg/mL of C26 as an internal standard and analysed using a Hewlett-Packard (now Agilent; Santa Clara CA) 6890 gas chromatograph (GC) interfaced to an H-P 5973 mass selective detector. The GC was fitted with a DB-17 30 m × 0.25 mm i.d. capillary column (J&W Scientific, Folsom CA) with splitless injections and a temperature program of initial temperature 100°C/0 min, 20°C per min to 160°C/0 min, 4°C per min to 280°C, hold 25 min. Detailed methods were described previously.70 Each individual CHC peak was quantified using peak area.

Phylogenetic relationships of reductases at the DsFAR locus
Amino acid sequences were downloaded from Flybase (www.flybase.com), based on reported homologues from Finet et al.35 and aligned using MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle/). Evolutionary history of these reductases was inferred by using the Maximum Likelihood method and performing bootstrapping and obtain 1000 bootstrapped trees. All analyses were conducted in MEGA X.52

QUANTIFICATION AND STATISTICAL ANALYSIS
Desiccation resistance
The time elapsed until 10 flies were dead (completely immobile) was recorded and defined as LT50. As the number of dead flies neared 10, vials were checked for death every 20 min. In total 78 trials were conducted for the *northern* phenotype (females: \(n = 33\), males: \(n = 45\)) and 79 trials for the *common* phenotype (females: \(n = 37\), males: \(n = 42\)). Desiccation resistance was analysed using mixed effects linear models fitted using the MIXED (Gaussian response variables) procedures of SAS (version 9.3, SAS Inst. Cary, NC). Desiccation resistance was modelled as:

\[
Y = \mu + S + C + |S \times C| + |S \times GC| + \varepsilon
\]  

(Equation 1)

where, \(Y\) was the time taken for 50% of the test flies to die, \(\mu\) is the intercept, \(S\) is the sex of the flies, \(C\) is the CHC phenotype of the fly, \(GC\) is the random effect of genotype nested within CHC phenotype and \(\varepsilon\) is the random error. When the sexes are analysed together, the experiment is effectively a split-plot design and so line nested within CHC phenotype and an interaction between sex and line within CHC phenotype were also fitted. The error term for the CHC polymorphism effect is \(\varepsilon\) in this linear model. The sex x genotype interaction was tested via an F-test and within-sex effects via t-tests. Statistical significance was defined as a p-value < 0.05.

Heat shock stress survival
The number of survivors per replicate vial was scored. In total 118 trials were conducted for the *northern* phenotype (females: \(n = 71\), males: \(n = 47\)) and 130 trials for the *common* phenotype (females: \(n = 69\), males: \(n = 61\)). Heat shock stress survival was analysed using mixed effects linear models fitted using the GLIMMIX (binary responses) procedures of SAS (version 9.3, SAS Inst. Cary, NC). Heat shock stress survival was modelled as:

\[
Y = \mu + S + C + |S \times C| + GC + |S \times GC| + \varepsilon
\]  

(Equation 2)

where \(Y\) was survival (dead or alive) after 24 hours post-exposure to heat shock, \(\mu\) is the intercept, \(S\) is the sex of the flies, \(C\) is the ‘northern’ status of the flies, \(GC\) is the genotype for the ‘northern’ status of the flies and \(\varepsilon\) is the random error. The sex x genotype interaction was tested via an F-test and within-sex effects via t-tests. Statistical significance was defined as a p-value < 0.05.

Competitive mating success
A total of 785 mating trials were conducted; 308 female choice trials and 377 male choice trials. Mating success was analysed using mixed effects linear models fitted using the GLIMMIX (binary responses) procedures of SAS (version 9.3, SAS Inst. Cary, NC). The mating assay data was analysed in males and females separately. Mating success was assessed using a binary mixed model through Proc GLIMMIX as:

\[
Y = \mu + SC + FC + |SC \times FC| + G_{SC} + G_{FC} + \varepsilon
\]  

(Equation 3)

where \(Y\) was mating success of the focal fly (accepted or rejected), \(\mu\) is the intercept, \(SC\) is the CHC phenotype of the suitor flies, \(FC\) is the CHC phenotype of the focal fly, \(G_{SC}\) is the random effect of genotype nested within the suitor CHC phenotype, \(G_{FC}\) is the random effect of genotype nested within focal fly CHC phenotype and \(\varepsilon\) is random error. The interaction was tested via an F-test and within-sex effects via t-tests. Statistical significance was defined as a p-value < 0.05.

QTL mapping
QTL mapping was performed by analysing the CHC polymorphism as a binary trait. The trait 5,9-C25 was used to categorise flies as having either the *northern* or *common* phenotype; it is the largest CHC peak in *D. serrata* and is the easiest to score. For QTL analysis, we recoded the data, classifying flies as either *northern* (1) or *common* phenotype (0) based on their 5,9-C25 phenotype. 5,9-C25 had a clear bimodal distribution in the F2 cross with two distinct phenotypes. A 3:1 segregation ratio could not be rejected using Chi-Square analysis (\(\chi^2 = 2.181, p = 0.14\)), which is consistent with a single recessive segregating factor. Although the F2 progeny came from a
reciprocal cross that included both male and female offspring, we found no effect of either factor (cross or sex) on variability in the binary phenotype score. Thus, they were not fitted in subsequent QTL analyses. Standard interval mapping in R/qtl was then performed using the scanone() function and the model="binary" option.71 Genome-wide significance thresholds (95%) for the detected QTLs were determined through permutation (n=1000) tests.72

**Bulked segregant analysis**
Read counts for each variable site in the northern and common pools were analysed in R where we assessed genetic differentiation between the pools for each SNP using Fisher’s exact tests and a false discovery rate threshold of 5%.73

**Genome resequencing**
As the northern allele is recessive to the common type and we were working with inbred lines, we developed a diagnostic criterion for whether a SNP could be the causal variant. First, SNPs had to be fixed and identical across all northern lines. Second, the homozygous genotype of all northern lines must differ from any genotype of the common lines. Although in the cases of a tri-allelic SNP, it was possible for multiple common lines to possess different bases at a potentially causal variant as long as they differed from the northern lines, we applied the conservative test of homozygous fixed differences only. This rule was applied and we summed the number of candidate SNPs within 40kbp windows along the genome to look for an overrepresentation of positive hits using R.57

**Population genetic analysis**
Because Tajima’s D is affected by both selection and demography, we estimated Tajima’s D at all polymorphic exons within the genome using the PopGenome package in R and used this empirical distribution as a point of comparison. The expectation was that a locus under balancing selection would be in the 95th percentile of the empirical distribution, that is, 95% or more of the genome-wide Tajima’s D values would be less than the value for the locus. One-tailed significance (p-values) of Tajima’s D being a genome-wide outlier was therefore assessed by calculating that proportion of all exons that were greater than the value of interest.
Supplemental Information

Natural variation at a single gene generates sexual antagonism across fitness components in *Drosophila*

Bosco B. Rusuwa, Henry Chung, Scott L. Allen, Francesca D. Frentiu, and Stephen F. Chenoweth
Figure S1 Crossing design for QTL mapping experiments, Related to Figure 3. Crossing scheme for F2 inter-cross between inbred lines fixed for alternate CHC phenotypes. Schematic also indicates the highly advanced 60-generation mass bred population that was used for bulked segregant analysis.
Figure S2 Evolutionary relationships among reductase genes at the DsFAR locus, Related to Figure 4.

A) Gene copy number in the CG17562/CG17560/CG14893 locus vary from one to five in different Drosophila genomes. Figure adapted from Finet et al. S1.

B) Maximum likelihood tree of the reductases in the CG17562/CG17560/CG14893 suggest a that the duplication of FAR2 locus in D. serrata is a recent one.
Figure S3 Oenocyte specific RNAi of the DsFAR2 ortholog CG17562 in male *D. melanogaster*, Related to Figure 4 Showing both sides of the reciprocal cross and all integrated CHC peaks. The oenocyte GAL4 driver line was the mother in reciprocal cross 1 and the father in cross 2. Error bars are 95% confidence intervals. Asterisks indicate significant (p < 0.05) differences in mean CHC expression between the Control and RNAi treatment within each reciprocal cross (*t*-test).
Figure S4. Oenocyte specific RNAi of the DsFAR2 ortholog CG17562 in female *D. melanogaster*, Related to Figure 4. Showing both sides of the reciprocal cross and all integrated CHC peaks. The oenocyte GAL4 driver line was the mother in reciprocal cross 1 and the father in cross 2. Error bars are 95% confidence intervals. Asterisks indicate significant (p < 0.05) differences in mean CHC expression between the Control and RNAi treatment within each reciprocal cross (*t*-test).
<table>
<thead>
<tr>
<th>Trait</th>
<th>Effect</th>
<th>d.f.</th>
<th>F-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desiccation Resistance</td>
<td>Sex</td>
<td>1,10</td>
<td>357.22</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Phenotype (<em>northern/common</em>)</td>
<td>1,10</td>
<td>2.41</td>
<td>0.1513</td>
</tr>
<tr>
<td></td>
<td>Sex × Phenotype</td>
<td>1,10</td>
<td>5.29</td>
<td>0.0442</td>
</tr>
<tr>
<td>Heat Shock Stress Survival</td>
<td>Sex</td>
<td>1,10</td>
<td>24.69</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>Phenotype</td>
<td>1,10</td>
<td>0.53</td>
<td>0.4814</td>
</tr>
<tr>
<td></td>
<td>Sex × Phenotype</td>
<td>1,10</td>
<td>5.7</td>
<td>0.0381</td>
</tr>
<tr>
<td>Male Mating Success</td>
<td>Chooser Phenotype</td>
<td>1,10</td>
<td>3.12</td>
<td>0.1078</td>
</tr>
<tr>
<td></td>
<td>Suitor Phenotype</td>
<td>1,10</td>
<td>21.9</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>Chooser Phenotype × Suitor Phenotype</td>
<td>1,284</td>
<td>1.76</td>
<td>0.1860</td>
</tr>
<tr>
<td>Female Mating Success</td>
<td>Chooser Phenotype</td>
<td>1,10</td>
<td>0.0037</td>
<td>0.9525</td>
</tr>
<tr>
<td></td>
<td>Suitor Phenotype</td>
<td>1,10</td>
<td>0.83</td>
<td>0.3841</td>
</tr>
<tr>
<td></td>
<td>Chooser Phenotype × Suitor Phenotype</td>
<td>1,353</td>
<td>2.92</td>
<td>0.0885</td>
</tr>
</tbody>
</table>

Table S1 Statistical tests of fixed effects for desiccation resistance, heat shock survival and mating success, Related to Figure 2.

<table>
<thead>
<tr>
<th>Generation</th>
<th>CHC Polymorphism Phenotype</th>
<th>Sequence Type (insert size)</th>
<th>Read Length</th>
<th>Sequencing Provider</th>
<th>No. of Reads</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>F60(1)</td>
<td>Common</td>
<td>Paired-end (500bp)</td>
<td>100bp</td>
<td>BGI</td>
<td>66,631,462</td>
<td>31.4566</td>
</tr>
<tr>
<td>F60(1)</td>
<td>Northern</td>
<td>Paired-end (500bp)</td>
<td>100bp</td>
<td>BGI</td>
<td>62,418,518</td>
<td>29.2918</td>
</tr>
<tr>
<td>F60(2)</td>
<td>Common</td>
<td>Paired-end (500bp)</td>
<td>100bp</td>
<td>BGI</td>
<td>59,914,718</td>
<td>27.5689</td>
</tr>
<tr>
<td>F60(2)</td>
<td>Northern</td>
<td>Paired-end (500bp)</td>
<td>100bp</td>
<td>BGI</td>
<td>60,005,830</td>
<td>27.6818</td>
</tr>
<tr>
<td>CTN42</td>
<td>Northern</td>
<td>Paired-end (500bp)</td>
<td>100bp</td>
<td>AGRF</td>
<td>116,022,310</td>
<td>57.7446</td>
</tr>
</tbody>
</table>

Table S2 Sequencing libraries for pooled DNA samples of a CTN42-FORS4 D. serrata F60 intercross, related to Figure 3. Data used for bulked segregant analysis of CHC polymorphism.
Supplemental References