www.nature.com/hdy

ORIGINAL ARTICLE

Identification of quantitative trait loci function through analysis of multiple cuticular hydrocarbons differing between *Drosophila simulans* and *Drosophila sechellia* females

JM Gleason[1], RA James[2], C Wicker-Thomas[3] and MG Ritchie[2]

[1]Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, USA; [2]School of Biology, University of St Andrews, St Andrews, Fife, Scotland and [3]LEGS, CNRS, Université Paris-XI, Gif sur Yvette Cedex, France

The genetics of sexual isolation, behavioral differences between species that prevent mating, is understood poorly. Pheromonal differences between species can influence sexual isolation in many animals and in some cases a single locus can cause large functional changes in pheromonal mating signals. Drosophila cuticular hydrocarbons (CHCs) can function as pheromones and consequently affect mate recognition. In a previous study of the two major CHCs in females that affect mating discrimination between *Drosophila simulans* and *D. sechellia*, quantitative trait loci (QTL) were identified on the X and third chromosome, and a few candidate genes were potentially implicated. Here we specifically test candidate genes for CHC biosynthesis and determine the genetic architecture of four additional CHCs that differ in abundance between *D. simulans* and *D. sechellia*

females. The same QTL, and new ones, were found for additional CHCs. By examining all these CHCs and exploring their covariance, we were able to ascribe putative function to the major QTL. Although desaturases have received considerable attention for their role in CHC biosynthesis, evidence here implies that elongases may be just as important. Sex determination genes do not seem to have a role in this species difference although *D. sechellia* is sexually dimorphic in CHCs, whereas *D. simulans* is not. Epistatic interactions, only detected for CHCs limited to *D. sechellia*, imply that complex interactions among loci may also be having a role in these compounds that affect mating isolation.
Heredity (2009) **103**, 416–424; doi:10.1038/hdy.2009.79; published online 5 August 2009

Introduction

Identifying genes contributing to adaptive divergence and reproductive isolation between species is a crucial step for understanding the process of speciation. Reproductive isolation can be caused by mechanisms acting before or after mating. One contributing factor to premating isolation is behavioral or sexual isolation, but unlike postmating isolation, few behavioral genes influencing speciation have been identified (Noor, 2003; Orr et al., 2004; Noor and Feder, 2007). Mating signals in multiple modalities may influence sexual isolation. Chemosensory reception, which includes olfaction and gustation, may have a large role so that differences in pheromones functioning as mating signals can influence sexual isolation (Smadja and Butlin, 2009).

Evidence exists for the effects of both single and multiple genes on pheromonal differentiation. A change in the expression of a single desaturase in the Asian Corn Borer results in a novel pheromonal blend compared with the European Corn Borer (Roelofs et al., 2002), and there is evidence for a single gene change in sibling species of *Helicoverpa* moths resulting in a pheromonal

blend change (Wang et al., 2008). Although in both of these cases single genes are implicated, the allelic difference between species results in a change in the blend of compounds, not just a single pheromone. As there are multiple compounds requiring different biosynthetic pathways, there is the potential for the interaction of multiple genes through both additive and epistatic effects: some enzymes may be required for the synthesis of precursor molecules, whereas other enzymes may modify precursors to final products. In Drosophila, multiple gene effects and epistatic interactions have been found for interspecific differences between *D. simulans* and *D. sechellia*, as well as between *D. pseudoobscura* and *D. persimilis* (Coyne, 1996; Noor and Coyne, 1996; Gleason et al., 2005).

Drosophila cuticular hydrocarbons (CHCs) can function as pheromones and affect mate recognition (Cobb and Jallon, 1990; Coyne et al., 1994; Etges and Ahrens, 2001; Ferveur, 2005). CHCs are derived from long-chain fatty acids and influence desiccation (Toolson, 1982; Lockey, 1988; Rouault et al., 2004), cold-tolerance (Ohtsu et al., 1998) and starvation resistance (Hoffmann et al., 2001). Species within the *D. melanogaster* subgroup are either sexually monomorphic or dimorphic for CHCs. Monomorphic species (such as *D. simulans* and *D. mauritiana*) have high levels of monoenes, especially 7-tricosene (7-T, which has 23 carbons and one double bond, C23:1). Males of sexually dimorphic species (*D. melanogaster* and *D. sechellia*) have monoenes, but

Correspondence: Dr JM Gleason, Department of Ecology and Evolutionary Biology, University of Kansas, 1200 Sunnyside Avenue, Haworth Hall 6006, Lawrence, KS 66045-7534, USA.
E-mail: jgleason@ku.edu

the females have dienes, primarily 7,11-heptacosadiene (7,11-HD, C27:2, Cobb and Jallon, 1990). Males of monomorphic species rarely court females of dimorphic species (Cobb and Jallon, 1990). Although it has not been demonstrated definitively that 7-T and 7,11-HD are the active compounds involved in sexual isolation between *D. simulans* and *D. sechellia*, 7,11-HD has been shown to be the major stimulant for *D. melanogaster* courtship (Antony *et al.*, 1985).

7-T and 7,11-HD are not the only CHCs differing between these two species, as other CHCs are present that differ in their number of carbons and double bonds (Table 1). The overall pattern is that CHCs in *D. sechellia* have longer chains and are less saturated than those found in *D. simulans*. In a previous study we used a quantitative trait locus (QTL) approach to examine differences in the amount of 7-T and 7,11-HD in females of *D. simulans* and *D. sechellia* (Gleason *et al.*, 2005). Through composite interval mapping (CIM), we found at least four QTLs for the monoene 7-T, three on chromosome 3 and one on the X, and two QTLs for the diene 7,11-HD on chromosome 3. Although genes have been identified that may influence CHC biosynthesis (candidate genes), very few were included as markers in that study, though the approximate locations of these genes were determined (Table 6 in Gleason *et al.*, 2005). Of the possible candidate genes, only those on chromosome 3 were close to QTL. Gleason *et al.* (2005) did not study the other CHCs that differ between these two species.

All Drosophila CHCs are synthesized *de novo* from acetate- and linear-saturated fatty acids by elongation, decarboxylation and desaturase pathways (Figure 1, Chan Yong and Jallon, 1986; Pennanec'h *et al.*, 1997). Analysis of multiple CHCs may give an insight into the genetic basis of the two major compounds influencing mate recognition in these two species through establishing the relative contributions of various enzymes associated with CHC biosynthesis (Legendre *et al.*, 2008). Very little is known about the genes encoding carboxylases, but the elongases and desaturases have received more attention. Recent work has shown that *elongase F* (*eloF*), expressed in female *D. melanogaster* and *D. sechellia*, but not *D. simulans*, is involved in the elongation of long hydrocarbons, especially diene pheromones (Chertemps *et al.*, 2007). RNA interference knockdown of *eloF* in oenocytes is sufficient to reduce the expression of long dienes (Chertemps *et al.*, 2007).

The role of other elongases in Drosophila CHC expression has not been explored.

Much more is known about the desaturases, which comprises a gene family with effects on pheromones (Roelofs and Rooney, 2003). Desaturases add double bonds to existing compounds and, hence, may influence the relative amounts of dienes, monoenes and saturated compounds. There are at least 10 fatty acid desaturases in *D. melanogaster*, seven of which are on chromosome 3. Three of these third chromosome desaturases, *desat1*, *desat2* and *desatF*, have been shown to be involved in pheromone biosynthesis in *D. melanogaster* (Figure 1 for example, Dallerac *et al.*, 2000; Legendre *et al.*, 2008).

In *D. melanogaster*, induced mutations in *desat1* have reduced levels of all unsaturated CHCs (including 7,11-HD and 7-T) but increased levels of saturated hydrocarbons (Labeur *et al.*, 2002); one mutation also influences males perception of pheromones (Marcillac *et al.*, 2005). Genetically adjacent to *desat1* on chromosome 3, *desat2* has been implicated in geographical differentiation in CHCs (7,11-HD vs 5,9-HD) in *D. melanogaster* (Coyne *et al.*, 1999; Dallerac *et al.*, 2000; Takahashi *et al.*, 2001), though molecular transformation
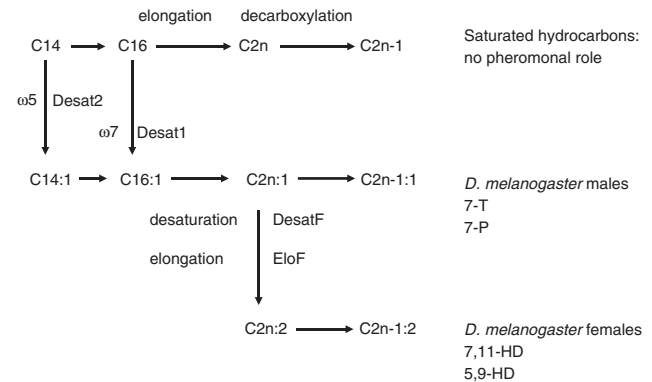

Figure 1 Hydrocarbon biosynthesis in *Drosophila melanogaster*. The first ω7 desaturation, carried out by Desat1, is common to males and females. The ω5 desaturation, carried out by Desat2, is only expressed in *D. melanogaster* African females (those that have 5,9-heptacosadiene (HD)). *D. simulans* synthesis for both males and females follows the *D. melanogaster* male pathway. *D. sechellia* hydrocarbon biosynthesis is presumably similar to that of *D. melanogaster*, though this has not been shown. Adapted from Legendre *et al.*, (2008).

Table 1 Cuticular hydrocarbons studied and parameters from composite interval mapping

Abbreviation	Compound	Number of carbons	Number of double bonds	Species[a]	Quantitative trait loci (QTL) significance level[b]	Number of background markers[c]	Number of individuals in QTL mapping
T	Tricosane	23	0	D simulans	13.357	8	484
7-T	7-Tricosene	23	1	D. simulans	13.572	8	484
7-P	7-Pentacosene	25	1	D. simulans	13.981	5	484
7-H	7-Heptacosene	27	1	D. sechellia	14.387	5	484
7,11-PD	7,11-Pentacosadiene	25	2	D. sechellia	13.781	8	483
7,11-HD	7,11-Heptacosadiene	27	2	D. sechellia	14.492	7	482

[a]Species in which the compound is more abundant.
[b]Significance level (likelihood ratio) at $P = 0.05$ for each compound is derived from 1000 permutations of the dataset and composite interval mapping (Figure 3).
[c]The number of significant background markers used in composite interval mapping. All the significant markers derived from interval mapping were used.

experiments are in disagreement regarding the contribution of *desat2* to ecological adaptation in *D. melanogaster* (Greenberg *et al.*, 2003, 2006; Ritchie and Noor, 2004; Coyne and Elwyn, 2006).

Also on chromosome 3 is *desatF* (syn. *Fad2*). The RNA interference knockdown of this gene reduces diene production to <5% of normal and increases the level of monoenes (Chertemps *et al.*, 2006). Males with 'feminized' hydrocarbons through the targeted expression of *transformer*, express *desatF* and are courted by wild-type males. *desatF* is expressed in *D. sechellia* females (but not *D. simulans*), and hybrids between *D. melanogaster* and *D. simulans* have dienes only if the *desatF* allele from *D. melanogaster* is present (Legendre *et al.*, 2008). *desatF* is thus a strong candidate gene for increased levels of dienes.

In addition to genes directly in biosynthesis pathways, genes influencing their regulation may be potential candidate genes. As CHCs are sexually dimorphic in *D. sechellia*, genes in the sex determination pathway are potential candidates. These include *Sexlethal* (*Sxl*), *transformer* (*tra*), *intersex* (*ix*), *doublesex* (*dsx*) and *fruitless* (*fru*). In our previous work (Gleason *et al.*, 2005), *Sxl*, *tra* and *ix* were not near QTL. The remaining two genes are both on chromosome 3. *dsx*, which influences the production of female dienes in *D. melanogaster*, (Jallon *et al.*, 1988) was on the edge of a QTL (Gleason *et al.*, 2005). The location of *fru* was not determined in that study.

In this study, we expand our previous investigation (Gleason *et al.*, 2005) using the same individuals to include more CHCs in addition to 7–T and 7,11-HD, and refine the mapping with the addition of marker loci in candidate genes. By examining six CHCs with varying saturation and carbon chain length, we examine the covariance of CHC production to help identify common genetic pathways among CHCs. In addition, by scoring markers in the potential candidate genes *desat1*, *desatF*, *doublesex* (*dsx*) and *fruitless* (*fru*), we are able to further assess if these contribute to QTL.

Materials and methods

Mapping population

This study used the same individuals as Gleason *et al.* (2005), with the introduction of new markers and additional CHCs in the analysis. Strain maintenance, crosses, molecular markers, cuticular hydrocarbon extraction and genetic map construction are described in Gleason *et al.* (2005). For the mapping population, female *D. simulans* were crossed to male *D. sechellia* and backcrossed to *D. simulans* to obtain backcross individual females for analysis.

Cuticular hydrocarbons were isolated by washing 5-day-old females in hexane containing 800 ng hexacosane, as a control for extraction efficiency. The quantities of the CHCs (listed in Table 1) were calculated through gas chromatography (described in Gleason *et al.*, 2005) as the area under each peak on the chromatograms. By dividing the CHC peak areas by the quantity of hexacosane recovered, we were able to account for differences in extraction and gas chromatography analysis. Eight extreme data outliers were removed and one was added before the resulting values were natural log transformed to remove a right skew. Remaining residuals

were normally distributed. All subsequent analyses were done with the transformed variables. In addition to the backcross females, *D. simulans*, *D. sechellia* and F1 hybrid females were also measured. Final sample sizes are given in Table 1.

After hexane washes, DNA was isolated from the backcross females in a 100 µl single fly DNA prep (Gloor and Engels, 1992). For each individual, five morphological markers and 37 molecular markers were scored (described in Gleason *et al.*, 2005). For this study, four new markers in the candidate genes *desat1*, *desatF*, *dsx* and *fru*, all located on chromosome 3, were also scored using the same procedures (Supplemental Table 1) on the same DNA that had been stored at −20 °C. A fifth candidate gene, *desat2* is not presented because the gene is only 7 kb from desat1 in *D. simulans* and *D. sechellia* (Tweedie *et al.*, 2009), and is not resolvable as different from *desat1* in this analysis. Markers were scored by PCR amplification and visualization with gel electrophoresis. For markers that did not differ in size between *D. simulans* and *D. sechellia*, restriction digests were used to distinguish the two alleles (Supplemental Table 1).

Genetic mapping and QTL analysis

A genetic linkage map was assembled from the 46 markers using Antmap with the Kosambi map function (Iwata and Ninomiya, 2006). The map was subsequently used in QTL analyses using QTL Cartographer v 1.17j (Basten *et al.*, 2005). A marker on the fourth chromosome (*eyeless*) was not significantly associated with any trait and thus results for this chromosome, which comprises about 1% of the genome, are not shown. Composite interval mapping was carried out for the rest of the genome using all significant background markers and a window size of 10 cm. Significance levels at $P = 0.05$ (Table 1) for all CHC were calculated from 1000 permutations of the trait data among marker classes (Churchill and Doerge, 1994). Boundaries of QTL were determined by finding the highest peak and then extending the QTL for 2-LOD (logarithm of odds) around the peak, which is a conservative estimate of the location (van Ooijen, 1992; Manichaikul *et al.*, 2006). When peaks were adjacent, if there was not a 2-LOD drop between them, the QTL region was combined. Nearest markers to the peaks were determined from the map.

Statistical tests

Differences in quantities for each CHC were compared among pure species, hybrids and backcross hybrids using analysis of variance with a Tukey *post-hoc* test with a 5% family error rate. Epistatic interactions between QTL were detected using generalized linear analysis of variance models with loci as covariates and all first order epistatic interactions. Genotype state (homozygous *D. simulans* or hybrid) was taken from the marker nearest to the likelihood peak of each QTL. As we examined six traits, an appropriate Bonferroni corrected significance level is 0.008.

Results

QTL analysis

Six of the most abundant CHCs were measured (Table 1). These varied greatly in abundance both within and between *D. simulans* and *D. sechellia* females (Figure 2).

These differences were not likely to be an effect of a difference in body size between the two species because most of these CHCs were present in one species and absent in the other. For most CHCs, the hybrid mean was intermediate. 7,11-HD showed a dominance effect of *D. simulans* alleles and 7-T of *D. sechellia* alleles, and there was transgressive segregation (in the form of over-dominance) for 7-H and 7,11-PD. The backcross mean was more similar to *D. simulans* than to *D. sechellia* and not statistically significantly different from *D. simulans* for 7-T, 7-pentacosene (P), 7-heptacosene (H) and 7,11-HD, though 7-P showed transgressive segregation. The backcross mean was more like *D. simulans* for tricosane, although the backcrosses had significantly less of the compound. For 7,11-PD, the backcross mean was similar to the hybrid mean and *D. sechellia*.

QTL analysis showed 2–5 QTL for each CHC (Figure 3 and Table 2). Effects for tricosane and 7-T were all in the expected positive direction relative to *D. sechellia* and for 7,11-PD and 7,11-HD were all negative. For the other two CHCs, the directions of the effects were mixed (Table 2). The total phenotypic variance explained by all QTL for individual CHCs ranged from 27.95–85.93%.

The ranges of the 2-LOD drop intervals for the QTL were either unique to a CHC (three for 7,11-PD and one for T) or shared among 4–6 CHCs (Figure 3). The unique QTL included the only QTL found on the second chromosome, whereas all the others were on the third chromosome. The QTL regions shared by multiple CHCs were in three distinct areas. The first was on the X chromosome, closest to the *forked* (*f*) gene and was shared among the most saturated compounds. No known candidate genes were in this region. The second region was on the left arm of the third chromosome and was shared by both dienes (7,11-PD and 7,11-HD), as well as 7-T and tricosane. The two dienes were very close to *desatF*, suggesting that *desatF* was a relevant candidate

Figure 2 Amounts of cuticular hydrocarbons (CHCs) per individual. Nanogram quantities of each CHC are given with standard errors for *Drosophila simulans* and *D. sechellia* females, hybrids of the cross between *D. simulans* females with *D. sechellia* males and backcrosses of hybrid females to *D. simulans*. All values were corrected for the amount of standard recovered for each individual. The sample sizes are given below each bar. Different letters above each bar indicate statistically significant differences between pairs of female types by Wilcoxon's signed-ranks tests with Bonferroni correction for multiple tests.

gene for the dienes. No other candidate genes were in this region.

The final shared QTL was on 3R and influenced all of the CHC. The bounding markers for this region are *prospero* (*pros*) and *Metallothionein A* (*MtnA*). The two candidate genes *desat1* and *desat2* were adjacent to this interval and were within the 2-LOD drop QTL interval for tricosane and 7,11-PD, but were not within the QTL for the two CHC which were the most abundant CHCs in *D. simulans* and *D. sechellia*, 7-T or 7,11-H*D*.

Epistasis

Pairwise analyses of QTL for each compound was carried out to identify epistatic interactions (Table 3). QTL for only three compounds had epistatic interactions; all of these compounds were more abundant in *D. sechellia* than *D. simulans* (Table 1). For 7,11-PD and 7,11-HD, all epistatic interactions were between a QTL on the left arm of chromosome 3 and one either on the right arm of chromosome 3 or on chromosome 2. 7,11-HD is present in large amounts only when both the QTLs on chromosome 3 are heterozygous (Figure 4).

In the case of 7-H, which lacks a QTL on the left arm of chromosome 3, the epistatic interaction detected was between the QTL on the X chromosome and one on the right arm of chromosome 3. Without Bonferroni correction, these two QTL would have a significant ($P = 0.013$) interaction for 7-P as well (data not shown).

Covariance of traits

Quantities of CHC compounds were related as the amounts present showed high covariance and strong positive or negative correlations (Table 4). The potential for covariance to reflect common QTL is an important and rarely tested issue (Gardner and Latta, 2007), and here the greatest covariance (Table 4) was for the traits sharing the most QTLs (3 peaks in common between 7-T and tricosane; Figure 3). 7-P and 7-H were also strongly correlated (and share two QTL peaks; Figure 3). A compound found only in *D. sechellia*, 7,11-HD, was strongly correlated with all the other CHCs, and was negatively correlated with those CHCs that are most abundant in *D. simulans*.

Analysis of covariance between pairs of traits showed a unique, non-linear relationship for 7-T and 7-P, which formed two distinct clusters with different slopes (Figure 5). The majority of the individuals comprising the cluster with the steeper of the two slopes (that is, a higher ratio of 7-P to 7-T than the other cluster) were heterozygous for *pros* (Figure 5) and *MtnA* (data not shown). The majority of the individuals comprising the second cluster were homozygous for these two loci. The mean value of this ratio for *D. simulans* falls within the cluster for the homozygotes (Figure 5). The difference between homozygotes and heterozygotes was consistent with an elongase function in the QTL region near *pros* and *MtnA*, as 7-P is longer than 7-T.

To examine this relationship further, the ratio of 7-P to 7-T was used in composite interval mapping for QTL (Supplementary Figure 1). The ratio of 7-P to 7-T was used to compare a difference in chain length (25 vs. 27 carbons), but not saturation (one double bond each), between CHCs. In all, 3 QTL were identified that account for nearly 94% of the phenotypic variance (Table 2).
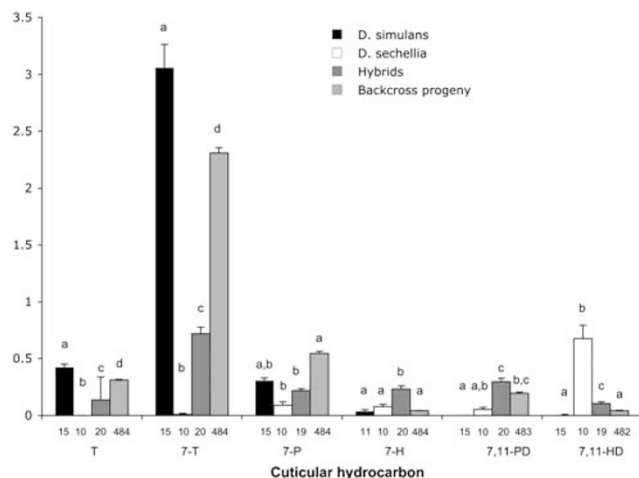
Table 2 Locations and effects of QTLs for each cuticular hydrocarbon and ratios

Compound[a]	Chrom.	Position[b]	Markers in region[c]	Candidate gene	Effect	%V_p[d]
T	X	1.3490–1.4763 (1.3749)	*nonA-f-Zw*		0.1018	21.23
	3	0.7644–1.1164 (1.0164)	*Sod–Est6-**Acp70A**-st*		0.0336	2.16
	3	1.6532–1.8440 (1.7438)	*Gld-dsx-e-nos*	*dsx*, CG2781	0.0562	4.90
	3	2.3692–2.6341 (2.4492)	*gl-**desat1**-pros—MtnA*	*desat1, desat2*, CG8630	0.0773	9.95
				Total		38.24
7-T	X	1.3149–1.3949 (1.3549)	*nonA-**f***		0.2532	16.99
	3	0.7644–1.1164 (1.0164)	*Sod–Est6-**Acp70A**-st*		0.1099	2.98
	3	2.5341–2.6341 (2.5941)	*pros–**MtnA***	*eloF*	0.2146	9.68
				Total		29.65
7-P	X	1.2949–1.3949 (1.3549)	*nonA-**f***		0.1718	11.73
	3	2.5741–2.6141 (2.5941)	*pros–**MtnA***	*eloF*	−0.3379	45.11
	3	2.7271–2.7671 (2.7471)	***MtnA**–slo*	CG31141	−0.3847	29.09
				Total		85.93
7-H	X	1.2549–1.4963 (1.3749)	*nonA-f-Zw*		0.0247	5.27
	3	2.5541–2.6341 (2.5941)	*pros–**MtnA***	*eloF*	−0.0512	22.68
				Total		27.95
7, 11-PD	2	0.2572–0.4558 (0.3358)	*nt-**odd**-Mst26Aa*		−0.0582	2.24
	3	0.1412–0.4698 (0.3498)	*cdc37-**tipE**-LanB2*		−0.0680	2.38
	3	0.6573–0.6973 (0.6573)	*LanB2-**desatF***	*desatF, Elo68α, Elo68β*	−0.2296	11.25
	3	0.8044–0.8244 (0.8244)	*Sod–**Est6***		−0.2570	27.46
	3	2.3492–2.6341 (2.5741)	*gl-desat1-**pros**–MtnA*	*desat1, desat2, eloF*	0.0710	2.89
				Total		73.47
7,11-HD	3	0.7644–0.8644 (0.8044)	***Sod**–Est6*		−0.0644	16.73
	3	2.6141–2.7471 (2.6671)	*pros—**MtnA**-slo*	*eloF*, CG31141	−0.0634	18.99
				Total		35.72
7-P/7-T	X	1.2549–1.4563 (1.3549)	*nonA-f-Zw*		0.0849	3.12
	3	2.5741–2.6141 (2.5941)	*pros–**MtnA***	*eloF*	−0.3531	53.76
	3	2.7271–2.7671 (2.7471)	***MtnA**-slo*	CG31141	−0.3920	36.75
				Total		93.63
7,11-PD/7-P	X	1.0415–1.4149 (1.3149)	*cac-**nonA**-f*		−0.3628	2.85
	3	0.6573–0.8444 (0.7102)	*LanB2-**desatF**–Sod–Est6*	*desatF, Elo68α, Elo68β*	−0.7588	11.73
	3	1.3158–1.7238 (1.4758)	*Cat–**Gld**–dsx-e*	*dsx*, CG2781	−0.4764	4.19
	3	2.5341–2.6341 (2.5941)	*pros—**MtnA***	*eloF*	0.6875	9.39
				Total		28.16

Abbreviations: H, heptacosene; HD, heptacosadiene; MtnA, metallothionein A; P, pentacosene; PD, pentacosadiene and QTL, quantitative trait loci.
[a]See Table 1 for compound abbreviations.
[b]Postions are given in Morgans and correspond to Figure 3. The position in parentheses is the point of the highest peak. The range is the 2-LOD change from the peak.
[c]Flanking markers are given for the 2-LOD range. The marker in bold is the closest marker to the peak.
[d]Percentage of the phenotypic variance explained.

These QTL were the same as the intervals found for multiple QTL in the single trait mapping, one on the X chromosome and two to either side of *MtnA*. No epistatic interactions were found for the ratio of 7-P to 7-T.

To study the difference in saturation (two vs. one double bond), but not chain length (both 25 carbons), the ratio of 7,11-PD to 7-P was used in QTL mapping (Supplementary Figure 1). In this case 4 QTL were found that explain only approximately 28% of the phenotypic variance. Again, the X-chromosome QTL was present, as was one near *MtnA*. In addition, the largest peak was near *desatF*. A fourth peak is unique, located between *Catalase* (*Cat*) and *Glucose dehydrogenase* (*Gld*). The QTL nearest *desatF* had strong epistatic interactions with both the QTL on the X chromosome and the major QTL on the right arm of chromosome 3 (Table 3).

Discussion

Drosophila CHCs have important roles in behavior and ecological adaptation (Lockey, 1988; Cobb and Jallon, 1990; Coyne *et al.*, 1994; Ohtsu *et al.*, 1998; Hoffmann *et al.*, 2001; Fang *et al.*, 2002; Rouault *et al.*, 2004). CHCs result from the action of enzymes in biochemical path-

ways that start from a common substrate and, therefore, may be phenotypes that are likely to have shared QTL, covariance and epistasis because the underlying enzymes influence multiple traits or relationships between traits. By examining six CHC compounds differing between *D. simulans* and *D. sechellia* females, we have found evidence for the following three effects: common QTL regions, covariance and epistasis.

QTL, candidate genes and QTL function
In our previous study (Gleason *et al.*, 2005), we identified QTL for the two CHC, 7-T and 7,11-HD, that have a large influence on sexual isolation between *D. simulans* and *D. sechellia* (Cobb and Jallon, 1990; Coyne, 1996). By including additional CHC, here we were able to assign functions to QTL through inferences about shared structure (chain length and saturation). Associations with candidate genes reinforce the hypothesized functions. In addition to the four new markers in candidate genes, we identified other candidate genes through the literature and a search of Flybase (Tweedie *et al.*, 2009; Wicker-Thomas and Chertemps, 2010) for genes with the molecular processes related to CHC production (for
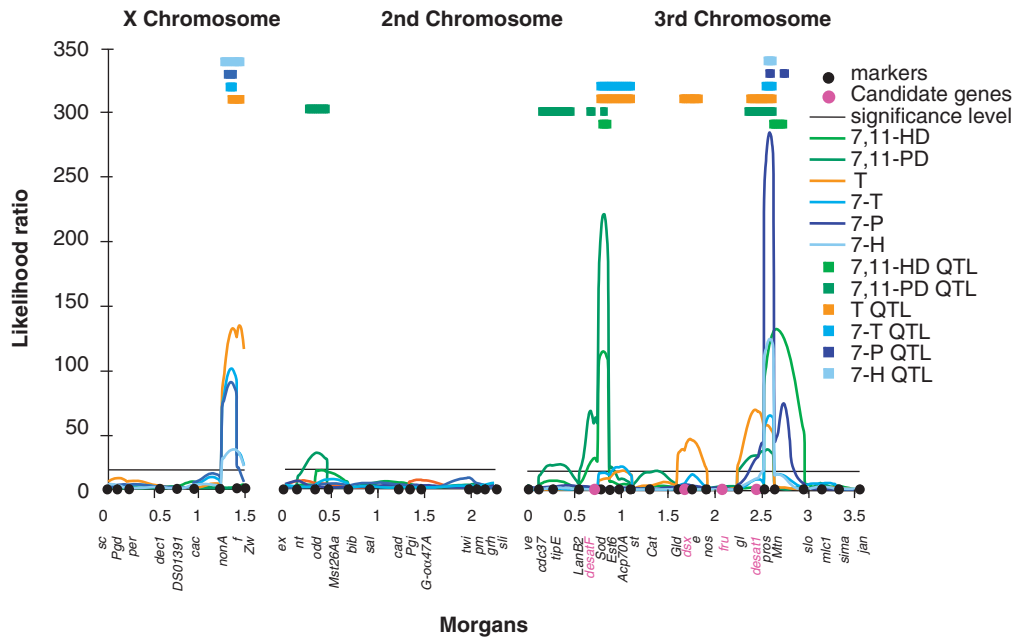
Figure 3 Quantitative trait loci (QTL) mapping of all six cuticular hydrocarbons (CHCs). Composite interval mapping (CIM) was carried out with a walking speed of 2 cM, a window size of 10 cM and forward/backwards regression on all significant background markers (see Table 1 for the number of significant markers for each CHC). The locations of the marker loci are indicated on the *x* axis in the conventional left to right order for Drosophila. The experiment-wide significance level of $P < 0.05$ obtained by 1000 permutations of the trait data among marker classes is shown with a horizontal line. Only the highest significance level, for 7,11-heptacosadiene (HD) is shown. Values for the significance levels are given in Table 1. Candidate loci, which are new markers in this study, are in pink. Across the top of the figure are the boundaries of the 2-LOD intervals for each QTL.

Table 3 Tests for Epistasis

Compound[a]	QTL pair[b]	P-value[c]
7-H	*f-MtnA*	**<0.001**
7,11-PD	*odd-tipE*	**0.005**
	odd-desatF	**<0.001**
	odd-Est6	0.598
	odd-pros	0.022
	tipE-desatF	0.346
	tipE-Est6	0.113
	tipE-pros	0.013
	desatF-Est6	0.993
	desatF-pros	**<0.001**
	Est6-pros	**<0.001**
7,11-HD	*Sod-MtnA*	**<0.001**
7,11-PD/7-P	*nonA-desatF*	**<0.001**
	nonA-Gld	0.410
	nonA-MtnA	0.044
	desatF-Gld	0.060
	desatF-MtnA	**<0.001**
	Gld-MtnA	0.035

Abbreviations: H, heptacosene; HD, heptacosadiene; MtnA, metallothionein A; P, pentacosene; PD, pentacosadiene and QTL, quantitative trait loci.
[a]See Table 1 for compound abbreviations. Only compounds showing epistatic interactions are presented.
[b]The nearest marker to the likelihood peak of each QTL was used in the analysis (see text).
[c]Epistatic interactions between QTL were detected using generalized linear models with loci as covariates and all first order interactions for allelic data at the nearest marker to the QTL. The appropriate Bonferroni significance level is 0.008.

example, stearoyl-CoA desaturase activity and fatty-acid elongase activity). There are 10 desaturases and 20 elongases, many of which occur in tandem pairs or in clusters (Supplementary Table 2).
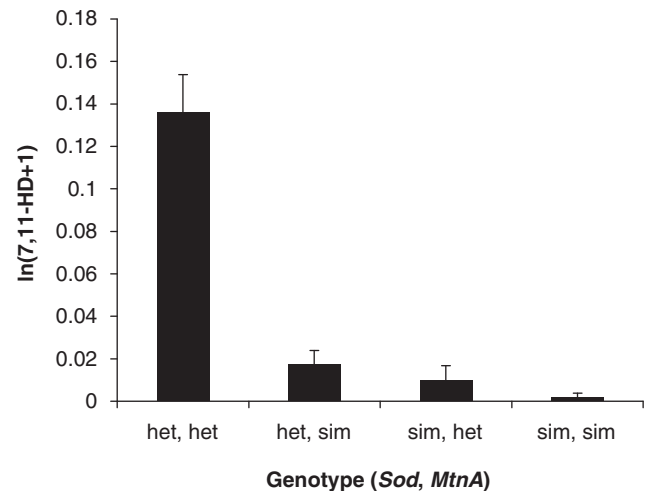

Figure 4 Epistatic interaction for 7,11-heptacosadiene (HD). The amount of the cuticular hydrocarbon (CHC) is given for each allelic state (heterozygote or homozygote *Drosophila simulans*) at the *Sod* and *Metallothionein A* (*MtnA*) loci in the recombinant backcross females. Error bars are 95% confidence intervals.

Three chromosome arms had major QTL that overlapped for multiple CHCs: 3L, 3R and the X chromosome. The total proportion of the phenotypic variance explained by QTL was high for the two 25 carbon compounds, both of which show transgressive segregation in the F1 or F2 backcross populations. A major inference from these shared QTL is that the same pathways are involved in the production of multiple CHCs and that phenotypic covariance is reflected in these shared QTL.

Table 4 Covariances and correlation coefficients between quantities of CHC compounds [a]

	T	7-T	7-P	7-H	7,11-PD	7,11-HD
T		**0.844**[b]	0.088	−0.075	0.041	**−0.198**
7-T	0.0121		0.110	−0.043	0.085	**−0.192**
7-P	0.0024	0.0084		**0.737**	−0.150	**0.401**
7-H	0.0004	−0.0007	0.0099		0.023	**0.362**
7,11-PD	0.0009	0.0050	−0.0072	−0.0011		**0.359**
7,11-HD	−0.0016	−0.0043	0.0072	0.0014	0.0050	

Abbreviations: CHC, cuticular hydrocarbon; H, heptacosene; HD, heptacosadiene; MtnA, metallothionein A; P, pentacosene and PD, pentacosadiene.
[a]Covariances are in the lower diagonal and correlation coefficients are in the upper diagonal.
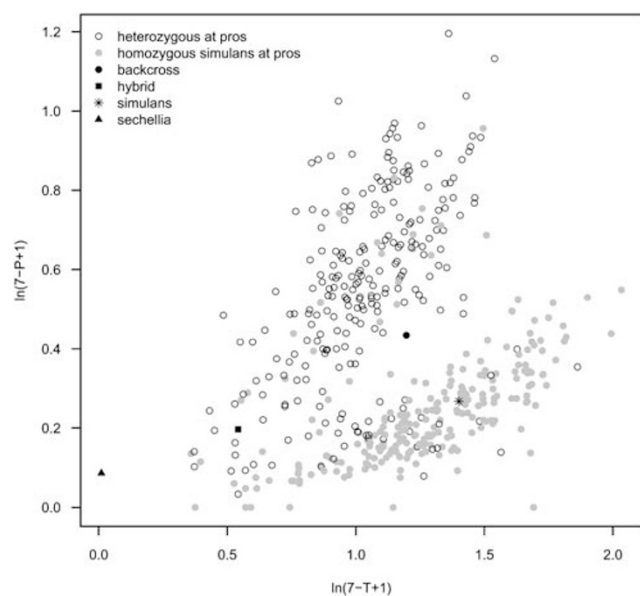[b]Correlation coefficients in bold have $P < 0.008$

Figure 5 The relationship between 7-pentacosene (P) and 7-tricosene (T). Natural log transformed quantities of 7-P vs. 7-T is plotted for all the backcross individuals, as well as the mean values for *Drosophila simulans*, *D. sechellia*, their hybrids and the back-crosses.

Starting with 3L, QTL for the two dienes were located between *Sod* and *Est6* (Figure 3, Table 2). These QTL overlapped with broad QTL located near *Acp790A* for tricosane and 7-T. The proportion of the phenotypic variance (Vp) explained by the QTL for the dienes greatly exceeded that of the other two compounds (Table 2). The best candidate gene for this region was the desaturase *desatF*, given that saturation differences mapped most strongly here (Supplementary Figure 1), as shown by the ratio of 7,11-PD to 7-P, two compounds that differ in saturation but not chain length. However, two elongases are adjacent to *desatF* (Table 2, Supplementary Table 2). As neither of these elongases is expressed in *D. melanogaster* females (Chertemps *et al.*, 2005), neither are likely to be involved in female CHC biosynthesis and are probably not the causative gene(s) here if expression patterns have not changed for these species, though this remains to be confirmed for *D. simulans* and *D. sechellia*.

A QTL on 3R was shared by all of the CHCs implying that the QTL covers one or more fundamental genes involved in the production of CHCs. The most well-studied candidate genes among our markers were the pair *desat1* and *desat2*; however, these fell outside of the QTL for most of the CHCs. The 3R QTL explained a large proportion of Vp for the ratio of 7-P to 7-T, two compounds differing in chain length and not saturation implying that there was an elongase function for this QTL. Two clusters of genes, all elongases, were located in this region. The first cluster containing five genes fell between *pros* and *MtnA*, and included *eloF*, which has female-biased expression in *D. melanogaster*. Furthermore, *eloF* affects the relative production of dienes of different chain lengths and is not expressed in *D. simulans* (Chertemps *et al.*, 2007), thus is an excellent candidate gene. The second cluster of four genes fell between *MtnA* and *slo*; one or more of these elongases is likely to be important in the difference between *D. simulans* and *D. sechellia*. Most of the literature discussing genetic control of pheromonal differences among species has focused on the role of desaturases (e.g., Coyne *et al.*, 1999; Rosenfield *et al.*, 2001; Fang *et al.*, 2002; Roelofs *et al.*, 2002; Greenberg *et al.*, 2003; Roelofs and Rooney, 2003; Coyne and Elwyn, 2006; Legendre *et al.*, 2008). Although a desaturase has been implicated for 3L, there is at least an equal role in CHC for elongases in the difference between *D. simulans* and *D. sechellia*.

The other major QTL region identified affected tricosane and all the monoenes (on the X chromosome, Figure 3) implying a common pathway for these compounds. This QTL was not associated with any known candidate genes. Examination of the region between *nonA* and *f* in Flybase (Tweedie *et al.*, 2009) identified 148 genes, the majority of which do not have a known function. Of the genes with inferred or tested functions, none have functions that are known to be involved in CHC production. Thus, as yet undescribed loci are implicated here.

The sex determination genes used here as candidate genes are not implicated in the shared pathways of these CHCs, as these genes are not associated with the shared QTL. *dsx* was within the bounds of the unique tricosane QTL (Figure 3) that overlapped with a QTL for the ratio 7,11-PD to 7-P (Supplementary Figure 1), although neither of these compounds have a QTL in this location when mapped individually. Our search for additional candidate genes on Flybase identified a fatty-acid elongase, CG2781, tightly linked to *dsx* (Supplementary Table 1) and thus may be the relevant gene. In the original report of a *fru* allele affecting male CHC production (Gailey and Hall, 1989), the *fru* allele was actually an inversion and the CHC effect mapped not to what was later shown to be *fru* locus, but to a breakpoint that in our map (Figure 3) is located just to the right of *gl*. This area is implicated in QTL for some compounds and thus *fru* is not the candidate gene for CHCs.

Sex determination pathway genes are not likely to have a role in the species differences here because most mutations in sex determination genes that affect pheromone production change the pheromone from that of one sex to that of another (for example, Savarit and Ferveur, 2002; Tompkins and McRobert, 1995). In the evolution of these two species, it is not clear if differentiation of *D. sechellia* female CHCs were gained or if dimorphism were lost in *D. simulans* (Jallon and David, 1987). If dimorphism were lost, sex determination might have a role in CHC production by 'masculinizing'

D. simulans females. If sexual dimorphism were gained and the sex determination pathway were involved, then the pathway would be responsible for *de novo* synthesis of new compounds, which is unlikely. Regardless, there is no strong evidence for the role of sex determination pathway genes here.

Epistasis and covariance of traits

Strong epistatic interactions are prevalent between the left arm of chromosome 3 and other loci for the dienes. In general, QTL for the monoenes do not have strong epistatic interactions, with the exception of 7-H, which is marginally more abundant in *D. sechellia* (Figure 2). One implication of the epistatic interactions for the dienes is that we have identified QTL in a single pathway such that the production of the dienes is dependent upon multiple enzymatic reactions, and thus, dienes are not produced in large quantities unless multiple genes have a *D. sechellia* allele (Figure 4).

Complex interactions are further exemplified by those of 7,11-HD, as it was correlated with all the other CHCs (Table 4). As the predominant CHC in *D. sechellia*, it was negatively correlated with the *D. simulans* CHCs (with the exception of 7-P) and positively correlated with the other *D. sechellia* CHCs. Thus, there is more evidence for common pathways for these CHCs.

All the epistatic interactions detected involved QTL of CHCs, which are more abundant in *D. sechellia* than in *D. simulans*. As these compounds are both longer and less saturated than the predominant ones of *D. simulans*, they are products from further along the biosynthesis pathway (Figure 1). Thus, the difference in epistatic interactions between these two species may reflect the more complicated biosynthesis required for *D. sechellia* rather than *D. simulans*. Detection of epistatic interactions for CHCs more abundant in *D. sechellia* than in *D. simulans* may be an artifact of the experimental design in that we used backcross progeny, and thus, cannot contrast all possible genotypes (no individuals were homozygous for *D. sechellia* alleles). However, the epistatic interactions were detected for the *D. sechellia* alleles in the heterozygous state, thus it is unlikely that we have missed epistatic interactions that would be evident only in a homozygous state.

Multiple candidate genes per QTL region, the presence of epistasis and the correlation of multiple CHCs may imply that isolating individual genes involved in these species differences may be difficult. Testing a single gene (for example, in a complementation test) may fail to reveal effects, although a particular candidate gene may be an important component of the biosynthetic pathway for a CHC. Thus, there may be support for a profound effect of genetic background on alleles that affect CHCs as suggested by Greenberg *et al.*, 2006.

Major models for the evolution of post-mating isolation between species rely on epistatic interactions within co-adapted gene complexes, such as the Dobzhansky–Muller models of hybrid incompatibilities (thought to underly Haldane's Rule; Turelli and Orr, 2000). Genes affecting sexual isolation have been argued to be more likely under direct rather than indirect selection, and thus may be less prone to such interaction effects than genes affecting post-mating isolation (Ritchie and Phillips, 1998). However, in this QTL study, we find that genes responsible for important species-specific divergence in pheromones are also affected by epistasis. The nature of the production of these important mating signals (by alternative biosynthetic pathways from a common precursor) may make CHCs prone to few genes of large effect with strong epistasis.

Acknowledgements

References

Antony C, Davis TL, Carlson DA, Pechiné J-M, Jallon J-M (1985). Compared behavioral responses of male *Drosophila melanogaster* (Canton S) to natural and synthetic aphrodisiacs. *J Chem Ecol* **11**: 1617–1629.

Basten CJ, Weir BS, Zeng Z-B (2005). *QTL Cartographer: A reference manual and tutorial for QTL mapping.* Department of Statistics North Carolina State University: Raleigh, NC, 189 pp.

Chan Yong TP, Jallon J-M (1986). Synthèse *de novo* d'hydrocarbures potentiellement aphrodisiaques chez les Drosophiles. *CR Acad Sci, Ser III* **303**: 197–202.

Chertemps T, Duportets L, Labeur C, Ueda R, Takahashi K, Saigo K *et al* (2007). A female-biased expressed elongase involved in long-chain hydrocarbon biosynthesis and courtship behavior in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **104**: 4273–4278.

Chertemps T, Duportets L, Labeur C, Ueyama M, Wicker-Thomas C (2006). A female-specific desaturase gene responsible for diene hydrocarbon biosynthesis and courtship behaviour in *Drosophila melanogaster*. *Insect Mol Biol* **15**: 465–473.

Chertemps T, Duportets L, Labeur C, Wicker-Thomas C (2005). A new elongase selectively expressed in *Drosophila* male reproductive system. *Biochem Biophys Res Commun* **333**: 1066–1072.

Churchill GA, Doerge RW (1994). Empirical threshold values for quantitative trait mapping. *Genetics* **138**: 963–971.

Cobb M, Jallon J-M (1990). Pheromones, mate recognition and courtship stimulation in the *Drosophila melanogaster* species subgroup. *Anim Behav* **39**: 1058–1067.

Coyne JA (1996). Genetics of a difference in male cuticular hydrocarbons between two sibling species, *Drosophila simulans* and *D sechellia*. *Genetics* **143**: 1689–1698.

Coyne JA, Crittenden AP, Mah K (1994). Genetics of a pheromonal difference contributing to reproductive isolation in *Drosophila*. *Science* **265**: 1461–1464.

Coyne JA, Elwyn S (2006). Does the *desaturase-2* locus in *Drosophila melanogaster* cause adaptation and sexual isolation? *Evolution* **60**: 279–291.

Coyne JA, Wicker-Thomas C, Jallon J-M (1999). A gene responsible for a cuticular hydrocarbon polymorphism in *Drosophila melanogaster*. *Genet Res* **73**: 189–203.

Dallerac R, Labeur C, Jallon J-M, Knipple DC, Roelofs WL, Wicker-Thomas C (2000). A Δ9 desaturase gene with a different substrate specificity is responsible for the cuticular diene hydrocarbon polymorphism in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **97**: 9449–9454.

Etges WJ, Ahrens MA (2001). Premating isolation is determined by larval-rearing substrates in cactophilic *Drosophila*

mojavensis. V. Deep geographic variation in epicuticular hydrocarbons among isolated populations. *Am Nat* **6**: 585–598.

Fang S, Takahashi A, Wu C-I (2002). A mutation in the promoter of *desaturase 2* is correlated with sexual isolation between Drosophila behavioral races. *Genetics* **162**: 781–784.

Ferveur J-F (2005). Cuticular hydrocarbons: their evolution and roles in *Drosophila* pheromonal communication. *Behav Genet* **35**: 279–295.

Gailey DA, Hall JC (1989). Behavior and cytogenetics of *fruitless* in *Drosophila melanogaster*: different courtship defects caused by separate, closely linked lesions. *Genetics* **121**: 773–785.

Gardner KM, Latta RG (2007). Shared quantitative trait loci underlying the genetic correlation between continuous traits. *Mol Ecol* **16**: 4195–4209.

Gleason JM, Jallon J-M, Rouault J-D, Ritchie MG (2005). Quantitative trait loci for cuticular hydrocarbons associated with sexual isolation between *Drosophila simulans* and *D sechellia*. *Genetics* **171**: 1789–1798.

Gloor G, Engels W (1992). Single-fly DNA preps for PCR. *Dros Inf Serv* **71**: 148–149.

Greenberg AJ, Moran JR, Coyne JA, Wu C-I (2003). Ecological adaptation during incipient speciation revealed by precise gene replacement. *Science* **302**: 1754–1757.

Greenberg AJ, Moran JR, Wu CI (2006). Proper control of genetic background with precise allele substitution: A comment on Coyne and Elwyn. *Evolution* **60**: 623–625.

Hoffmann AA, Hallas R, Sinclair C, Mitrovski P (2001). Levels of variation in stress resistance in *Drosophila* among strains, local populations, and geographic regions: patterns for desiccation, starvation, cold resistance and associated traits. *Evolution* **55**: 1621–1630.

Iwata H, Ninomiya S (2006). AntMap: Constructing genetic linkage maps using an ant colony optimization algorithm. *Breeding Sci* **56**: 371–377.

Jallon J-M, David JR (1987). Variations in cuticular hydrocarbons among the eight species of the *Drosophila melanogaster* subgroup. *Evolution* **41**: 294–302.

Jallon J-M, Lauge G, Orssaud L, Antony C (1988). Female pheromones in *Drosophila melanogaster* are controlled by the *doublesex* locus. *Genet Res* **51**: 17–22.

Labeur C, Dallerac R, Wicker-Thomas C (2002). Involvement of *desat1* gene in the control of *Drosophila melanogaster* pheromone biosynthesis. *Genetica* **114**: 269–274.

Legendre A, Miao X-X, Da Lage J-L, Wicker-Thomas C (2008). Evolution of a desaturase involved in female pheromonal cuticular hydrocarbon biosynthesis and courtship behavior in Drosophila. *Insect Biochem Mol Biol* **38**: 244–255.

Lockey KH (1988). Lipids of the insect cuticle: origin, composition and function. *Comp Biochem Physiol B Biochem Mol Biol* **89**: 595–645.

Manichaikul A, Dupuis J, Sen S, Broman KW (2006). Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. *Genetics* **174**: 481–489.

Marcillac F, Grosjean Y, Ferveur JF (2005). A single gene mutation alters production and discrimination of Drosophila sex pheromones. *Proc R Soc B Biol Sci* **272**: 303–309.

Noor MAF (2003). Genes to make new species. *Nature* **423**: 699–700.

Noor MAF, Coyne JA (1996). Genetics of a difference in cuticular hydrocarbons between *Drosophila pseudoobscura* and *D persimilis*. *Genet Res* **68**: 117–123.

Noor MAF, Feder JL (2007). Speciation genetics: evolving approaches. *Nat Rev Genet* **7**: 851–861.

Ohtsu T, Kimura MT, Katagiri C (1998). How Drosophila species acquire cold tolerance: Qualitative changes of phospholipids. *Eur J Biochem* **252**: 608–611.

Orr HA, Masly JP, Presgraves DC (2004). Speciation genes. *Curr Opin Genet Dev* **14**: 675–679.

Pennanec'h M, Bricard L, Kunesch G, Jallon J-M (1997). Incorporation of fatty acids into cuticular hydrocarbons of male and female *Drosophila melanogaster*. *J Insect Phys* **43**: 1111–1116.

Ritchie MG, Noor MAF (2004). Evolutionary genetics-gene replacement and the genetics of speciation. *Heredity* **93**: 1–2.

Ritchie MG, Phillips SDF (1998). The genetics of sexual isolation. In: Howard DJ and Berlocher SH (eds). *Endless Forms: Species and Speciation*. Oxford University Press: New York. pp 291–308.

Roelofs WL, Liu WT, Hao GX, Jiao HM, Rooney AP, Linn CE (2002). Evolution of moth sex pheromones via ancestral genes. *Proc Natl Acad Sci USA* **99**: 13621–13626.

Roelofs WL, Rooney AP (2003). Molecular genetics and evolution of pheromone biosynthesis in Lepidoptera. *Proc Natl Acad Sci USA* **100**: 9179–9184.

Rosenfield C-L, Man You K, Marsella-Herrick P, Roelofs WL, Knipple DC (2001). Structural and functional conservation and divergence among acyl-CoA desaturases of two noctuid species, the corn earworm, *Helicoverpa zea*, and the cabbage looper, *Trichoplusia ni*. *Insect Biochem Mol Biol* **31**: 949–964.

Rouault J-D, Marican C, Wicker-Thomas C, Jallon J-M (2004). Relations between cuticular hydrocarbon (HC) polymorphism, resistance against desiccation and breeding temperature; a model for HC evolution in D melanogaster and *D. simulans*. *Genetica* **120**: 195–212.

Savarit F, Ferveur JF (2002). Genetic study of the production of sexually dimorphic cuticular hydrocarbons in relation with the sex-determination gene transformer in *Drosophila melanogaster*. *Genet Res* **79**: 23–40.

Smadja C, Butlin RK (2009). On the scent of speciation: the chemosensory system and its role in premating isolation. *Heredity* **102**: 77–97.

Takahashi A, Tsaur S, Coyne JA, Wu C-I (2001). The nucleotide changes governing cuticular hydrocarbon variation and their evolution in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **98**: 3920–3925.

Tompkins L, McRobert SP (1995). Behavioral and pheromonal phenotypes associated with expression of loss-of-function mutations in the *Sex-lethal* gene of *Drosophila melanogaster*. *J Neurogenet* **9**: 219–226.

Toolson EC (1982). Effects of rearing temperature on cuticular permeability and epicuticular lipid composition in *Drosophila pseudoobscura*. *J Exp Zool* **222**: 249–253.

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

Tweedie S, Ashburner M, Falls K, Leyland P, McQuilton P, Marygold S *et al* (2009). FlyBase: enhancing Drosophila Gene Ontology annotations. *Nucleic Acids Res* **37**: D555–D559.

van Ooijen JW (1992). Accuracy of mapping quantiative trait loci in autogamous species. *Theor Appl Genet* **84**: 803–811.

Wang H-L, Ming Q-L, Zhao C-H, Wang CZ (2008). Genetic basis of sex pheromone blend difference between *Helicoverpa armigera* (Hübner) and *Helicoverpa assulta* (Guenée) (Lepidoptera: Noctuidae). *J Insect Phys* **54**: 813–817.

Wicker-Thomas C, Chertemps T (2010). Molecular biology and genetics of hydrocarbon production. In: Blomquist GJ and Bagnères AG (eds). *Insect Hydrocarbons: Biology, Chemistry and Chemical Ecology,* (in press).

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)