Genetic Dissection of Hybrid Incompatibilities Between *Drosophila simulans* and *D. mauritiana*. II. Mapping Hybrid Male Sterility Loci on the Third Chromosome

Yun Tao,*,†,[1] Zhao-Bang Zeng,‡ Jian Li,‡ Daniel L. Hartl† and Cathy C. Laurie*,[2]

**DCMB and Department of Zoology, Duke University, Durham, North Carolina 27708, ‡Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695 and †Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138*

Manuscript received January 22, 2003
Accepted for publication March 21, 2003

ABSTRACT

Hybrid male sterility (HMS) is a rapidly evolving mechanism of reproductive isolation in Drosophila. Here we report a genetic analysis of HMS in third-chromosome segments of *Drosophila mauritiana* that were introgressed into a *D. simulans* background. Qualitative genetic mapping was used to localize 10 loci on 3R and a quantitative trait locus (QTL) procedure (multiple-interval mapping) was used to identify 19 loci on the entire chromosome. These genetic incompatibilities often show dominance and complex patterns of epistasis. Most of the HMS loci have relatively small effects and generally at least two or three of them are required to produce complete sterility. Only one small region of the third chromosome of *D. mauritiana* by itself causes a high level of infertility when introgressed into *D. simulans*. By comparison with previous studies of the *X* chromsome, we infer that HMS loci are only ~40% as dense on this autosome as they are on the *X* chromosome. These results are consistent with the gradual evolution of hybrid incompatibilities as a by-product of genetic divergence in allopatric populations.

ALLOPATRIC speciation, or genetic divergence of subpopulations in geographic isolation, may be the most common mechanism of speciation (Mayr 1942). According to this mechanism, two populations gradually diverge through the fixation of different alleles at a number of loci and eventually acquire reproductive isolation. One common form of reproductive isolation is postzygotic hybrid incompatibility (HI), in which hybrids between the divergent populations are inviable or sterile. According to the Dobzhansky-Muller model, HI is an "accidental" by-product of divergence due to negative epistatic interactions between loci fixed for different alleles in the two populations. These negative interactions may occur in hybrids, but not in the parental populations, because the newly fixed alleles have never occurred together or been "tested" by natural selection within the same genome (Dobzhansky 1937, p. 256; Muller 1942). Under this mechanism of speciation, HI is expected to be polygenic and attributable to genes with a variety of different functions.

Investigation of the genetic architecture of HI has a long history (*e.g.*, Dobzhansky 1936, 1975) and much of this work has been done with Drosophila (reviewed in Coyne 1992; Coyne and Orr 1998; Wu and Hollocher 1998). One of the major results, referred to as the "faster male" phenomenon, is that hybrid male sterility (HMS) seems to evolve much more rapidly than hybrid female sterility or hybrid inviability (True *et al.* 1996b; Wu *et al.* 1996; Tao *et al.* 2003, this issue). Genetic dissection of HMS has been very fruitful, particularly for the species pair of *Drosophila simulans* and *D. mauritiana*. The *X* chromosome has been carefully analyzed to detect and map HMS loci, using an introgression approach, and nine loci were localized to ~40% of the chromosome (Coyne and Charlesworth 1986, 1989; Perez *et al.* 1993; Cabot *et al.* 1994; Perez and Wu 1995; Davis and Wu 1996; reviewed in Wu and Hollocher 1998). A high point of this series of introgression studies was the cloning of *Ods*H, a putative hybrid male sterility gene (Ting *et al.* 1998), and some interesting insights have been obtained from analyzing its molecular evolution (Ting *et al.* 2000).

Comparison of the genetic architecture of HI loci on the *X* and autosomes is potentially very valuable because population genetic theories predict different evolutionary dynamics for genes on these two types of chromosome (Charlesworth *et al.* 1987; Orr and Betancourt 2001). However, so far, genetic dissection of autosomal genes affecting HI has not been done at the same level of detail as the *X* chromosomal studies. Two studies of HMS loci on the autosomes of *D. simulans* × *D. mauritiana* hybrids have been published (Hollocher and Wu 1996; True *et al.* 1996b), but these had limited resolution and gave ambiguous results concerning the relative density of HMS factors on the *X* and autosomes. In a companion article (Tao *et al.* 2003, this issue) we report a higher-resolution analysis of the whole third chromo-

[1]*Corresponding author:* Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Ave., Cambridge, MA 02138. E-mail: ytao@oeb.harvard.edu

[2]*Present address:* Monsanto Company, 800 N. Lindbergh Blvd., N2NA, St. Louis, MO 63167.

some through a series of overlapping introgressions. In that study, a statistical model was used to arrive at the conclusion that the *X* chromosome hosts 2.5 times as many HMS loci per unit of euchromatin as the third chromosome. Here we extend this analysis of *D. simulans* × *D. mauritiana* hybrids to identify and localize HMS loci on the third chromosome through qualitative genetic mapping and quantitative trait locus (QTL) mapping.

MATERIALS AND METHODS

Introgression lines, markers, fertility assay, and single-fly genotyping: A full account of the methods for constructing the introgression lines, developing allele-specific oligonucleotide (ASO) markers, assaying male fertility, and genotyping single flies has been given in the companion article (Tao *et al.* 2003, this issue). Briefly, we constructed introgression lines in which each line has a single $P[w^+]$-tagged *D. mauritiana* segment on the third chromosome in an otherwise isogenic pure *D. simulans* (*sim*B) background. ASO markers were used to define the extent of the introgressions. The whole third chromosome was covered by multiple overlapping introgressions. All heterozygous introgressions are fertile and the lines were maintained by crossing five heterozygous $P[w^+]$-insert males (orange eyed) with ~12–15 *sim*B females in each generation.

Homozygous or *trans*-heterozygous males of the $P[w^+]$-inserts (red-eyed) obtained from these lines were assayed for male fertility in a mating test and were also genotyped with ASO markers to detect possible recombination within the *D. mauritiana* segments (Tao *et al.* 2003, this issue). Females for the mating test were *D. simulans w; e*, thus providing markers to eliminate the possibility of falsely declaring male fertility due to nonvirginity of the tester females. The genotyping of individual males in fertility tests started with the two outermost ASO markers defining an introgressed *D. mauritiana* segment inherited maternally. If both were *D. mauritiana* alleles, it was assumed that no crossover occurred within the segment. Otherwise, all inner ASO markers were genotyped and the crossover point was localized between two markers. The phenotypic and genotypic data thus obtained were used in mapping the HMS loci.

Qualitative genetic mapping of HMS: The crossing scheme in Figure 1A was used to construct the mapping population. Ten lines with introgressions that are not fully fertile when homozygous and that cover all of 3R (marked with an asterisk in Figure 2) were used to generate a total of 4461 progeny. For each line, several hundred red-eyed progeny (range of 311–509) were tested for fertility and genotyped for ASO markers. Males with increased fertility were usually recombinants. HMS loci were inferred by associating fertility shifts with genotypes, as described further in RESULTS. Additional evidence for HMS loci was obtained by comparing the positions of introgressed segments having different fertility levels (Figure 2). In this process we implicitly assume that *D. mauritiana* alleles always have a negative effect on fertility and that they interact with each other additively and synergistically but not antagonistically.

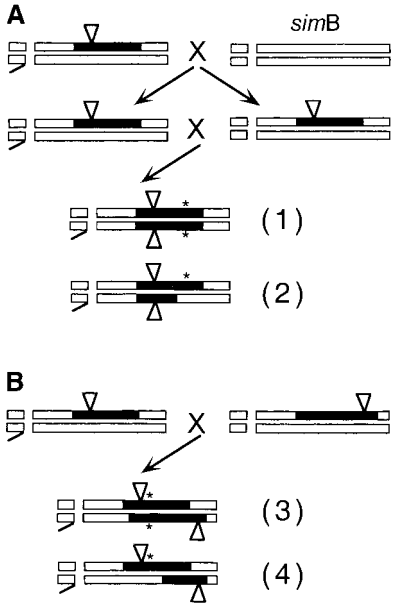
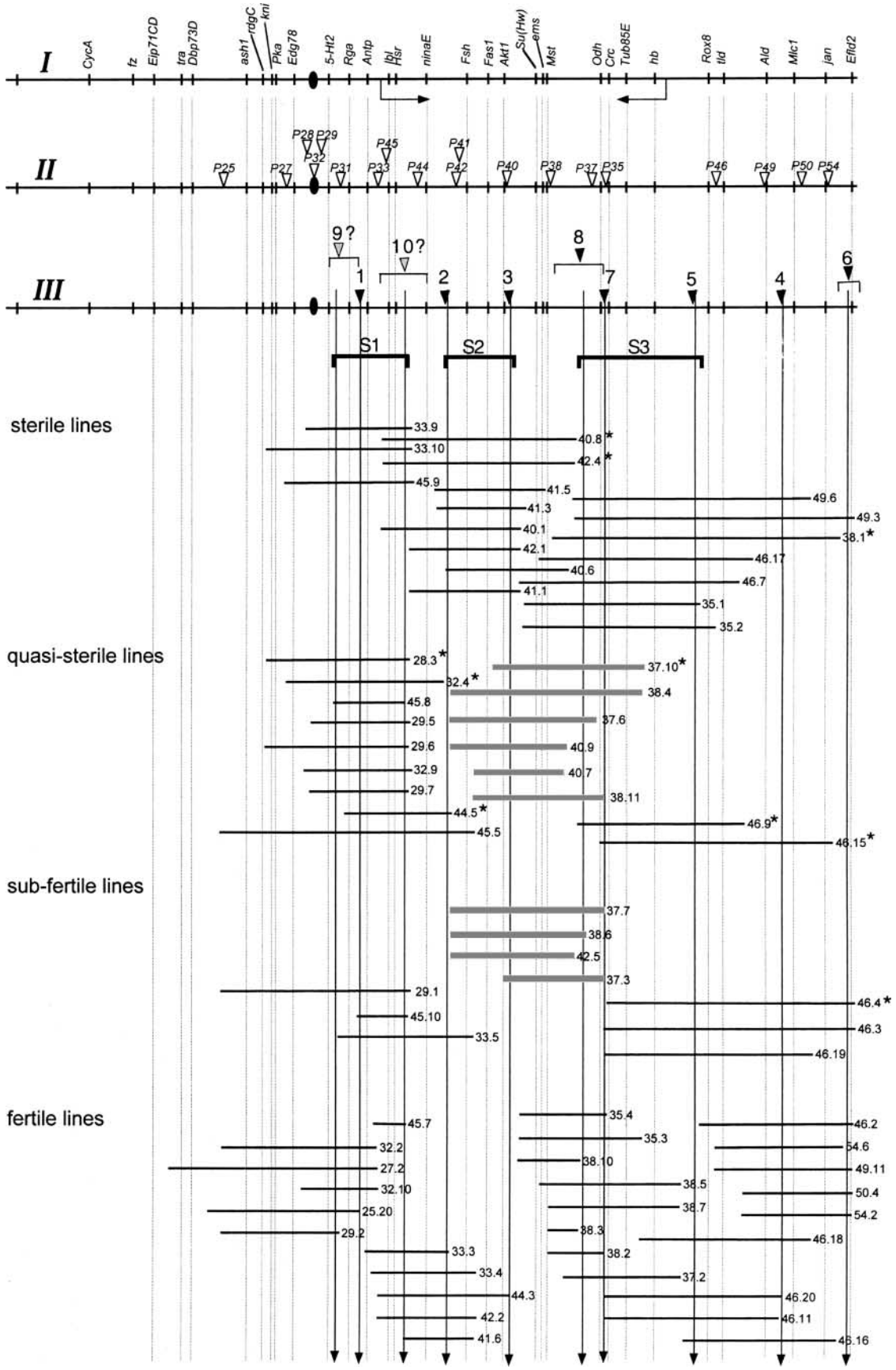


FIGURE 1.—Crossing schemes for HMS mapping. Chromosomes are represented by bars. Short bar, the *X* chromosome; short hooked bar, the *Y* chromosome; open bars, *D. simulans*; solid bars, *D. mauritiana*. The second chromosome, which was marked by *nt*, is not shown. The $P[w^+]$-inserts are represented by open triangles (Tao *et al.* 2003, this issue). (A) Making homozygous males from a single introgression line. The homozygous males are usually of genotype 1. Occasionally recombinants are obtained as genotype 2. If the two genotypes have different fertility, HMS factor(s) must be in the region as indicated by *. The location of a crossover on the maternal chromosome can be mapped with ASO markers. (B) Two different introgression lines (tagged by different $P[w^+]$-inserts) are used to produce males homozygous for a certain region. A similar rationale as in A is applied to infer HMS factor(s) from the fertility differences in genotypes 3 and 4.

Both the sample sizes and progeny counts for most genotypes were small, thus rendering routine statistical methods (parametric or nonparametric) unsuitable to ascertain fertility differences among genotypes. We overcame this problem by a simple permutation test: Two samples of offspring counts

FIGURE 2.—Ten HMS factors implicated on 3R. The proximal part of 3L, the centromere (solid oval), and the whole of 3R are shown for (I) the ASO markers used in this experiment, (II) positions of the $P[w^+]$-inserts (open triangles) in the introgression lines, and (III) positions of the 10 HMS factors mapped on 3R. Each horizontal bar represents an introgressed *D. mauritiana* segment, with the serial number of each line indicated at the right. An asterisk indicates lines used in the qualitative genetic mapping. Males are classified as sterile, quasi-sterile, subfertile, or fertile (Tao *et al.* 2003, this issue). Thick bars represent introgression lines in which the males sired female-biased progeny. By the mapping scheme of Figure 1A, eight HMS loci (solid triangle,1–8 on III) are mapped as elaborated in Figures 4 and A1–A7. For adequately explaining all observed sterility on 3R, two more loci (9 and 10, stippled triangles) need to be invoked (see text for details). For the four factors (6, 8, 9, and 10) not narrowed down within two ASO markers, the ranges of their possible locations are also shown. Three regions (S1, S2, and S3) that each can cause full sterility are also indicated. Note that although line 46.9 is formally classified as quasi-sterile, it is essentially sterile (Figure A5).


I
CycA
tz
Eip71CD
tra
Dbp73D
ash1
rdgC
kni
Pka
Edg78
5-Ht2
Rga
Antp
lbl
Hsr
ninaE
Fsh
Fas1
Akt1
Su(Hw)
ems
Mst
Odh
Crc
Tub85E
hb
Rox8
tld
Ald
Mlc1
jan
Efld2
II
P25
P27
P28
P29
P32
P31
P33
P45
P44
P41
P42
P40
P38
P37
P35
P46
P49
P50
P54
III
9?
10?
8
6
1
2
3
7
5
4
S1
S2
S3
sterile lines
33.9
33.10
40.8*
42.4*
45.9
41.5
49.6
41.3
49.3
40.1
38.1*
42.1
46.17
40.6
46.7
41.1
35.1
35.2
quasi-sterile lines
28.3*
32.4*
45.8
29.5
29.6
32.9
29.7
44.5*
45.5
37.10*
38.4
37.6
40.9
40.7
38.11
46.9*
46.15*
sub-fertile lines
37.7
38.6
42.5
37.3
29.1
45.10
33.5
46.4*
46.3
46.19
fertile lines
45.7
32.2
27.2
32.10
25.20
29.2
33.3
33.4
44.3
42.2
41.6
35.4
35.3
38.10
38.5
38.7
38.3
38.2
37.2
46.2
54.6
49.11
50.4
54.2
46.18
46.20
46.11
46.16

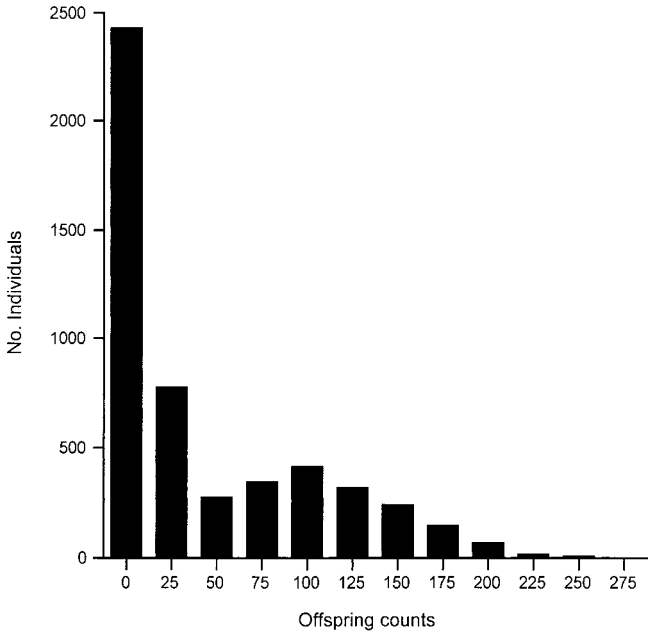


FIGURE 3.—The distribution of offspring counts from 5025 introgression males used in the QTL mapping. One outlier siring 491 offspring is not shown. There were 2423 sterile males (column 0) and the rest (2602) sired at least one offspring.

were merged and then split randomly into two samples of the original sizes. The mean offspring difference of these random samples was calculated and compared to the observed difference. This procedure was repeated 1000 times and the empirical significance P_0 (one-tailed) was obtained. A one-tailed test was used because we assume that short introgressions derived from long introgressions by recombination are more fertile (Figure 1A).

QTL mapping: The phenotypic and genotypic data for QTL mapping were collected from a total of 5025 male progeny from the two crossing schemes in Figure 1. Introgression lines covering the whole third chromosome were used. From the scheme in Figure 1A, 3801 males from 196 introgression lines were used. From the scheme in Figure 1B, 1224 males were generated from 87 crosses involving 58 introgression lines. Figure 3 shows the distribution of offspring counts in the males used for QTL mapping. Among the 5025 males tested, 2423 were sterile and 2602 sired at least one offspring.

Male flies for the qualitative mapping and for QTL mapping are from independent collections, although the introgression lines used are not. Eight of the 10 introgression lines (except 44.5 and 37.10) used in qualitative mapping were also used in QTL mapping, with a total of 279 males.

Two approaches were taken to mapping QTL. In one approach, all males were analyzed and fertility was treated as a binary trait (z), with males characterized as either fertile (z = 1 for at least one offspring) or sterile (z = 0 for no offspring). In the other approach, only nonsterile males were analyzed and fertility was treated as a continuous trait (y = number of offspring produced in a mating test). In both cases, QTL mapping was performed by the method of multiple-interval mapping (MIM; KAO *et al.* 1999).

The binary-trait model: A threshold model was used to account for z with an assumption that the ith male has a certain value of an underlying fertility score (x) that renders the male fertile when it reaches a threshold θ,

$$z_i = \begin{cases} 1 & \text{when } x_i > \theta \\ 0 & \text{when } x_i < \theta. \end{cases}$$

We further assume that the underlying fertility score is affected by m QTL according to the following model,

$$x_i = \mu_x + \sum_{r=1}^{m} \alpha_r u_{ir} + \sum_{r=1}^{m} \delta_r v_{ir} + \sum_{r<s}^{E_1} (\alpha\alpha)_{rs} u_{ir} u_{is} + \sum_{r<s}^{E_2} (\alpha\delta)_{rs} u_{ir} v_{is} + \sum_{r<s}^{E_3} (\delta\alpha)_{rs} v_{ir} u_{is} + \sum_{r<s}^{E_4} (\delta\delta)_{rs} v_{ir} v_{is} + e_{xi},$$

where $e_{xi} \sim N(0,\sigma_x^2)$; α_r and δ_r are additive and dominant effects on fertility score from the rth QTL; u_{ir} and v_{ir} are the effect variables corresponding to the genotype of the rth QTL for the ith individual (M, *D. mauritiana*; S, *D. simulans*),

$$u_{ir} = \begin{cases} 1 (M_r M_r) \\ 0 (M_r S_r) \\ -1 (S_r S_r) \end{cases} \quad \text{and} \quad v_{ir} = \begin{cases} -\tfrac{1}{2} (M_r M_r) \\ \tfrac{1}{2} (M_r S_r) \\ -\tfrac{1}{2} (S_r S_r) \end{cases};$$

the terms $(\alpha\alpha)_{rs}$, $(\alpha\delta)_{rs}$, $(\delta\alpha)_{rs}$, and $(\delta\delta)_{rs}$ are epistatic effects (additive × additive, additive × dominant, dominant × additive, and dominant × dominant); and E_1, E_2, E_3, and E_4 are the numbers of pairs of QTL that show significant epistatic effects of the four kinds through a model selection process. See KAO and ZENG (2002) for details about parameterization of the epistatic effects.

The conditional probability of z given a QTL genotype specified by u_{ir} and v_{ir} is

$$\begin{aligned} P(z_i = 1 | u_{ir}, v_{ir}) &= \int_{\theta}^{\infty} f(x_i)\, dx_i \\ &= \int_{\theta}^{\infty} \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{(x_i - \mu_x - \Delta U_i)^2}{2\sigma_x^2} \right) dx_i \\ &= 1 - \Phi\left(\frac{\theta - \mu_x - \Delta U_i}{\sigma_x} \right) \end{aligned}$$

and $P(z_i = 0 | u_{ir}, v_{ir}) = 1 - P(z_i = 1 | u_{ir}, v_{ir}) = \Phi((\theta - \mu_x - \Delta U_i)/\sigma_x)$, where $\Phi(\cdot)$ is a cumulative normal density function and

$$\Delta U_i = \sum_{r=1}^{m} \alpha_r u_{ir} + \sum_{r=1}^{m} \delta_r v_{ir} + \sum_{r<s}^{E_1} (\alpha\alpha)_{rs} u_{ir} v_{is} + \sum_{r<s}^{E_2} (\alpha\delta)_{rs} u_{ir} v_{is} + \sum_{r<s}^{E_3} (\delta\alpha)_{rs} v_{ir} u_{is} + \sum_{r<s}^{E_4} (\delta\delta)_{rs} v_{ir} v_{is}.$$

In this notation, Δ is a vector of QTL-effect parameters (α_r, δ_r, . . .), and U_i is a vector of effect variables (u_{ir}, v_{ir}, . . .). In this analysis, the threshold θ and QTL effects Δ are estimated by assuming that $\mu_x = 0$ and $\sigma_x^2 = 1$.

Of course, QTL genotypes are not observed, but we can calculate the probability of each possible QTL genotype for an individual, conditional on the observed marker genotypes, if we assume a specific model of QTL number and positions. There are three possible genotypes for each QTL and 3^m possible joint genotypes for m QTL. Thus

$$P(z_i = 1 | \text{markers}) = \sum_{j=1}^{3^m} p_{ij} (1 - \Phi(\theta - \Delta U_{ij})),$$

where p_{ij} is the conditional probability of the jth multilocus QTL genotype given marker genotypes and a specific QTL model (see JIANG and ZENG 1997 for the calculation with missing and dominant markers). Although theoretically there are 3^m possible joint genotypes for m QTL, only a handful of joint QTL genotypes have nonzero conditional probabilities for any individual in this introgression experiment, so the numerical analysis is not particularly prohibitive.

The likelihood of the data (z) is

$$L(z|\theta, \Delta) = \prod_{i=1}^{n} P(z_i = 1|\text{markers})^{z_i} P(z_i = 0|\text{markers})^{1-z_i} = \prod_{i=1}^{n} \sum_{j=1}^{3^m} p_{ij}(1 - \Phi(\theta - \Delta U_{ij}))^{z_i} \Phi(\theta - \Delta U_{ij})^{1-z_i}.$$

This is a multiple-interval mapping version (Kao *et al.* 1999) of the threshold model discussed in Xu and Atchley (1996).

The continuous-trait model: For offspring counts y, we use the following model,

$$y_i = \mu_y + \sum_{r=1}^{m} a_r u_{ir} + \sum_{r=1}^{m} d_r v_{ir} + \sum_{r<s}^{E_1} (aa)_{rs} u_{ir} u_{is} + \sum_{r<s}^{E_2} (ad)_{rs} u_{ir} v_{is} + \sum_{r<s}^{E_3} (da)_{rs} v_{ir} u_{is} + \sum_{r<s}^{E_4} (dd)_{rs} v_{ir} v_{is} + e_{yi},$$

where $e_{yi} \sim N(0, \sigma_y^2)$ and a_r, d_r, $(aa)_{rs}$, $(ad)_{rs}$, $(da)_{rs}$, and $(dd)_{rs}$ are additive, dominant, and epistatic effects of QTL on y. The likelihood of the data (y) is

$$L(y|\mu_y, \sigma_y^2, D) = \sum_{i=1}^{n} \sum_{j=1}^{3^m} p_{ij} \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(y_i - \mu_y - DU_{ij})^2}{2\sigma_y^2}\right) = \prod_{i=1}^{n} \sum_{j=1}^{3^m} p_{ij} \phi\left(\frac{y_i - \mu_y - DU_{ij}}{\sigma_y}\right),$$

where $\phi(\cdot)$ is a normal density function and D is a vector of QTL-effect parameters (a_r, d_r, . . .).

Parameter estimates and hypothesis testing: Maximum-likelihood parameter estimates were obtained with an expectation-maximization (EM) algorithm (as in Kao and Zeng 1997 for continuous trait; J. Li and Z-B. Zeng, unpublished results for categorical trait). The test for each QTL effect was performed by a likelihood ratio (LOD score), conditional on other selected QTL effects. In this analysis, the additive and dominance effects of a QTL were searched and tested together first and significant epistatic effects were identified subsequently. For QTL r, the null hypothesis is H_0: $a_r = 0$ and $d_r = 0$, and the alternative hypothesis is H_1: $a_r \neq 0$ and $d_r \neq 0$. The test statistic is

$$\text{LOD} = \log_{10}[L_1(a_r \neq 0, d_r \neq 0)/L_0(a_r = 0, d_r = 0)],$$

where L_0 and L_1 are the maximum likelihoods under H_0 and H_1, respectively, conditional on all other selected QTL effects.

Model selection: The initial model selection (IMS) was performed by regression of the trait value on marker genotypes (logistic regression for z and linear regression for y), using a backward stepwise selection process. The procedure is as follows:

IMS-a: The number of markers n_m in the model is set equal to the total number of markers n_M. This model is considered as the full model, in which each marker is fitted with an additive and a dominant effect.

IMS-b: Each marker in the full model is eliminated in turn to produce n_m reduced models. Note that when one marker in the full model is eliminated, two corresponding effects are eliminated.

IMS-c: For each reduced model, the test statistic for the dropped marker is computed, which can be either the ratio of log-likelihood under the full and reduced models (for continuous trait) or Wald statistics (for binary trait).

IMS-d_1: If the likelihood-ratio test statistic is <2 (on LOD score) for the least significant marker, that marker is dropped from the full model. Then the corresponding reduced model is set as the full model and n_m(new) = n_m(old) − 1, and the process returns to step IMS-b.

IMS-d_2: Otherwise (LOD > 2 for all markers), the process is stopped and the current full model with n_m markers is used as the initial model for subsequent analysis.

After the initial model is selected (with the number of QTL $n_Q = n_m$), model selection is refined using MIM, as outlined in the following:

MIM-a: The positions of QTL in the current model are optimized. This optimization is performed for each QTL in turn in a sequential order. For each QTL, the model likelihood is evaluated in the vicinity of the previous position, conditional on the current positions of other QTL. The position that gives the maximum likelihood is chosen for the new positions of the QTL.

MIM-b: A new QTL may be added. The chromosome is scanned for the position of a new QTL that gives the maximum likelihood of the model. If the LOD score for this putative QTL is >2, the QTL is added into the model and then the process goes to step MIM-c; otherwise no new QTL is added, the QTL identification phase stops, and the process moves on to analyze epistatic effects.

MIM-c: All QTL in the current model are tested for significance again in a sequential order. This procedure is similar to steps IMS-b and IMS-c in the initial model selection. The likelihood of the full model is compared with that of a reduced model in which one QTL is dropped. If a QTL is not significant (LOD < 2), it is dropped from the model. If all QTL are significant, the model is unchanged. The process then returns to step MIM-a.

Here we use a LOD score of 2 as a threshold for QTL detection. In principle, a residual permutation test could be used to empirically estimate the threshold for model selection (Zeng *et al.* 1999). However, an overwhelming amount of computer time would be required in this case, because of the large sample size (>5000) and number of detected QTL involved. For one chromosome, a LOD score of 2 is reasonably conservative to adjust for the multiple testing.

Epistasis: In this introgression experiment, only closely linked QTL segregate in a given cross (Figure 1), so only epistatic effects between close QTL are considered. Here "close" QTL mean pairs of QTL with no more than two other QTL between them. For each QTL pair considered, one of the pairwise epistatic interactions between additive and dominance effects (four interactions for each QTL pair) was either present or absent from the model.

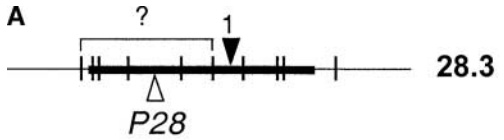
a. At each step, each of the epistatic effects (except those already selected) was added to the current model in sequence. The one that increased the likelihood the most was selected for inclusion in the model.

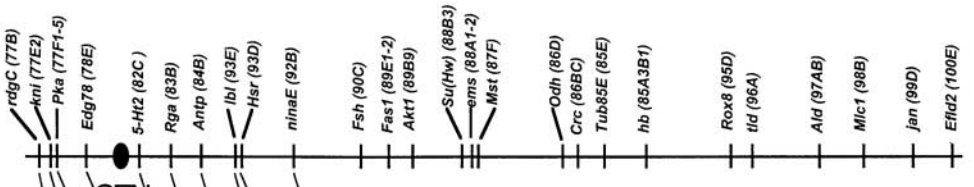
b_1. If the likelihood increase exceeded a preset criterion (LOD = 1), the epistatic effect was retained in the model, and the process returned to step a.

b_2. If the likelihood increase was less than the preset criterion, the search process was stopped.

Throughout this article, we use HMS and QTL to refer to the loci for hybrid male sterility mapped through the qualitative genetic and the QTL mapping procedures, respectively. When necessary, we use QTL_x and QTL_y to distinguish the QTL mapped from the binary or continuous traits, respectively.

Genetic marker map: Genetic distances between ASO markers are required in the QTL mapping algorithm. These distances were obtained by analyzing data from the QTL-mapping population itself. For any two adjacent ASO markers on a maternally transmitted haplotype, the parental type (PT) is that of *D. mauritiana* alleles at both loci, whereas the nonparental type (NPT) is that of one *D. mauritiana* allele at one locus (proximal to $P[w^+]$-insert) and a *D. simulans* allele at the other (distal to $P[w^+]$-insert). The crossover rate is calculated as NPT/(NPT + PT), from which the genetic distance was calculated using the Haldane mapping function.





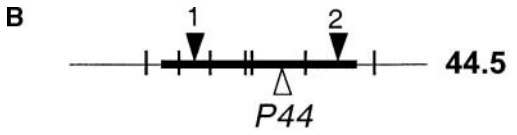
		Sample size	Siring male (%)	Mean fertility (±s.d.)	
A.	S M M M M M M M S	442	5.9	0.2±1	
B.	S M M M M M M H S	5	0	0	
C.	S M M M M M M H H S	9	0	0	
D.1	S M M M M M H H H S	2	0	0	*①
D.2	S M M M M M H H H S	1		3	*①
D.3	S M M M M M H H H S	1		27	*①
D.4	S M M M M M H H H S	1		116	*①
D.5	S M M M M M H H H S	1		77	*①
D.6	S M M M M M H H H S	1		173	*①
E.	S M M M M H H H H S	1		77	
F.	S H M M M M M M M S	1		0	
G.1	S H H M M M M M M S	2	0	0	
G.1	S H H M M M M M M S	1		7	
H.1	S H H H M M M M M S	5	0	0	
H.2	S H H H M M M M M S	1		3	

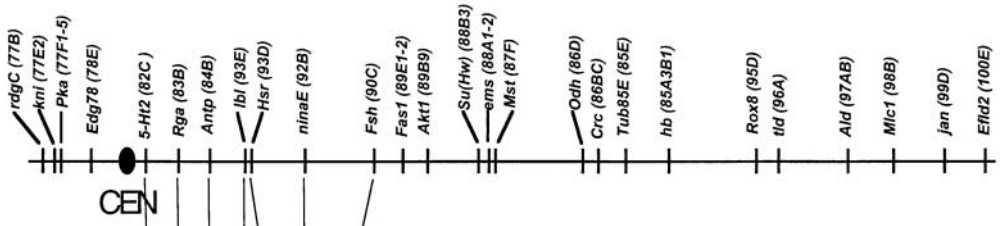
FIGURE 4.—Mapping HMS factors by a qualitative method. (A–C) Top, the introgressed *D. mauritiana* segment. Thin line, *D. simulans*; thick line, *D. mauritiana*. Open triangles point to the positions of $P[w^+]$-insert. Solid triangles point to the positions of putative HMS factors. Middle, The ASO markers and their cytological positions on part of the third chromosome. The solid oval represents the centromere (CEN). Bottom, The genotype and phenotype of individuals used in mapping. Each genotype is defined by ASO markers: S (homozygous for *D. simulans* allele), M (homozygous for *D. mauritiana* allele), and H (heterozygous). The sample size and fertility (percentage of males that sired any progeny and mean offspring ± SD) of each genotype are shown. The fertility shift that is the key evidence for declaring an HMS locus is indicated by a star followed by a circled number identifying that HMS locus. (A) Line 28.3. The $P[w^+]$-insert *P28* is to the left of the centromere. The HMS factor marked with the question mark is inferred by comparing H in A with A.3 and B in B. (B) Line 44.5. Note that categories A.1 and A.2 have the same genotype, but were handled somewhat differently. All of the males in A.1 are sterile; only four were genotyped and the rest are assumed to be of the same genotype. All of the males in A.2 sired some offspring and all were genotyped. (C) Line 32.4. The position of the $P[w^+]$-insert *P32* relative to the centromere is not determined. It is unclear whether HMS 2 is present in this line or not (see text for details).

RESULTS

The recombination rate within introgressed segments: In this study, the map distance is 104.8 cM for the third chromosome when estimated from segments of *D. mauritiana* introgressed into the genetic background of *D. simulans*. Compared to the 211.3 cM estimated in a pure *D. mauritiana* background (TRUE *et al.* 1996a), the recombination rate is reduced ~50%. The introgression map is also much shorter than that measured in F_1 hybrids (174.6 cM; ZENG *et al.* 2000). However, it is longer than a previous estimate that was also made in an introgression background (~54.9 cM; TRUE *et al.* 1996b). The previous introgression estimate may not be very accurate, since it was inferred statistically from the distribution of introgression lengths after 15 generations of backcrossing (TRUE *et al.* 1996b). An underestimate of the map length would result if the distribution of introgression lengths was biased upward, which might have been the case, since selection would favor longer introgressions due to heterosis in the highly inbred genetic background used in that experiment. In any case, the shorter map of introgressed *D. mauritiana* segments in a *D. simulans* rather than in a *D. mauritiana* background indicates the presence of *trans*-acting factors affecting the rate of recombination.

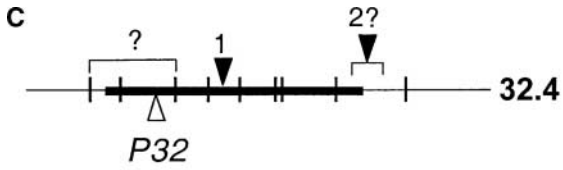
HMS loci identified on 3R by qualitative genetic mapping: A total of 218 introgression lines covering the whole third chromosome have been tested for male fertility. There were 115 fertile, 25 subfertile, 30 quasi-sterile, and 48 sterile lines. The fertility class is based on the average number of offspring that one homozygous male can sire: sterile (0 per male), quasi-sterile (0–10), subfertile (10–45), and fertile (>45; TAO *et al.* 2003, this issue). Eight HMS loci on 3R were detected by analyzing the progeny of 10 crosses using the scheme in Figure 1A. Two additional HMS are required to ex-

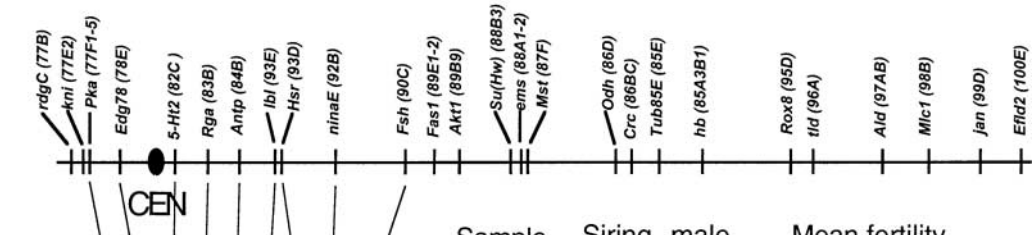




		Sample size	Siring male (%)	Mean fertility (±s.d.)	
A.1	S M M M M M S	438	9.7	0.7±3	*②
A.2	S M M M M M S	47			
A.3	S M M M M M S	1		83	*②
B.1	S M M M M H S	1		13	
B.2	S M M M M H S	1		55	
C.1	S H M M M M S	1		7	*①
C.2	S H M M M M S	1		60	*①
D.	S H H M M M S	13	100	87±30	
E.	S H H M M H S	3	100	107±33	
F.	S H H H M M S	2	100	86±39	
G	S H H H H M S	1		77	

FIGURE 4.—*Continued.*





		Sample size	Siring male (%)	Mean fertility (±s.d.)	
A.	S M M M M M M M S	408	14	0.6±3	*②?
B.	S M M M M M M H S	17	18	3±10	*②?
C.	S M M M M H H H S	4	25	1±3	*①
D.	S M M M H H H H S	1		4	*①
E.1	S M M H H H H H S	1		101	*①
E.2	S M M H H H H H S	1		38	*①
E.3	S M M H H H H H S	1		97	*①
F.1	S M H H H H H H S	1		55	
F.2	S M H H H H H H S	1		60	
F.3	S M H H H H H H S	1		1	
G	S H M M M M M M S	6	50	5±7	

TABLE 1

Summary of HMS qualitative mapping

HMS loci	Location (flanking markers)	Major evidence (lines and genotypes where fertility shifts)
1	*Rga-Antp*	28.3, quasi-sterile → fertile within D 44.5, quasi-sterile (C.1) → fertile (C.2) 32.4, quasi-sterile (C) → fertile (E)
2	*ninaE-Fsh*	44.5, quasi-sterile → fertile within A Mapped and named as *broadie* (Tao *et al.* 2001) 32.4, quasi-sterile (A) → quasi-sterile (B) ($P_{\bar{\sigma}} = 0.014$)
3	*Akt1-Su(Hw)*	42.4, quasi-sterile → fertile within D 40.8, quasi-sterile → fertile within D 37.10, quasi-sterile → fertile within D, sex ratio (k) shifts from 0.72 to 0.48 Mapped and named as *tmy* (Tao *et al.* 2001)
4	*Ald-Mlc1*	46.15, fertility increases gradually from quasi-sterile (A and B) to subfertile (C, D, and E). 46.4, fertility increases from C to D ($P_{\bar{\sigma}} < 0.001$), although both are fertile. The current location of HMS 4 is most parsimonious to explain the data, but several loci, each with smaller effects, are also likely.
5	*Hb-Rox 8*	46.15, subfertile (G) → fertile (H) ($P_{\bar{\sigma}} = 0.002$) 46.4, subfertile (H) → fertile (I) ($P_{\bar{\sigma}} < 0.001$) 38.1, 30% of males siring progeny within F
6	Distal to *jan*	46.4, A → B ($P_{\bar{\sigma}} = 0.004$), although both are subfertile
7	*Odh-Crc*	Comparison of 46.4 and 46.15; the left end of 46.15 must carry HMS factor(s) Also implicated in 37.10 as recombination distal to *P37* increases fertility (A → B and C)
8	*P38-HMS 7*	Comparison of 46.15 and 46.9 and 46.4; the left end of 46.9 must carry HMS factor(s) Also implicated in 46.9, quasi-sterile (A) → subfertile (B) ($P_{\bar{\sigma}} < 0.001$)

Arrows represent (1) fertility shifts between classes and (2) significant increases in fertility as detected by permutation test ($P_{\bar{\sigma}}$ indicated).

plain all differences in fertility level among introgression lines covering 3R. Figure 2 summarizes the inferred locations of these 10 HMS loci, the 10 lines used to generate the mapping data (marked by asterisks), and other informative introgression lines covering 3R.

The existence of HMS factor(s) is inferred by a fertility shift among genotypes that differ in one or few markers (Figure 1). A fertility shift between two genotypes was declared by one of two criteria: They belong to different fertility classes, or the numbers of their offspring differ significantly as shown by the permutation test.

In the following text, as well as in Figure 4, we describe the details of three examples of mapping HMS loci using the scheme of Figure 1A. The other mapping details for seven additional lines are summarized in the appendix. The critical evidence for declaring the eight HMS loci is summarized in Table 1.

Introgression 28.3: The genotypes and the corresponding phenotypes are displayed in Figure 4A. Genotype A is for the parental, nonrecombinant homozygous type, which is classified as quasi-sterile. Genotypes B and C were tested as sterile, but their sterility may not really be different from A, given the small sample size. A fertility shift from quasi-sterile to fertile obviously occurs within genotype D, suggesting an HMS factor (1) between markers *Rga* and *Antp*. Additional evidence for HMS factor 1 is noted in Table 1. This factor is not necessarily a single gene. Here and throughout this article, we consider that each HMS "locus" may actually be a cluster of genes of small effect.

On the other end of the 28.3 introgression, recombinants between *P28* and *rdgC* are all similar to the parental genotype, providing no evidence for a major factor in that region. However, because the sample size is small and sterile or quasi-sterile males are minimally informative, these data do not rule out the possibility of HMS factor(s) in that region. Indeed, a comparison with line 44.5 described below provides some evidence for the existence of at least one factor in that interval, hence the bracket marked as "?" in Figure 4A.

Introgression 44.5: This line and line 37.10 (see appendix) were the first two lines used for qualitative mapping. Unfortunately the sterile males (A.1 in Figure 4B and Figure A3) from these two lines were not saved for genotyping. So, here we assume that most or all of the

sterile males were of parental type (A.1). Some males of the same genotype (A.2) sired small numbers of offspring, so line 44.5 is classified as quasi-sterile. Interestingly, a single male (A.3) with the same genotype as A.1 and A.2 was very fertile. All markers were scored twice to ensure correct genotyping and nonvirginity of tester females is excluded, although mislabeling of the male cannot be excluded. One may invoke a double crossover between *Rga* and *Antp* to explain the data, but for such a narrow range (0.79 cM) this is unlikely ($P = 0.03$ with Bonferroni correction). A more likely scenario is that an HMS factor ("2") to the right of *ninaE* is present. The fertility of two flies of genotype B supports this interpretation and additional evidence is noted in Table 1. On the other end of the introgression, HMS 1 was confirmed (genotype C). Obviously, HMS 2 alone is not enough to cause sterility (C–G in Figure 4B).

A comparison between the two introgressions 28.3 and 44.5 is also informative. In the analysis of 44.5, it appears that both HMS 1 and HMS 2 are needed for quasi-sterility. Yet genotype A of 28.3 (Figure 4A), hosting HMS 1 but clearly not HMS 2, is quasi-sterile. This observation suggests an HMS factor(s) to the left of *Rga* ("?" in Figure 4A). If a factor is located between *rdgC* and *Edg78*, it appears to have a dominant effect when combined with HMS 1 (F, G, and H in Figure 4A). Still other evidence discussed below indicates that the factor "?" in Figure 4A may consist of more than two HMS factors, including HMS 9.

Introgression 32.4: Genotypes A, B, C, D, and G were all quasi-sterile, but significant fertility differences were detected from between-group comparisons: A *vs.* B ($P_{\bar{o}} = 0.014$) and A *vs.* G ($P_{\bar{o}} = 0.024$; Figure 4C). HMS 2 or other factors may contribute to the A *vs.* B difference (also compare to line 44.5, Figure 4B). The obvious shift from quasi-sterile (genotypes C and D) to fertile (genotype E) again supports the localization of HMS 1. On the other end of the introgressed segment, the fact that A and G differ in their fertility suggests that there might be some factor(s) to the left of *5-Ht2* (marked by "?"). However, it is unclear on which side of the centromere this factor(s) would be located. We therefore do not declare an HMS locus on 3R here.

Below we show that a factor (HMS 9) is indicated between HMS 1 and *5-Ht2*. As in the case of 28.3 (Figure 4A), the observations made in Figure 4C are consistent with this interpretation. Without another factor in this region, genotypes B, C, and D in Figure 4C, which clearly are not homozygous for HMS 2, would be fertile because we noted earlier that HMS 1 alone cannot cause quasi-sterility (see also A.3 and B in Figure 4B).

Similar mapping analyses from crosses involving seven other lines are detailed in the APPENDIX and the evidence from all 10 lines suggesting a total of eight HMS loci is summarized in Table 1.

Are the eight HMS loci listed in Table 1 sufficient to account for male sterility observed in all introgression lines covering 3R? To answer this question, each of the introgression lines was examined for the presence or absence of the putative HMS loci. Lines informative for this purpose are displayed in Figure 2. Most of the sterility detected on 3R can be attributed to the eight loci. However, several lines seem to require at least two more HMS loci to account for the male sterility observed.

In the region around the centromere, at least one HMS factor seems to be in the *Pka-Rga* region as suggested in the legend of Figure 4, A and C. Additional evidence comes from comparisons among the lines 45.7, 45.8, 45.9, and 45.10. These lines were derived as recombinants from the same progenitor, and their right ends are very likely the same. An HMS factor ("9" in Figure 2) is implicated on the basis of the fact that line 45.10 was more fertile (16 progeny per male) than line 45.8 (4 progeny, $P_{\bar{o}} < 0.001$).

To account for the sterility of lines 33.9, 33.10, and 45.9 and the quasi-sterility of lines 28.3, 32.4, 45.8, 29.5, 29.6, 32.9, and 29.7, another HMS factor ("10" in Figure 2) is implicated. This factor should fall within the region from the right end of line 27.2 (or 32.2 and 32.10) to the right end of the just-mentioned sterile or quasi-sterile lines.

An HMS equivalent consists of two or three minor HMS loci: Figure 2 shows three regions on 3R (labeled S1, S2, and S3) that each can cause complete male sterility. Each region is by definition one "HMS equivalent" (TAO *et al.* 2003, this issue). Here we elaborate on the HMS loci contained within each region.

The S1 region contains HMS loci 1, 9, and 10. A combination of all three loci can account for almost all sterile or quasi-sterile lines, with one or two exceptions. HMS 9 and 1 together cannot cause significant sterility, while HMS 1 and 10 together may render a male subfertile (line 45.10). However, it is unclear what causes the fertility difference between the sterile and quasi-sterile lines covering the S1 region. Furthermore, line 45.5 is rather strange because it covers the whole S1 region and HMS 2, yet it is still not fully sterile (line 29.1 could also contain the whole S1 region but its right end may not contain HMS 10). Two *ad hoc* explanations are provided here to explain these observations, but further experimental evidence is needed to clarify these issues.

First, genetic background variation could underlie the fertility variations among different lines. Although these introgression lines were constructed deliberately in a genetic background as uniformly as possible, some variation cannot be avoided (TAO *et al.* 2003, this issue). Spontaneous mutations during line construction also may affect fertility.

Second, it is possible that some *D. mauritiana* alleles act antagonistically, rather than synergistically, in causing sterility. In other words, two HMS factors can suppress each other. For example, the line 45.5 evidently covers the whole length of 45.9, yet the latter is fully sterile. It is possible that some loci on both ends of 45.5,

not covered by 45.9, could be this kind of antagonistic gene. For example, it is possible that HMS 9 and 2 could suppress one another (lines 33.5 and 45.5). This notion of antagonistic interactions is supported by QTL mapping (see below).

In the S2 region, HMS 2 and 3 together can cause complete sterility (Figure 2, lines 41.3 and 41.5). A notable aspect of HMS 3 (called *tmy*) is that no fertile introgression lines ever cover it. Indeed, *tmy* is the only HMS locus that may contain a "major gene" effect. No other locus can cause significant sterility on its own, because the third chromosome is fully covered by fertile homozygous introgressions except in the region of *tmy* (Tao *et al.* 2003, this issue). On the other hand, HMS 2 (*broadie*) enhances the sterilizing effect of *tmy*, rendering any *tmy* males completely sterile (*e.g.*, lines 41.3 and 41.5 compared to lines 40.7 and 40.9).

Distal to *tmy*, HMS 7 may also enhance the sterilizing effect of *tmy*. For example, line 37.10 is quasi-sterile while line 37.7 is subfertile. The latter may not contain HMS 7. It is unclear, however, whether HMS 8 has an enhancing effect because the mapping of HMS 8 is too rough to make any reliable inference.

In region S3, HMS 8, 7, and 5 together can cause complete sterility (Figure 2, lines 46.17, 46.7, 35.1, and 35.2). The effects of HMS 8 and 7 can be demonstrated best by lines 46.4, 46.15, and 46.9, as shown in Figures A4–A6. In the S3 region, if HMS 8 is not included in the introgression, the fertility will recover a little but still be quasi-sterile (*e.g.*, line 46.15 in Figure A4). If both HMS 7 and 8 are not included, the fertility will recover substantially (Figure A6, line 46.4).

The other two factors, HMS 4 and 6, have no apparent effect by themselves (Figure 2, 46.2, 49.11, etc.). If combined with HMS 5, they may cause significant fertility reduction but not full sterility (Figure 2, lines 46.4, 46.3, and 46.19).

QTL of hybrid male fertility: An MIM procedure was used to carry out two analyses for identifying QTL in a population of introgression genotypes (Figure 3). In one case, 5025 males were assigned a binary trait classification (sterile or fertile), which was treated as a threshold trait with an underlying continuous fertility variable. In the other case, the offspring counts from a subsample of 2602 nonsterile males were treated as a continuous variable for QTL mapping. We use QTL_x and QTL_y to represent the results obtained from these two analyses, respectively (Figure 5). The additive, dominance, and epistatic effects for 19 QTL_x are summarized in Tables 2 and 3 and the corresponding effects for 18 QTL_y are given in Tables 4 and 5. Although variances of the estimates were not obtained in this analysis, we believe they are large because each estimate tends to depend on a small number of introgression lines. Therefore, these results provide only a rough estimate of the genetic architecture of hybrid male fertility. Nevertheless, some general patterns can be inferred from Tables 2–5.

First, the *D. mauritiana* allele (M) does not always have a negative effect on fertility (*i.e.*, α or $a < 0$) as might be expected for hybrid incompatibility factors. In fact, the *mauritiana* allele has a positive effect in 5 of 19 QTL for the binary sterility trait and in 8 of 18 QTL for the continuous fertility trait. This could mean that not all QTL detected are due to hybrid incompatibilities, but may be due to polymorphism of deleterious alleles that would have similar effects in both homo- and heterospecific genetic backgrounds.

Second, the dominance relationships between the M and S alleles are mixed, having no apparent trend. There are 8 QTL_x and 8 QTL_y, where the M allele is dominant over the S allele [*i.e.*, $\alpha(a)$ and $\delta(d)$ have the same sign], and 11 QTL_x and 10 QTL_y, where the M allele is recessive to the S allele [*i.e.*, $\alpha(a)$ and $\delta(d)$ have the opposite sign] (Tables 2 and 4). Furthermore, overdominance of the M allele was estimated for 5 QTL_x and 5 QTL_y [$|\delta| < |\alpha|$ or $|d| < |a|$, $\alpha(a)$ and $\delta(d)$ have the same sign], while underdominance was estimated for 3 QTL_x and 10 QTL_y [$|\delta| < |\alpha|$ or $|d| < |a|$, $\alpha(a)$ and $\delta(d)$ have the opposite sign]. These results may suggest complex interactions among loci causing hybrid male sterility.

Third, epistatic effects are pervasive (Tables 3 and 5). Ten out of 24 significant epistatic interactions for QTL_x (and 5 out of 18 for QTL_y) are positive. The signs of the epistatic effects cannot be predicted from the signs of the two effects [$\alpha(a)$ and $\delta(d)$] involved. These observations suggest both synergistic and antagonistic relationships among the loci causing hybrid male sterility.

Figure 5 shows the LOD profiles for QTL mapping of the two traits. There are several points of correspondence, but the high density of QTL, the pervasiveness of epistasis, and the large sampling variances of estimates make it difficult to identify one-to-one relationships with high confidence. Nevertheless, it is clear that each type of analysis identifies a similar number of QTL within

Figure 5.—QTL of hybrid male sterility on the third chromosome. (A) As a comparison, a summary of the 10 HMS loci mapped on 3R by qualitative genetic mapping is also shown. (B) A total of 19 QTL_x on the third chromosome are implicated by MIM from the complete sample of 5025 males, which are coded as either sterile (0) or fertile (1). (C) A total of 18 QTL_y are implicated from the subsample of 2602 nonsterile males. For each putative QTL, the likelihood profile is represented by a numbered curve. Note that the scale of the LOD score is adjusted for QTL_x 13 in B. The centromere is indicated by a solid oval. ASO markers are shown with their positions either on a genetic map (I), which is obtained from the introgression background in this study, or on a cytological map (II) in polytene bands.

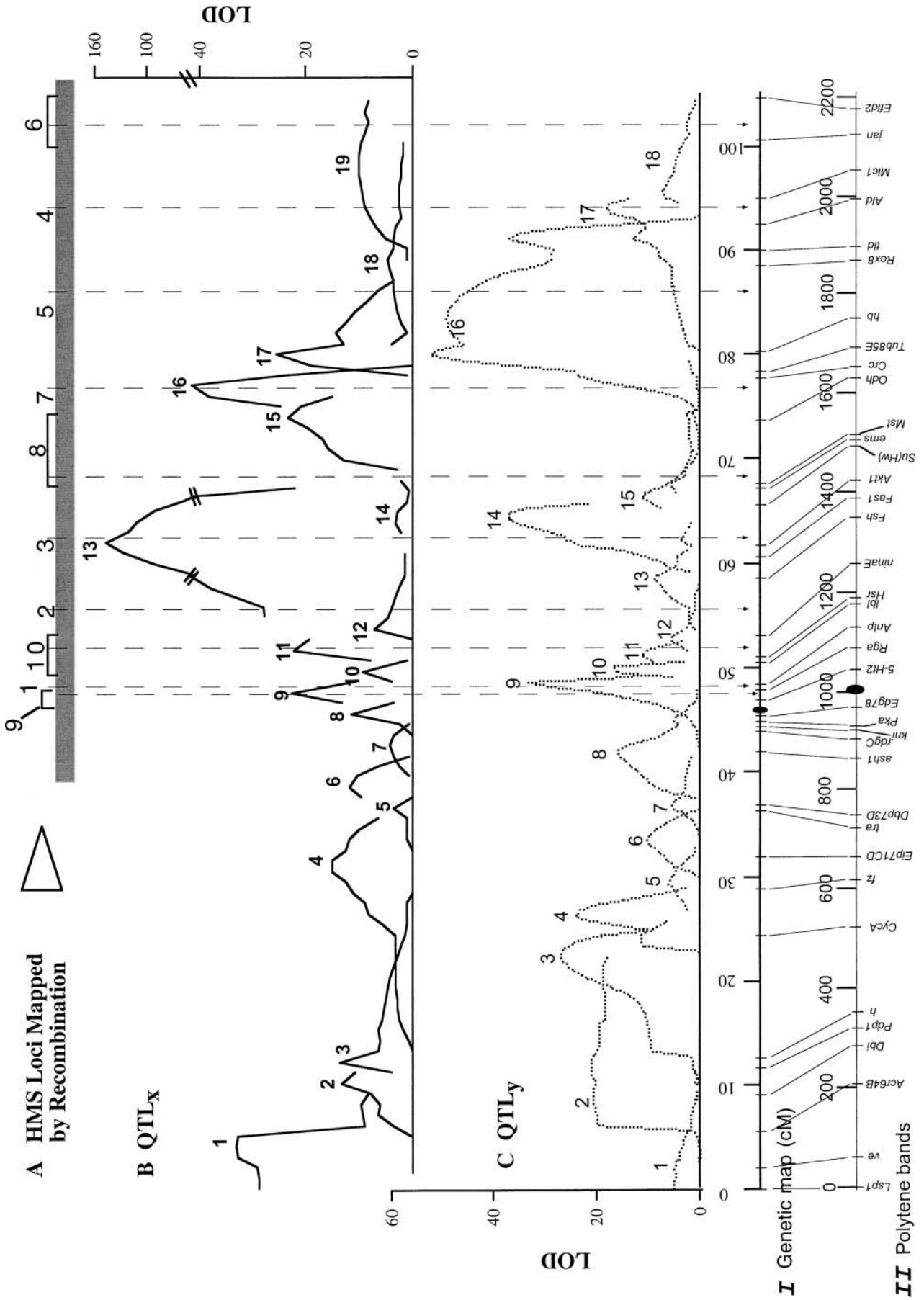

A HMS Loci Mapped by Recombination
B QTLx
C QTLy
LOD
I Genetic map (cM)
II Polytene bands

TABLE 2

Estimates of QTL_x positions and effects

QTL_x	Position (cM)	Additive effect (α)	Dominance effect (δ)
1	4	−1.241	−1.604
2	10	1.820	−1.103
3	12	−1.578	1.135
4	30	−0.969	0.490
5	36	0.727	−0.555
6	38	−1.493	0.793
7	42	0.542	0.649
8	45	−0.321	1.094
9	47	−0.971	0.411
10	49	0.051	1.387
11	51	−0.910	−1.803
12	53	−0.062	1.024
13	61	−1.696	1.705
14	67	−0.028	−0.774
15	73	−1.596	−0.935
16	76	1.308	2.233
17	79	−1.640	1.048
18	87	−0.634	−0.077
19	100	−0.678	0.206

TABLE 3

Estimates of epistasis between QTL_x pairs

No.	QTL_x1	QTL_x2	Type of epistasis	Effect
1	3	4	α × δ	0.672
2	4	5	α × α	−0.972
3	5	8	α × α	−0.176
4	5	7	α × δ	−0.477
5	6	8	α × α	−0.527
6	7	9	δ × δ	−0.067
7	8	9	α × δ	0.898
8	8	11	α × α	−0.177
9	8	10	δ × α	−0.198
10	9	11	α × α	−1.053
11	9	11	δ × α	1.061
12	9	12	δ × α	−0.362
13	10	12	δ × α	−0.710
14	10	13	δ × α	0.636
15	11	14	α × α	−1.318
16	13	14	α × α	0.457
17	13	14	α × δ	0.371
18	13	14	δ × δ	0.090
19	14	16	α × α	−0.897
20	15	16	α × δ	0.756
21	15	17	α × δ	0.211
22	16	18	α × α	−1.433
23	16	19	α × α	−0.310
24	17	18	α × α	0.466

the same regions and both analyses show that hybrid male sterility in these species is highly polygenic.

In the qualitative trait mapping, we implicitly assumed that the *D. mauritiana* allele at each locus confers a negative effect on fertility. However, we also mentioned an indication of antagonistic effects in the region around the centromere where longer introgressions such as 45.5 and 29.1 were actually more fertile than shorter ones such as 45.9 and 29.7 (Figure 2). The QTL mapping results are consistent with this indication, because QTL_x 7 and 10 and QTL_y 8 and 12 (near the centromere) have negative additive effects (*i.e.*, positive effects on fertility of the *D. mauritiana* allele).

In another comparison between qualitative and QTL mapping, there is a puzzling disagreement between QTL_x 16 and HMS 7, which occur at approximately the same position. HMS 7 was identified as a negative *M* effect by comparing lines 46.15 and 46.4 (Figures A4 and A6). However, QTL_x 16 has a very large positive α—strongly suggesting that its *M* allele has a fertilizing effect. How to reconcile this contradiction? The additive effect of QTL_x 16 has three significant negative interactions with other QTL_x effects (Table 3). If we calculate the net effect of QTL_x 16 as the value of α plus half of each of these three epistatic effects, the value is actually −0.012, suggesting a sterilizing *M* allele. The interpretation of the additive effect is not straightforward when epistatic effects are present. This conclusion may apply to several other QTL with positive α or *a* estimates.

Taking 19 as the number of HMS loci on the third chromosome, the density of HMS loci per polytene band is ~40% of that on the *X* chromosome, where 9 HMS loci have been mapped on ~40% of that chromosome (reviewed by Wu and Hollocher 1998; Tao *et al.* 2003, this issue). This density ratio is almost the same as what we estimated using a different method (Tao *et al.* 2003). The number (three) of "HMS equivalents" (*i.e.*, S1, S2, and S3) identified on 3R is also consistent with the previous genome-wide estimate of ~15 such equivalents (Tao *et al.* 2003). Thus, although detailed information for each of the HMS loci is still far from adequate, the overall view of the number and distribution of the HMS loci on the third chromosome appears to be reasonably reliable.

DISCUSSION

The genetic architecture of hybrid male sterility: *D. simulans* and *D. mauritiana* are geographically isolated and diverged from a common ancestor ~0.3 million years ago (Kliman *et al.* 2000). These two populations appear to represent a classic case of allopatric speciation, and hybrids produced in the laboratory show complete male sterility. The results of the genetic analysis reported here, along with previous studies of hybrid male sterility in this species pair, clearly show a highly polygenic genetic architecture, which is expected under the neo-Darwinian view of allopatric speciation (as noted in the Introduction). Our study also shows that most HMS loci have relatively small effects, so that sterility requires the effects of at least two or three different

TABLE 4

Estimates of QTL_y positions and effects

QTL_y	Position (cM)	Additive effect (a)	Dominance effect (d)
1	2	2.46	−11.06
2	12	1.90	−28.22
3	24	−18.22	−8.21
4	26	5.60	2.80
5	29	16.28	23.19
6	33	−19.91	34.76
7	37	−12.64	−23.82
8	42	1.65	29.55
9	48	−20.73	25.41
10	50	−11.32	10.54
11	51	−16.35	−15.47
12	53	21.65	−15.12
13	58	−9.10	35.87
14	64	−59.06	26.61
15	66	27.14	−46.47
16	79	−39.44	38.66
17	93	11.81	70.24
18	95	−15.55	−58.81

TABLE 5

Estimates of epistasis between QTL_y pairs

No.	QTL_y1	QTL_y2	Type of epistasis	Effect
1	2	3	$a \times a$	−17.40
2	2	5	$a \times a$	20.41
3	3	4	$d \times d$	−89.01
4	4	5	$d \times d$	−9.66
5	5	7	$a \times a$	−6.58
6	7	8	$d \times d$	−55.83
7	8	11	$a \times a$	−2.71
8	9	10	$a \times a$	−23.11
9	9	10	$d \times a$	−17.30
10	9	12	$d \times a$	35.07
11	10	13	$d \times d$	−19.03
12	12	13	$d \times d$	−11.54
13	13	16	$a \times d$	−24.32
14	13	16	$d \times d$	37.91
15	15	16	$d \times a$	−34.55
16	16	17	$a \times a$	6.01
17	16	17	$d \times a$	1.21
18	16	18	$a \times a$	−16.91

factors. Only one small region of the third chromosome (HMS 3, *tmy*) is capable of producing a high level of sterility in isolation from other *D. mauritiana* alleles. This result suggests that speciation may require the accumulation of many allelic fixations in isolated populations with each fixation producing incompatibilities of small effect.

This polygenic view of reproductive isolation does not preclude an occasional incompatibility of large effect due to a single pair of allelic fixations (Orr and Coyne 1992; Orr 1998), but these seem to be rare. One possible example is the rescue by a fourth-chromosome segment of the male sterility caused by introgressing the *D. arizonae Y* chromosome into an otherwise pure *D. mojavensis* background (Pantazidis *et al.* 1993). However, other cases that originally seemed to be major factors of HMS turned out to be a cluster of minor factors after more genetic analysis was done (Coyne and Charlesworth 1986, 1989; Perez *et al.* 1993; Perez and Wu 1995).

In any case, the magnitude of effect of an incompatibility in hybrids is not necessarily related directly to the magnitude of effect on the primary trait within diverging subpopulations. For example, it is possible that different mutations fixed by selection in different subpopulations may have large effects on a primary trait within those subpopulations, but no incompatibility at all in hybrids. Similarly, it is possible that mutations of small positive effect on a trait within subpopulations could have very negative interactions in hybrids, producing an incompatibility of major effect.

In the study reported here, the QTL approach to genetic analysis of incompatibilities has provided strong evidence for complex patterns of epistasis. Such a result may not be surprising, since hybrid incompatibilities are, by definition, negative interactions between alleles at different loci. However, there could be multiple systems of incompatibility that do not interact with each other. For example, an HMS locus A in this study may represent a *D. mauritiana* allele at locus *i* (within an introgression) that has a negative interaction with a *D. simulans* allele at locus *j* (outside of the introgression). Another HMS locus B may consist of a negative interaction between a *D. mauritiana* allele at locus *k* and a *D. simulans* allele at locus *l*. The locus pair *i*, *j* does not necessarily interact with the locus pair *k*, *l* and the effects of the two incompatibilities may be additive with respect to a quantitative measure of fertility.

One view of the genetic basis of hybrid incompatibility is that it is due to numerous loci distributed throughout the genome with effects that are additive and interchangeable (Naveira and Maside 1998). A different view is that several minor HMS loci may each have negligible effect, but act "epistatically" together to sterilize an introgression hybrid (Wu and Hollocher 1998; Orr and Irving 2001). Many, but not all, of the loci studied here show epistasis on a quantitative scale, suggesting that both views have merit.

D. simulans and *D. mauritiana* are very similar morphologically. The only distinctive difference is in the shape of the posterior lobe of the genital arch in males (Liu *et al.* 1996). A QTL analysis of the interspecific difference in lobe shape has shown that it is highly polygenic (like hybrid male sterility), but there is very little evidence for epistasis on a quantitative scale (Zeng *et al.* 2000). Also, there is no evidence for dysfunctional interactions that result in abnormal lobe development in

hybrids. The posterior lobe in F_1 hybrids is well formed and intermediate in shape between the two parental species. This result suggests that many of the allelic fixations that occur in diverging subpopulations do not contribute to hybrid incompatibilities.

The introgression approach to genetic analysis of HMS: The use of introgression in the genetic analysis of quantitative traits has a long history. A classic technique was to isolate a chromosomal region of interest in an isogenic background and then apply conventional mapping (Breese and Mather 1957; Thoday 1961). This strategy continues to be of great value in quantitative trait analysis (*e.g.*, Eshed and Zamir 1995; Frary *et al.* 2000; Fridman *et al.* 2000).

In the analysis of hybrid sterility, the use of introgressions has been a necessity because, in many species pairs, the large number of genes contributing to sterility and the high level of sterility make backcross or F_2 designs unworkable. The study reported here and earlier studies of the species pair *D. simulans* and *D. mauritiana* (reviewed by Wu and Hollocher 1998) have proven the value of the introgression approach. Although most HMS loci have relatively small effects, an individual locus may have full penetrance when conditioned on an appropriate genetic background (*i.e.*, one in which the level of an underlying fertility score is just above the fertile/sterile threshold). In these cases, a quantitative trait may be reduced to a qualitative trait with Mendelian segregation (Ting *et al.* 1998). However, two possible drawbacks to the introgression approach should be mentioned.

First, the fact that the detection of HMS loci is contingent on the genetic background, while sometimes an advantage, can also be a weakness. Some loci may escape detection if the background level of underlying fertility is too low or too high. Also, if complex epistasis is present, there is no way to detect a particular HMS locus if it interacts with other loci only in a particular way or with loci that are not closely linked within an introgressed segment that is being dissected genetically. Therefore, the current estimate of 19 HMS loci on the third chromosome could be considerably lower than the actual number.

Second, chromosomal rearrangements involving introgressed segments may suggest incompatibilities that do not exist. For example, transposition or translocation of a gene in one species, but not the other, could result in different genic content of an introgressed segment relative to the segment it is replacing. This rearrangement could result in the deletion of an essential gene in the introgression genotype relative to both parental species. Although the polytene band maps of *D. simulans* and *D. mauritiana* are nearly identical (Lemeunier and Ashburner 1984), small rearrangements may have occurred. However, it is clear that most of the sterility found in the male F_1 hybrids between these two species cannot be due to this mechanism unless there are many *X*-autosome transpositions. Whole-genome sequencing and the cloning of HMS loci (*e.g.*, Ting *et al.* 1998) may soon reveal more details of the molecular basis of hybrid incompatibilities.

We thank Lynn Stam, Kristen Inoue, Jermaine Fuller, Lixia Jiang, and Jing Zhang for technical support during this study and C. Meiklejohn and J. Blumenstiel for critical reading of various drafts of this article. We also thank two anonymous reviewers for their suggestions on improving this manuscript. Y.T. thanks R. Fehon, D. Kiehart, J. Mercer, M. Rausher, and M. Uyenoyama for guidance. Y.T. was a recipient of a Catherine Stern Dissertation Year Fellowship of Duke University. This work was supported by National Institutes of Health grants 47292 and 65169.

LITERATURE CITED

Breese, E. L., and K. Mather, 1957 The organization of polygenic activity within a chromosome of *Drosophila*. I. Hair characters. Heredity **11:** 373–395.

Cabot, E. L., A. W. Davis, N. A. Johnson and C.-I Wu, 1994 Genetics of reproductive isolation in the *Drosophila simulans* clade: complex epistasis underlying hybrid male sterility. Genetics **137:** 175–189.

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

Coyne, J. A., 1992 Genetics and speciation. Nature **355:** 511–515.

Coyne, J. A., and B. Charlesworth, 1986 Location of an *X*-linked factor causing sterility in male hybrids of *Drosophila simulans* and *D. mauritiana*. Heredity **57:** 243–246.

Coyne, J. A., and B. Charlesworth, 1989 Genetic analysis of *X*-linked sterility in hybrids between three sibling species of *Drosophila*. Heredity **62:** 97–106.

Coyne, J. A., and H. A. Orr, 1998 The evolutionary genetics of speciation. Philos. Trans. R. Soc. Lond. B **353:** 287–305.

Davis, A. W., and C.-I Wu, 1996 The broom of the sorcerer's apprentice: the fine structure of a chromosomal region causing reproductive isolation between two sibling species of Drosophila. Genetics **143:** 1287–1298.

Dobzhansky, Th., 1936 Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. Genetics **21:** 113–135.

Dobzhansky, Th., 1937 *Genetics and the Origin of Species*, Ed. 1. Columbia University Press, New York.

Dobzhansky, Th., 1975 Analysis of incipient reproductive isolation within a species of *Drosophila*. Proc. Natl. Acad. Sci. USA **72:** 3638–3641.

Eshed, Y., and D. Zamir, 1995 An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics **141:** 1147–1162.

Frary, A., T. C. Nesbitt, A. Frary, S. Grandillo, E. van der Knaap *et al.*, 2000 fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. Science **289:** 85–88.

Fridman, E., T. Pleban and D. Zamir, 2000 A recombination hotspot delimits a wild species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proc. Natl. Acad. Sci. USA **97:** 4718–4723.

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

Jiang, C., and Z-B. Zeng, 1997 Mapping quantitative trait loci with dominant and missing markers in various crosses from two inbred lines. Genetica **101:** 47–58.

Kao, C.-H., and Z-B. Zeng, 1997 General formulations for obtaining the MLEs and the asymptotic variance-covariance matrix in mapping quantitative trait loci when using the EM algorithm. Biometrics **53:** 653–665.

Kao, C.-H., and Z-B. Zeng, 2002 Modeling epistasis of quantiative trait loci using Cockerham's model. Genetics **160:** 1243–1261.

Kao, C.-H., Z-B. Zeng and R. D. Teasdale, 1999 Multiple interval mapping for quantitative trait loci. Genetics **152:** 1203–1216.

Kliman, R. M., P. Andolfatto, J. A. Coyne, F. Depaulis, M. Kreitman *et al.*, 2000 The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics **156:** 1913–1931.

Lemeunier, F., and M. Ashburner, 1984 Relationships within the *melanogaster* species subgroup of the genus *Drosophila* (*Sophophora*). IV. The chromosomes of two new species. Chromosoma **89:** 343–351.

Liu, J. J., J. M. Mercer, L. F. Stam, G. C. Gibson, Z-B. Zeng *et al.*, 1996 Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila simulans* and *D. mauritiana*. Genetics **142:** 1129–1145.

Mayr, E., 1942 *Systematics and the Origin of Species*. Columbia University Press, New York.

Muller, H. J., 1942 Isolating mechanisms, evolution, and temperature. Biol. Symp. **6:** 71–125.

Naveira, H. F., and X. R. Maside, 1998 The genetics of hybrid male sterility in *Drosophila*, pp. 330–338 in *Endless Forms: Species and Speciation*, edited by D. J. Howard and S. H. Berlocher. Oxford University Press, New York.

Orr, H. A., 1998 The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution **52:** 935–949.

Orr, H. A., and A. J. Betancourt, 2001 Haldane's sieve and adaptation from the standing genetic variation. Genetics **157:** 875–884.

Orr, H. A., and J. A. Coyne, 1992 The genetics of adaptation: a reassessment. Am. Nat. **140:** 725–742.

Orr, H. A., and S. Irving, 2001 Complex epistasis and the genetic basis of hybrid sterility in the *Drosophila pseudoobscura* Bogata-USA hybridization. Genetics **158:** 1090–1100.

Pantazidis, A. C., V. K. Galanopoulos and E. Zouros, 1993 An autosomal factor from *Drosophila arizonae* restores normal spermatogenesis in *D. mojavensis* males carrying the *arizonae Y* chromosome. Genetics **134:** 309–318.

Perez, D. E., and C.-I Wu, 1995 Further characterizations of the *Odysseus* locus of hybrid sterility in Drosophila: one gene is not enough. Genetics **140:** 201–206.

Perez, D. E., C.-I Wu, N. A. Johnson and M.-L. Wu, 1993 Genetics of reproductive isolation in the *Drosophila simulans* clade: DNA marker-assisted mapping and characterization of a hybrid-male sterility gene, *Odysseus* (*Ods*). Genetics **134:** 261–275.

Tao, Y., D. L. Hartl and C. C. Laurie, 2001 Sex-ratio segregation distortion associated with reproductive isolation in Drosophila. Proc. Natl. Acad. Sci. USA **98:** 13183–13188.

Tao, Y., S. Chen, D. L. Hartl and C. C. Laurie, 2003 Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. I. Differential accumulation of hybrid male sterility effects on the *X* and autosomes. Genetics **164:** 1383–1397.

Thoday, J. M., 1961 Location of polygenes. Nature **191:** 368–370.

Ting, C.-T., S.-C. Tsaur, M.-L. Wu and C.-I Wu, 1998 A rapidly evolving homeobox at the site of a hybrid sterility gene. Science **282:** 1501–1504.

Ting, C.-T., S.-C. Tsaur and C.-I Wu, 2000 The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. Proc. Natl. Acad. Sci. USA **97:** 5313–5316.

True, J. R., J. M. Mercer and C. C. Laurie, 1996a Differences in crossover frequency and distribution among three sibling species of Drosophila. Genetics **142:** 507–523.

True, J. R., B. S. Weir and C. C. Laurie, 1996b A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. Genetics **142:** 819–837.

Wu, C.-I, and H. Hollocher, 1998 Subtle is nature, pp. 339–351 in *Endless Forms: Species and Speciation*, edited by D. J. Howard and S. H. Berlocher. Oxford University Press, New York.

Wu, C.-I, N. Johnson and M. F. Palopoli, 1996 Haldane's rule and its legacy: Why are there so many sterile males? Trends Ecol. Evol. **11:** 281–284.

Xu, S. Z., and W. R. Atchley, 1996 Mapping quantitative trait loci for complex binary diseases using line crosses. Genetics **143:** 1417–1424.

Zeng, Z-B., C.-H. Kao and C. J. Basten, 1999 Estimating the genetic architecture of quantitative traits. Genet. Res. **74:** 279–289.

Zeng, Z-B., J. Liu, L. F. Stam, C.-H. Kao, J. M. Mercer *et al.*, 2000 Genetic architecture of a morphological shape difference between two Drosophila species. Genetics **154:** 299–310.

Communicating editor: M. A. F. Noor

APPENDIX

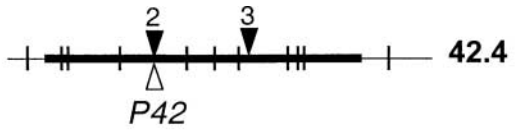
Figures A1–A7 provide additional details about the qualitative mapping of HMS loci on chromosome 3. Each figure describes the results of recombination analysis of a particular introgression. Also see the legend of Figure 4 for explanation of symbols.

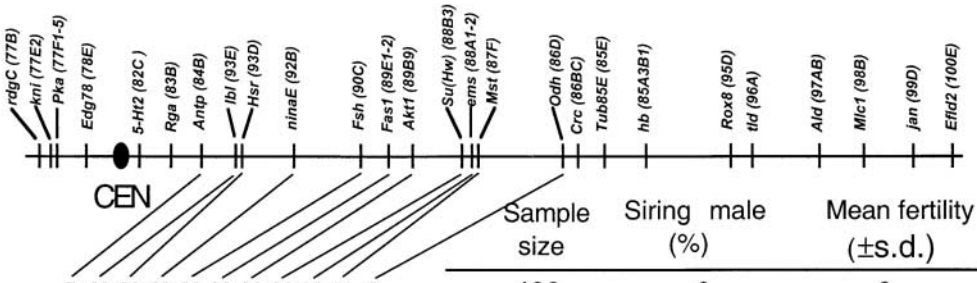
Sterile introgression 42.4 is summarized in Figure A1. One recombinant in the *Akt1* and *Su(Hw)* interval was fertile, indicating HMS factor 3, which was named *too much yin* (*tmy)* and mapped to a region of $< \sim 80$ kb (Tao *et al.* 2001). We use *tmy* and *Tmy* to denote the *D. mauritiana* and the *D. simulans* alleles, respectively. Homozygous *tmy* males are always semisterile (quasi-sterile or subfertile) and sire female-biased ($\sim$75%) progeny (Tao *et al.* 2001). It is the most significant HMS factor on the whole third chromosome (Tao *et al.* 2003, this issue). Although fertility shift seems to be within genotype E, it is not adequate to support the existence of an HMS factor between *Fas1* and *Akt1* because it is based on a single sterile male. This interpretation is consistent with the observations that the average fertility of E was similar to F ($P_{\bar{o}} = 0.873$) or G ($P_{\bar{o}} = 0.849$) and D.2 was very fertile. The position of *P42* relative to HMS factor 2, which was implicated in Figure 4B and previously named *broadie* (Tao *et al.* 2001), has not been determined.

The sterile introgression 40.8 (Figure A2) also provides evidence for HMS 3 (genotype D). Males homozygous for *tmy* alone are semi-sterile, but fully sterile if coupled with *broadie* (Tao *et al.* 2001). HMS factor 2 (*broadie*) acts dominantly with *tmy* to render a male fully sterile (genotypes F, G, and H). Marker *Hsr* is not informative here because the *D. mauritiana* allele for this line is the same as that of *D. simulans*.

Line 37.10 is summarized in Figure A3. Also shown is the proportion of females (k) in the offspring. The 245 males in A.1 were not genotyped but are assumed to be of parental genotype (A.1 and A.2). HMS factor 3 (*tmy*) is implicated once again with fertility and sex ratio shifts within genotype D ($P_{\bar{o}} = 0.04$ for comparing fertility in D.1 and D.2). D.1 is homozygous for *tmy* while D.2 is heterozygous. At the other end of the introgression, it appears that recombination to the right of *P37* may reduce sterility because only 14.9% of all males in A sired any progeny whereas the three individuals of B and C were all fertile. Some sterility factor(s) might be in the region to the right of *P37* (marked as "?"), but the data here are not adequate to clarify this point. Indeed, later evidence supports at least one HMS factor in the *Odh-Crc* region (see Figures A4 and A6).

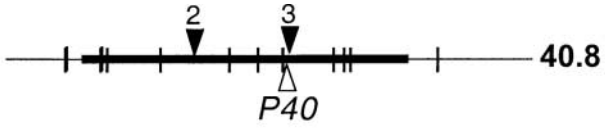
Introgression 46.15 is summarized in Figure A4. It is

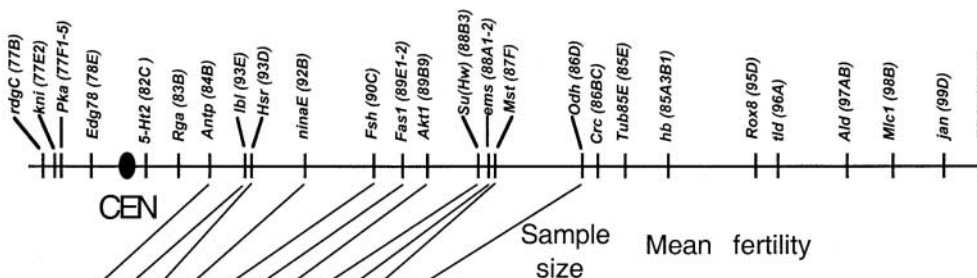




		Sample size	Siring male (%)	Mean fertility (±s.d.)
A.	S M M M M M M M M M S	428	0	0
B.	S M M M M M M M M H S	3	0	0
C.	S M M M M M M M H H S	8	0	0
D.1	S M M M M M M H H H S	15	0	0 *③
D.2	S M M M M M M H H H S	1		95*③
E.1	S M M M M M H H H H S	1		0
E.2	S M M M M M H H H H S	1		14
E.3	S M M M M M H H H H S	1		12
E.4	S M M M M M H H H H S	1		75
F.	S M M M M H H H H H S	7	100	51±46
G.	S H M M H H H H H H S	13	85	51±49
H.	S H M M M M M M M M S	2	0	0
I.	S H H M M M M M M M S	7	0	0
J.	S H H H M M M M M M S	13	0	0

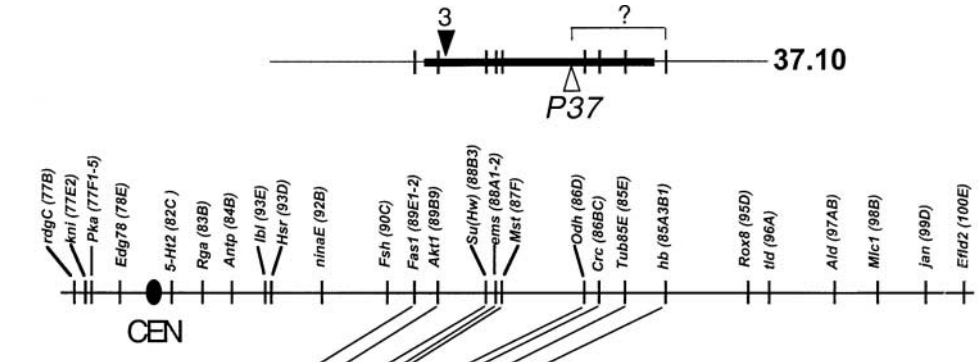
FIGURE A1.—Sterile introgression 42.4.





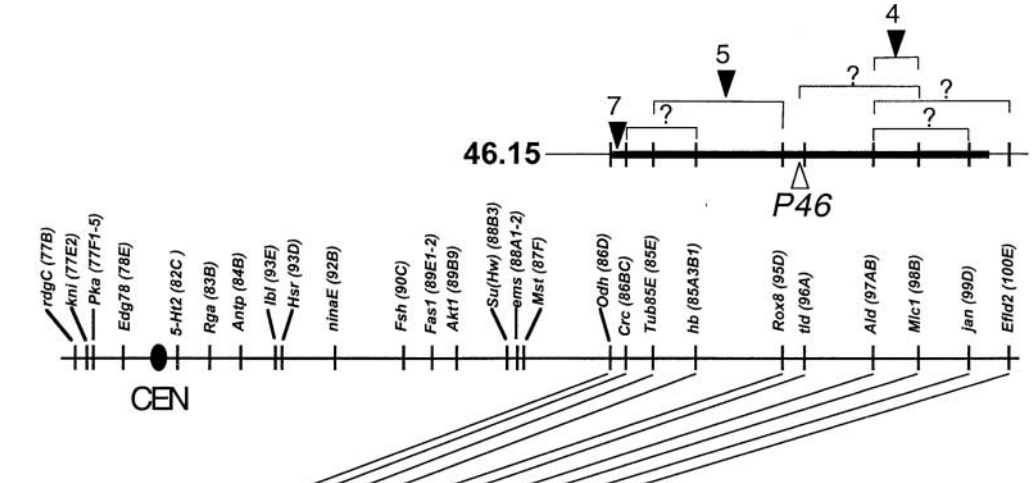
		Sample size	Mean fertility	
A.	S M - M M M M M M M S	275	0	
B.	S M - M M M M M M H S	1	0	
C.	S M - M M M M M H H S	5	0	
D.1	S M - M M M M H H H S	15	0	*③
D.2	S M - M M M M H H H S	1	9	*③
D.3	S M - M M M M H H H S	1	100	*③
E.	S H - M M M M M M M S	10	0	
F.	S H - H M M M M M M S	29	0	
G.	S H - H H M M M M M S	10	0	
H.	S H - H H H H M M M S	1	0	

FIGURE A2.—Sterile introgression 40.8.



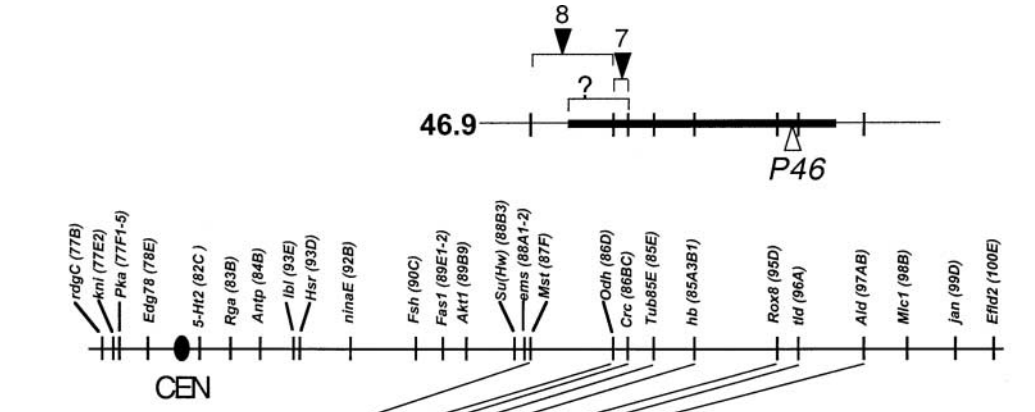
		Sample size	Siring male (%)	Mean fertility (±s.d.)	k	
A.1	S M M M M M M M S	245	14.9	2±7	0.763	
A.2	S M M M M M M M S	43				
B.1	S M M M M M H H S	1		30	0.833	
B.2	S M M M M M H H S	1		113	0.747	
C.	S M M M M H H H S	1		18	0.722	
D.1	S H M M M M M M S	4	100	6±4	0.720	*③
D.2	S H M M M M M M S	3	100	57±24	0.482	*③
E.	S H H M M M M M S	5	100	85±39	0.508	
F.	S H H H H M M M S	8	100	60±44	0.535	

FIGURE A3.—Line 37.10.



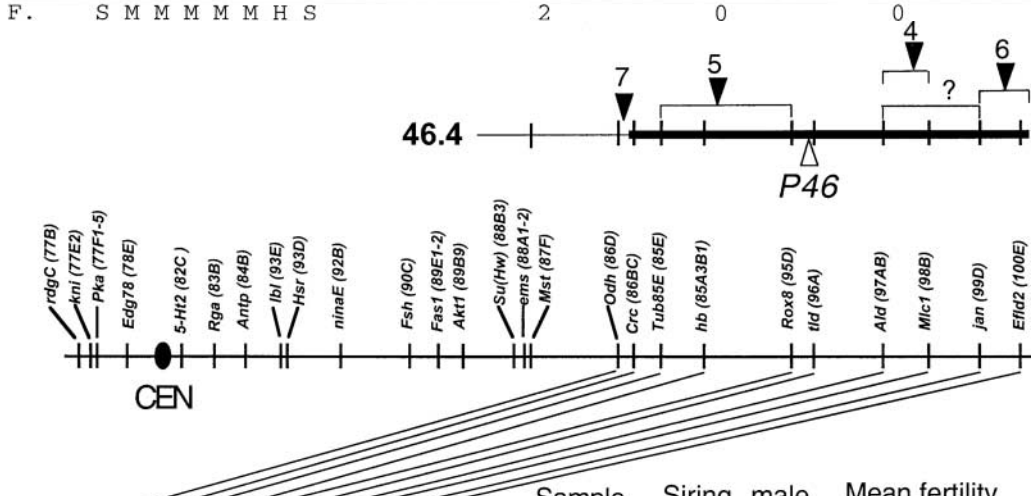
		Sample size	Siring male (%)	Mean fertility (±s.d.)	
A.	S M M M M M M M M S	367	21	2±6	*④
B.	S M M M M M M M H S	28	32	3±7	*④
C.	S M M M M M M H H S	10	70	11±13	*④
D.	S M M M M M H H H S	10	90	19±18	*④
E.	S M M M M H H H H S	4	75	36±30	*④
F.	S H M M M M M M M S	2	0	0	
G.	S H H M M M M M M S	13	92	33±29	*⑤
H.	S H H H M M M M M S	27	96	73±41	*⑤
I.	S H H H H M M M M S	1		62	
J.	S H H H M M M M H S	2	100	79±5	
K.	S H H H M M M H H S	1		33	
L.	S H H H H M M M M S	3	100	95±24	

FIGURE A4.—Introgression 46.15.



		Sample size	Siring male (%)	Mean fertility (±s.d.)	
A.	S M M M M M M S	398	1	1±1	*⑧
B.1	S H M M M M M S	5	0	0	*⑧
B.2	S H M M M M M S	1		1	*⑧
B.3	S H M M M M M S	1		2	*⑧
B.4	S H M M M M M S	1		4	*⑧
B.5	S H M M M M M S	1		5	*⑧
B.6	S H M M M M M S	1		59	*⑧
C.	S H H H M M M S	8	88	22±17	
D.	S H H H H M M S	27	89	40±41	
E.	S H H H H H M S	3	67	37±34	
F.	S M M M M M H S	2	0	0	

FIGURE A5.—Introgression 46.9.



		Sample size	Siring male (%)	Mean fertility (±s.d.)	
A.	S M M M M M M M M M	334	90	21±22	*⑥
B.	S M M M M M M M M H	23	96	34±28	*⑥
C.	S M M M M M M M H H	29	100	43±32	*④
D.	S M M M M M M H H H	9	100	94±22	*④
E.	S M M M M M H H H H	16	94	75±32	
F.	S M M M M H H H H H	5	100	69±27	
G.	S H M M M M M M M M	3	100	16±15	
H.	S H H M M M M M M M	8	100	44±28	*⑤
I.	S H H H M M M M M M	28	100	65±28	*⑤
J.	S H H H H M M M M M	1		70	
K.	S H H H M M M H H H	1		97	

FIGURE A6.—Introgression 46.4.

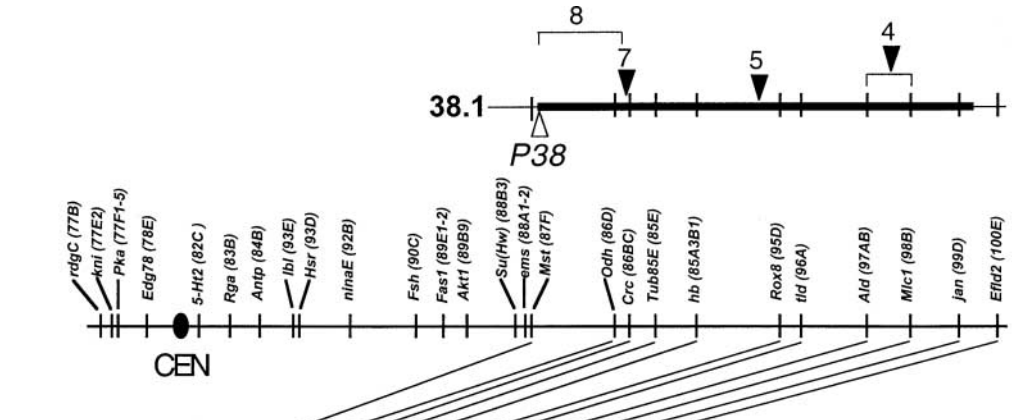


FIGURE A7.—Introgression 38.1.

		Sample size	Siring male (%)	Mean fertility (±s.d.)
A.	S M M M M M M M M M S	337	0	0
B.	S M M M M M M M M H S	26	0	0
C.	S M M M M M M M H H S	17	0	0
D.	S M M M M M M H H H S	12	0	0
E.	S M M M M M H H H H S	10	0	0
F.	S M M M M H H H H H S	53	30	9±22 *⑤
G.	S M M M H H H H H H S	10	60	6±11
H.	S M H H H H H H H H S	13	92	66±42
I.	S H H H H H H H H H S	25	96	62±36

best to examine this and two other introgressions, 46.9 (Figure A5) and 46.4 (Figure A6), together. At the right end of introgression line 46.15 (Figure A4), there was a significant increase in fertility from genotypes B to C ($P_{\bar{o}} = 0.03$; a shift from A to C is also significant with $P_{\bar{o}} = 0.002$). A sterility factor between *Ald* and *jan* is suggested. The mean fertility between C and E was different ($P_{\bar{o}} = 0.03$), implying an HMS factor in the region *P46-Mlc1*. These three regions that may contain HMS factor(s) are marked with "?" to the right of *P46* in Figure A4. It is clear that males were progressively more fertile when larger regions distal to *P46* were made heterozygous (A–E in Figure A4). This suggests that several HMS factors that dispersed distal to *P46* are responsible. Alternatively, a simpler explanation would be that one factor between *Ald* and *Mlc1* ("4") is responsible for the fertility change observed. This latter explanation is supported by more mapping data (see Figure A6).

At the left end of introgression 46.15, two phenotypic shifts can be detected. Compared to F, G became fairly fertile, implying an HMS factor in the region of *Crc* to *hb* ("?" to the left of *P46* in Figure A4). However, only two individuals were in F, so this interpretation may not be reliable. There was also a significant increase in fertility from G to H ($P_{\bar{o}} = 0.002$), implying another HMS factor between *Tub85E* and *Rox8* ("5" in Figure A4). HMS factor 7 is inferred by comparing this mapping with that of line 46.4 (see Figure A6).

Introgression 46.9 is summarized in Figure A5. This line is almost fully sterile and is definitely more sterile than line 46.15, as shown by comparing genotype A in Figures A4 and A5 ($P_{\bar{o}} < 0.001$). Since the introgression of line 46.15 is inclusive of that of line 46.9 at the right end, the sterility must be due to factor(s) at the left end and must be proximal to *Crc*. This region ("?" in Figure A5) is responsible for the phenotypic shift from A to B, where B had more fertile males ($P_{\bar{o}} < 0.001$; Figure A5). A parsimonious explanation is that it is HMS 8, narrowed down to the *Mst-Odh* region, as suggested by comparing this mapping with the two in Figures A4 and A6.

Introgression 46.4 is summarized in Figure A6. Two significant fertility increases were at the right end: from A to B ($P_{\bar{o}} = 0.004$) and from C to D ($P_{\bar{o}} < 0.001$). Two factors may explain the fertility changes. One is between *jan* and the telomere of 3R ("6" in Figure A6), and the other is between *Ald* and *jan* ("?" in Figure A6). The latter factor is most parsimoniously inferred to be HMS 4, which was previously implicated in Figure A4. Proximal to *P46*, HMS 5 is implicated again in a significant shift in mean fertility from H to I ($P_{\bar{o}} = 0.04$), but there is insufficient information to localize it precisely.

It is interesting to compare the mapping data from the three introgressions, 46.4, 46.9, and 46.15 (Figures A4–A6). First, a comparison of genotype A in Figures A4 and A6 ($P_{\bar{o}} < 0.001$) suggests that line 46.15 hosts a factor that line 46.4 does not (HMS 7). However, it is unlikely that line 46.15 hosts HMS 6 at the right end. Second, a combination of HMS factors 4, 5, and 7 can cause severe sterility (Figure A4, A in 46.15), yet 46.9 is more sterile than 46.15 even though 46.9 does not host HMS factors 4 and 6 (Figure A5, A *vs.* Figure A4, A; $P_{\bar{o}} < 0.001$). This phenotypic difference suggests that at least one more HMS factor ("8" in Figure A5) must be proximal to HMS factor 7 but distal to *Mst*, because the introgressed *D. mauritiana* segment in 46.9 does not extend beyond this point.

Introgression 38.1 is summarized in Figure A7. The mapping data for this introgression are consistent with previous analyses (Figures A4–A6). As in the case of 46.9, a combination of HMS factors 4, 5, 7, and 8 confers full sterility (*e.g.*, E in Figure A7). Recombinants implicate an HMS factor ("5"), which is localized to the region between *hb* and *Rox8*.