Copyright © 2004 by the Genetics Society of America

The Genetics of Inviability and Male Sterility in Hybrids Between *Anopheles gambiae* and *An. arabiensis*

M. Slotman,*,[1] A. della Torre[†] and J. R. Powell*

**Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06511-7444 and [†]Instituto di Parassitologia, Fondazione Pasteur-Conci Bolognetti, Università di Roma "La Sapienza," Rome, Italy*

Manuscript received July 16, 2003
Accepted for publication January 22, 2004

ABSTRACT

Male hybrids between *Anopheles gambiae* and *An. arabiensis* suffer from hybrid sterility, and inviability effects are sometimes present as well. We examined the genetic basis of these reproductive barriers between the two species, using 21 microsatellite markers. Generally, recessive inviability effects were found on the *X* chromosome of *gambiae* that are incompatible with at least one factor on each *arabiensis* autosome. Inviability is complete when the *gambiae* and *arabiensis* inviability factors are hemi- or homozygous. Using a QTL mapping approach, regions that contribute to male hybrid sterility were also identified. The *X* chromosome has a disproportionately large effect on male hybrid sterility. Additionally, several moderate-to-large autosomal QTL were found in both species. The effect of these autosomal QTL is contingent upon the presence of an *X* chromosome from the other species. Substantial regions of the autosomes do not contribute markedly to male hybrid sterility. Finally, no evidence for epistatic interactions between conspecific sterility loci was found.

THE analysis of the genetics of inviability and sterility of hybrids of closely related species has focused primarily on Drosophila. Although much progress has been made in this field over the last two decades, the generality of the conclusions that have been drawn needs to be tested in other taxa. The *Anopheles gambiae* complex represents a very interesting model in this respect, as it includes seven currently recognized sibling species, several of which hybridize in nature. Most of the species in this complex obey Haldane's rule, which says that sterility is found in hybrids of the heterogametic sex, whereas the homogametic hybrids are fertile (Haldane 1922). Moreover, due to their wide distribution, high vector competence, and their preference for human blood meals, two members of the complex, *An. gambiae s.s.* and *An. arabiensis*, are among the most important malaria vectors in Africa. Both species are widespread in sub-Saharan Africa and are sympatric over much of their range. They are found in the same localities and are sometimes present in the same mating swarms (Marchand 1984). Nonetheless, hybrids between these two species occur at very low frequency (<0.1%; Coluzzi *et al.* 1979; Petrarca *et al.* 1991; Touré *et al.* 1998; Tripet *et al.* 2001). Therefore, both pre- and postmating isolation mechanisms are incomplete, leaving open the possibility of introgression between these two species (della Torre *et al.* 1997). This prospect is particularly relevant for malaria vector control programs, since it is likely to affect the spread of insecticide resistance genes in natural populations, as well as the spread of newly introduced genes through the anticipated future release of genetically modified mosquitoes. To understand the possibility of introgression, the genetic architecture of reproductive incompatibilities must be understood. This is because the presence of inviability or sterility genes in a chromosomal region is expected to affect the probability of introgression of that particular region (Rieseberg *et al.* 1999; Noor *et al.* 2001a).

Studies of the genetic architecture of species isolation were initiated by Dobzhansky (1936), who examined chromosomal contributions to sterility in *Drosophila pseudoobscura* and *D. persimilis* hybrids. The origin of sterility genes was originally enigmatic, because it was thought that purifying selection prevented the fixation of genes causing inviability or sterility. However, the process was elucidated by Dobzhansky (1937) and Muller (1940) with the suggestion that incompatibilities between two or more heterospecific genes are responsible. The idea that the Dobzhansky-Muller model of incompatibility is the major cause of hybrid problems, at least in animals, is consistent with the data available so far (Orr 1995).

One much debated issue concerns the number of genetic changes necessary for speciation. Is speciation due to the gradual accumulation of numerous genes of small effect or are few substitutions of major effect responsible? Two much investigated forms of postzygotic reproductive isolation are hybrid inviability and hybrid sterility. On the basis of studies of Drosophila, hybrid inviability appears to evolve more slowly than

[1]*Corresponding author:* Department of Entomology, University of California, 1 Shields Ave., Davis, CA 95616.
E-mail: maslotman@ucdavis.edu

male hybrid sterility. As a consequence, typically only a few regions cause inviability in hybrids of closely related species, whereas many regions are involved in male hybrid sterility (Carvajal *et al.* 1996; True *et al.* 1996; Coyne *et al.* 1998; Tao *et al.* 2003a). This suggests that the number of genetic changes required for hybrid inviability is not large. Furthermore, several mutations that rescue inviability of hybrids between *D. melanogaster* and *D. simulans* have been discovered (Hutter 1997). The rescue of inviability by single mutations also suggests that hybrid inviability involves few genes.

Some studies on male hybrid sterility have suggested that few genes of major effect are responsible (Wu and Beckenbach 1983; Coyne 1984; Zouros *et al.* 1988; Orr 1992). However, as pointed out by Davis and Wu (1996), the resolution of these studies may not have been sufficient to rule out a polygenic model. A later study on *D. simulans* and *D. mauritiana* hybrids estimated the number of sterility genes to be ~120 (Wu *et al.* 1996), and several other studies have also found high numbers of sterility factors (*e.g.*, True *et al.* 1996; Tao *et al.* 2003a,b). Therefore, the number of sterility factors separating closely related species is probably large. However, interpreting these results in terms of the number of sterility loci required for isolation is problematic, since sterility factors continue to evolve after reproductive isolation has been achieved. As pointed out by Orr (1995), the number of isolation factors will increase exponentially with divergence, making this an issue of real concern. Therefore, to test whether hybrid sterility is polygenic when it first evolves, very recently diverged species need to be examined. Orr and Irving (2001) examined two subspecies of *D. pseudoobscura*, and they estimated the number of sterility factors to be ~15. *An. gambiae* and *An. arabiensis* are even less diverged than these subspecies (Bullini and Coluzzi 1978; Orr and Irving 2001) and a genome-wide analysis of sterility effects may shed some light on this issue.

Curtis (1982) examined hybrid sterility in backcrosses of *An. gambiae* and *An. arabiensis* using a single marker on the *X* chromosome. He concluded that the *X* chromosome has a very large effect on male hybrid fertility and that the role of the autosomes is very minor. Here we report the first results of a detailed analysis into the genetic architecture of postzygotic reproductive isolation between *An. gambiae* and *An. arabiensis*. One goal is to examine and describe complex inviability effects. Genotyping many backcross individuals with numerous markers allows the detection of inviability factors that are apparent only from the absence of particular genotypes. Furthermore, we use a quantitative trait locus (QTL) mapping approach to identify chromosomal regions that contribute to male hybrid sterility.

MATERIALS AND METHODS

Strains and crossing scheme: One *An. gambiae s.s.* and one *An. arabiensis* colony were chosen from those available in insectaries in the Institute of Parasitology of the University of Rome on the basis of highest number of shared inversions, to maximize recombination between them. Gasua-*2La* (G) is an *An. gambiae s.s.* colony originating from females collected in Suakoko, Liberia, in 1986, polymorphic for the *2Rb* and *2La* inversions, and later selected to be a *Xag, 2R+, 2La, 3R+, 3L+* homozygote. Armor (A) is an *An. arabiensis* colony derived from adult females collected in Moribabougou, Mali, in 1996, and characterized by the following polytene complement: *Xbcd, 2Rab/c/+, 2La, 3Ra/+, 3L+*. Larvae were reared in distilled water with 0.1% marine salt at 28° ± 1° and 70 ± 5% relative humidity and were fed on dry cat food pellets. Adult mosquitoes were kept in 50-cm^3 cages at 26° ± 1° and 70 ± 5% humidity and fed on 1% sugar solution. Females were blood fed twice on guinea pigs before oviposition. Both larvae and adults were kept on a 12-hr photoperiod.

G females were crossed with A males to produce F_1 hybrids (GA). The females were backcrossed to the males of both G and A to obtain two backcrosses: GA × G and GA × A. The males of these backcrosses were used for our analyses. The analyses were based on 430 GA × G males and 453 GA × A males. All backcross males were mass reared, combining the offspring of several females. Backcross males and females were also allowed to mate among themselves to produce what we refer to here as $(GA \times A)_{BC2}$ and $(GA \times G)_{BC2}$. In these generations introgressed chromosomal segments can be homozygous. These BC_2 populations were produced from a random mass mating. We also performed AG × A and AG × G backcrosses, in which the F_1 hybrids were derived from *arabiensis* females and *gambiae* males. This cross produced very low numbers of offspring and was not used for most of our analyses.

The phenotype: Sperm development of backcross males was classified in six categories that reflect the ratio of normal/abnormal sperm: 1, only normal sperm present; 2, mostly normal sperm present; 3, equal numbers of normal and abnormal sperm present; 4, mostly abnormal present; 5, only abnormal sperm present; and 6, no sperm development. Exact measurements of the ratio of normal/abnormal sperm were not feasible, and classification of males was therefore done by estimation. Sperm development of F_1 hybrids (GA) was checked in males of up to 7 days old to determine if a delay in sperm development was present. No delays were detected, but dissections of backcross males were performed no earlier than 24 hr after emergence.

Markers: Primers for amplification of microsatellite loci were taken mostly from Zheng *et al.* (1996). Names of loci follow Zheng *et al.* (1996), but without the prefixes *AGXH, AG2H*, or *AG3H*. Primer sequences for marker 32J0 were kindly provided by L. Zheng (unpublished data). In Figure 1 the position of the markers based on their position in the *An. gambiae* genome sequence (Holt *et al.* 2002) is represented. The *An. gambiae* genome is not complete and has gaps of unknown length. In the genome sequence these were assigned an average length on the basis of the proportion of missing sequence (Holt *et al.* 2002). Therefore, the exact position of the markers should be treated with some caution.

Microsatellite loci were tested on at least 20 individuals from the Gasua and Armor strain. Loci that shared alleles between the two strains were excluded. A later phase of the analyses revealed the presence of null alleles at two loci (11 and 525) in one of the two strains. These loci were used only for QTL analyses of the backcross in which the presence of the null allele is known due to the direction of the backcross. In these cases, a positive control was provided by the amplification of alleles from other loci in the same PCR reaction.

Molecular methods: DNA extractions were performed using the easy-DNA kit (Invitrogen, Carlsbad, CA). PCRs were performed with AmpliTag Gold (Applied Biosystems, Foster City, CA), with the following program: 94° for 12 min, 30 cycles of

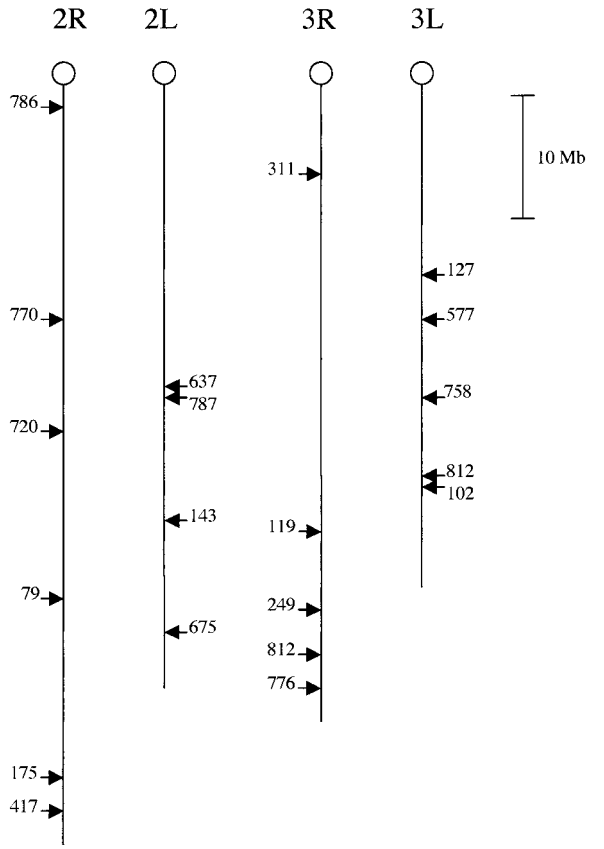

FIGURE 1.—Position of the microsatellite markers based on their position in the *An. gambiae* genome. The *An. gambiae* genome contains gaps of unknown length and the marker position should be treated with some caution.

94° for 40 sec, 55° for 30 sec, 72° for 30 sec, with a final extension step of 72° for 5 min. One to four loci were amplified in a single PCR reaction. Primers were labeled with fluorescent dye and PCR products were run on an ABI 373 automated sequencer (Applied Biosystems).

The strains were tested for the presence of Wolbachia symbionts using a diagnostic PCR method following ZHOU *et al.* (1998). DNA from Wolbachia-infected *D. melanogaster* was used as a positive control. PCRs were performed on DNA extractions of whole individuals, as well as of dissected ovaries.

Analyses: Extraction of genotypic data was performed using Genescan and Genotyper (Applied Biosystems) software. Genotypes were coded as homozygotes or heterozygotes only; *i.e.*, we did not distinguish between different alleles from the same species. Genotype data sets were visually inspected. Genotypes of any locus that differed from those of both flanking loci, which could be indicative of an experimental error, were reamplified. In a very few cases the genotype was not determined with confidence and these genotypes were removed from the data set.

Genetic maps were constructed with the Mapmaker software, version 3.0 (LANDER *et al.* 1987). Composite interval mapping (CIM; ZENG 1994) and multiple interval mapping (MIM; KAO *et al.* 1999) were performed to identify and calculate the effect of QTL affecting male hybrid sperm development using QTL cartographer (BASTEN *et al.* 2001). The experiment-wide 5% significance level for CIM was determined by permutating the marker and trait data 1000 times (DOERGE and CHURCHILL 1996). For all CIM analyses a window size of either 0.5 or 1 cM was used. Significant markers to use as cofactors in the analyses were identified using stepwise regression ($P < 0.05$). Because the proper significance criterion for maintaining a QTL in the model selection performed by MIM is unclear, CIM was used to identify significant QTL. The results from these analyses were used as an initial model in MIM analyses, using information criterion 1 (BASTEN *et al.* 2001). In only one case did the MIM analyses result in finding additional QTL. Here, MIM identified as significant two QTL that were below the significant threshold in the CIM analyses. These QTL are considered putative.

Additive effects of QTL, the percentage of variance explained by QTL, and epistatic interactions were calculated with MIM following the procedure outlined above. Epistatic interactions were further examined using EPISTAT (CHASE *et al.* 1997). This program uses likelihood ratios (LR) to statistically compare an epistatic and additive model. We performed 1,000,000 Monte Carlo simulations to determine the probability of exceeding a given log-likelihood ratio (CHASE *et al.* 1997).

RESULTS

The crosses: The GA cross between *An. gambiae* females and *An. arabiensis* males yielded only males with fully arrested sperm development. No normal or abnormal or immature sperm were observed in 40 dissected males. In the AG cross all F_1 males had abnormal or immature sperm present, and a small number of males (5 of 45 examined) had a few fully developed spermatozoa. In the first GA crosses (October 1999), the number of males and females in the F_1 generation was counted, and no sex-ratio distortion was detected (256 males *vs.* 239 females, $P = 0.44$). This cross was performed again in June 2000, and although the exact number of males and females was not determined, it produced large numbers of both males and females. However, when this cross was repeated in March 2001 a strong bias in sex ratio was observed, and of ~800 mosquitoes only a few were males. Whether this was due to embryonic death or an imbalance of *X*- or *Y*-bearing sperm was not established. However, larvae viability was high, yielding a great majority of females. The mosquito strains were subsequently tested for contamination with other strains, but no new alleles were observed. Both strains were also tested for Wolbachia infection, a group of bacterial symbionts that are known to cause cytoplasmic incompatibility (YEN and BARR 1971, 1973), using a diagnostic PCR procedure (ZHOU *et al.* 1998). However, no Wolbachia DNA was detected.

The genetics of hybrid inviability: A substantial number of backcross individuals were scored for markers covering much of the genome. An absence or deficiency of specific genotypes from the backcross population indicates inviability of those particular genotypes. In Table 1, the number of individuals carrying foreign alleles of the GA × G and GA × A backcrosses is represented. Throughout we use the term *native allele* for alleles that are derived from the strain that contributes the *Y* chromosome to the backcross generation, whose autosomes can be homozygous. Alleles of the other strain are referred to as *foreign alleles*. In the GA × G backcross there was a remarkable difference in the pattern shown

TABLE 1

Percentage of individuals carrying foreign alleles in backcross populations

		GA × G		GA × A	
Marker	Division	Female	Male	Female	Male
		X chromosome			
53	3D	52.0	53.7	47.3	9.7***
		Second chromosome			
417	7B	63.1***	60.7***	63.2***	60.2***
175	8C	63.4***	60.0***	64.9***	62.8***
79	11C	65.4***	62.1***	64.7***	64.1***
720	14B/C	66.4***	61.9***	64.3***	63.4***
770	15E	66.7***	61.3***	64.4***	63.9***
786	17C	66.7***	61.8***	64.3***	63.9***
143	25D/26A	62.4***	61.2***	61.7***	62.2***
787	23D	62.3***	59.9***	58.1***	63.4***
637	23D	60.6***	59.5***	58.3***	62.6***
675	28A/B	53.2	59.4***	56.1*	56.3**
		Third chromosome			
776	29B	49.8	53.5	60.4***	54.8*
812	29D/30A	50.0	53.7	61.3***	56.1**
249	30B	50.5	53.7	61.2***	56.1**
119	31B	51.0	53.0	61.8***	56.1**
311	35B	52.2	54.8*	63.0***	55.2*
127	41B	54.0	55.4*	64.8***	55.0*
577	42A	53.5	54.2*	62.8***	54.6
758	43B	52.3	54.9*	63.2***	55.2*
102	44B/C	52.7	51.2	61.5***	55.9**
817	44A	51.5	52.3	60.3***	55.8*

Significance was tested use a chi-square test, using 50.0 as the expected value. Division refers to the chromosomal division on the polytene chromosome. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$.

by the different chromosomes. A significant excess of heterozygotes occurred at almost all loci on the second chromosome in both females and males. On the *X* and the third chromosome the frequency of foreign alleles for most loci is not significantly different from the 50% Mendelian expectation. In the GA × A cross the females showed an excess of heterozygotes at all autosomal loci (56.1–64.9%). The X_G chromosome showed no evidence of segregation distortion or differences in viability in females. In the males, a significant excess of heterozygotes appeared at all but one locus on the two autosomes (54.6–64.1%). However, unlike in females, loci on the second chromosome have higher heterozygosity levels than those on the third chromosome.

One explanation for the heterozygote excess could be heterosis. Heterosis in crosses between these two species has been observed previously by DELLA TORRE *et al.* (1997). That heterosis is at least partially responsible for heterozygote excess in our crosses was confirmed by a comparison of the number of individuals entirely hetero- or homozygous for the autosomes. In the GA × A males, the number of ($X_A/2_{AA}/3_{AA}$) and ($X_A/2_{AG}/3_{AG}$) individuals was 17 and 43, respectively ($P < 0.001$, chi-square test). In the GA × A females, the number of ($X_{AA}/2_{AA}/3_{AA}$) and ($X_{AA}/2_{AG}/3_{AG}$) individuals was 10 and 31, respectively ($P < 0.01$, chi-square test). In the GA × G females, the number of ($X_{GG}/2_{GG}/3_{GG}$) and ($X_{GG}/2_{AG}/3_{AG}$) individuals was 4 and 14, respectively ($P < 0.05$, chi-square test). In the GA × G males, there was not a significant difference in the number of ($X_G/2_{GG}/3_{GG}$) and ($X_G/2_{AG}/3_{AG}$) individuals (11 and 13, respectively), but incompatibilities between the hemizygous X_G and the heterozygous *arabiensis* autosomes could have affected this comparison. Recessive inviability interactions of autosomes with the *X* chromosome or recessive-dominant two-locus autosomal incompatibilities could contribute to the observed excess of heterozygotes as well. Such incompatibilities could result in a deficiency of homozygotes in regions linked to recessive incompatibility factors.

The most conspicuous result, however, is the low frequency of the X_G chromosome in males. Only 9.7% of the GA × A males carry an X_G chromosome, indicating that the X_G chromosome causes inviability in an *arabiensis* genetic background. A small number of AG × A males and females (62 and 126, respectively) were also analyzed, and the frequency of the X_G in this cross was similar to that in the GA × A cross; *i.e.*, 10% of the males and 48.8% of the females have an X_G chromosome.

The low number of GA × A males carrying an X_G chromosome was examined in more detail. In Table 2, the percentage of homozygotes for alleles at autosomal loci is represented. There are five loci at which no individuals are homozygous for the background alleles. All other loci have very low percentages of homozygosity. This indicates that inviability is caused by an incompatibility between the X_G chromosome and at least one locus on each autosome of *arabiensis*. Furthermore, these *arabiensis* autosomal factors, when homozygous, cause complete inviability in combination with the X_G, since these genotypes were absent from our population. The same loci are heterozygous in the AG × A X_G males.

We also analyzed the offspring of matings between males and females of both the GA × G and GA × A backcrosses. These offspring, called BC_2 here, can be homozygous for both alleles. In Table 3 the number of observed foreign homozygotes is reported. The expected number is calculated on the basis of allele frequencies in the BC_2, assuming Hardy-Weinberg proportions. Our inference of complete inviability of two *arabiensis* homozygous autosomal regions with the X_G is supported by the complete absence of foreign homozygotes at linked markers in the $(GA \times G)_{BC2}$ population. A comparison of Tables 2 and 3 indicates that the inviability factors are closely linked to marker 786 on the *arabiensis* second chromosome and to markers 311 and 127 on the *arabiensis* third chromosome.

To investigate inviability interactions between the X_A and *gambiae* autosomes, the numbers of ($X_G/2_{GG}/3_{GG}$)

TABLE 2

Percentage of homozygosity in GA × A males with X_G chromosome

Marker	Division	%
	Second chromosome	
417	3D	12.5**
175	7B	6.3***
79	11C	3.1***
720	14B/C	3.1***
770	15E	3.1***
786	17C	0***
143	25D/26A	0***
787	23D	3.1***
637	23D	3.1***
675	28A/B	21.9*
	Third chromosome	
776	29B	18.8**
812	29D/30A	12.5***
249	30B	12.5***
119	31B	6.3***
311	35B	0***
127	41B	0***
577	42A	0***
758	43B	6.3***
102	44B/C	12.5***
817	44A	12.5***

Expected homozygosity values are based on the allele frequency of the total GA × A population, assuming Hardy-Weinberg equilibrium. A chi-square test was used to test for significant deviation. Division refers to chromosomal division of polytene chromosomes. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$.

TABLE 3

Percentage of foreign homozygotes in BC_2 populations

	GA × G		GA × A	
Marker	Observed	Expected	Observed	Expected
		X chromosome		
53	0	0.2	0	0
		Second chromosome		
417	0.5***	3.3	11.5*	14.3
175	0.5***	3.3	11.2**	14.9
79	0***	3.2	11.2***	18.2
720	0***	3.3	11.5***	17.6
770	0***	3.2	11.8***	17.7
786	0***	3.6	12.0***	17.1
143	0.4***	2.7	7.8***	12.4
787	0.4***	4.4	5.7**	8.6
637	0.7***	4.7	5.1***	8.6
675	1.1***	5.1	4.6**	8.4
		Third chromosome		
776	3.5	4.2	0.8*	2.4
812	1.4*	2.5	0.3***	2.3
249	0.2***	1.5	0.3**	2.1
119	0.2***	1.4	0.3***	2.2
311	0***	2.2	0.5**	2.4
127	0***	2.6	1.6	0.7
577	0.2***	2.9	1.3	1.0
758	0.5***	3.3	1.3	1.3
102	0.9***	4.2	0.5*	1.8
817	0.7***	3.6	0.5*	1.7

Expected values were calculated on the basis of allele frequencies in the BC_2 population, assuming Hardy-Weinberg equilibrium. A chi-square test was used to test for significant deviations. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. Sample size for $(GA \times G)_{BC2}$ ranges from 540 to 561 for each locus, the average being 558. For $(GA \times A)_{BC2}$, the sample size ranges from 365 to 376, the average being 373.

and ($X_A/2_{GG}/3_{GG}$) individuals in the GA × G males can be compared. Both numbers are very low (11 and 7, respectively) and they are not significantly different. However, they do indicate that no incompatibilities between the X_A and the *gambiae* autosomes cause complete inviability.

The interaction of the X_G is not solely with the homozygous *arabiensis* autosomal inviability loci. We found 170 GA × A backcross males that were heterozygous at loci 786, 311, and 127. Half of these are expected to carry an X_G chromosome, but only 45 do so ($P \ll 0.001$, chi-square test). The number of ($X_G/786_{AG}/311_{AG}$) and ($X_A/786_{AG}/311_{AG}$) individuals is 45 and 131, respectively ($P \ll 0.001$, chi-square test) in the GA × A males. This indicates that the *arabiensis* inviability factors are not completely recessive or that additional semilethal incompatibilities are present in ($X_G/786_{AG}/311_{AG}$) individuals. However, in the GA × G males the number of ($X_G/2_{GG}/3_{GG}$) and ($X_A/786_{AG}/311_{AG}$) individuals was 65 and 74, respectively, which is not significantly different. A comparison of X_G and X_A GA × G males with completely heterozygous autosomes (13 and 31, respectively; $P < 0.01$, chi-square test) does show a significant difference in viability of these genotypes. In some ($X_G/786_{AG}/311_{AG}$) individuals of the GA × A cross, parts of the autosomes are homozygous and this may explain the difference between the two observations. However, a comparison of X_G and X_A GA × A males with completely heterozygous autosomes (20 and 43, respectively; $P < 0.01$, chi-square test) indicates that this cannot be the sole explanation. As far as the X_G chromosome is concerned, the number of ($X_{AA}/786_{AA}/311_{AA}$) and ($X_{AG}/786_{AA}/311_{AA}$) individuals (32 and 22, respectively; $P = 0.161$) suggests that the inviability factors on the X_G chromosome tend to be recessive.

The inviability of particular male genotypes is expected to result in a biased sex ratio. However, we did not count the number of males and females in our backcrosses, so we are unable to verify that this was the case. The sex-ratio distortion observed in the last G × A cross could have been caused by the inviability effect of the X_G chromosome. Therefore, genotypes were obtained for the small number of available AG × G individuals that derived from the G × A cross from June 2000,

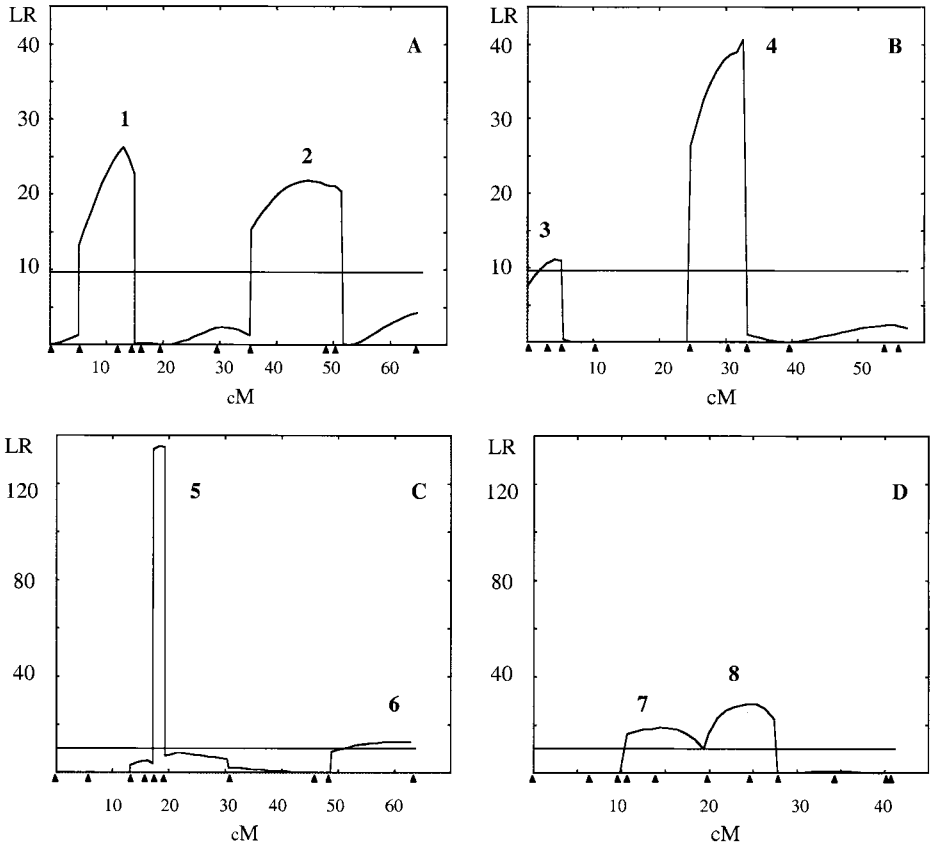

FIGURE 2.—QTL affecting male hybrid sterility in backcrosses between *An. gambiae* and *An. arabiensis*. The LR is plotted against the genetic map. The triangles indicate marker positions. The horizontal line represents the critical value ($P = 0.05$) for experiment-wide type I error. (A) Chromosome *2* of GA × G cross, (B) chromosome *3* of GA × G cross, (C) chromosome *2* of GA × A cross, and (D) chromosome *3* of GA × A cross are shown.

which produced no obvious sex-ratio distortion. The sample size was only 19, but only one of these mosquitoes carried an X_G chromosome. This is a significant deviation from the expected 8.5 ($P \ll 0.001$, chi-square test) and indicates that the inviability of the X_G chromosome was present even when the cross did not show any obvious sex-ratio distortion. However, AG × A individuals ($n = 46$) from the 1997 cross had an X_G frequency of 40%. In these individuals the *arabiensis* markers linked to the recessive inviability factors are sometimes homozygous. This clearly indicates that one or both of the strains were originally polymorphic for inviability factors.

TABLE 4

Additive effects and percentage of variance explained by QTL and their epistatic interactions in GA × G and GA × A crosses

GA × G			GA × A		
QTL	Additive[a]	%[a]	QTL	Additive[a]	%[a]
X	2.08	39.5	*X*	0.75	4.7
1	0.63	4.6	5	1.26	20.3
2	0.47	4.2	6	0.41	3.4
3	0.41	4.2	7	0.79	13.8
4	0.77	7.9	8	0.82	14.2
X–1	1.58	4.1	7–8	0.79	0.1
X–2	1.04	2.9			
X–3	1.19	5.5			
X–4	1.46	6.8			
3–4	0.43	0.3			

[a] As calculated by multiple interval mapping implemented in QTL cartographer.

Genetic map distances: Genotypes were determined for locus 53 in the *Xag* inversion, which covers ~75% of one end of the *X* chromosome, and marker 32J0, located near the centromere on the opposite end, in 764 individuals. Not a single recombinant was found, indicating that the presence of the *Xag* inversion suppresses recombination along almost the entire chromosome. A comparison of our genetic distances between markers and those published previously (ZHENG *et al.* 1996) suggests that recombination in our crosses was low in regions in which our strains were polymorphic for inversions, as expected. This can affect the detection of QTL, since they are more likely to be detected in regions of low recombination.

QTL mapping of male hybrid sterility factors in the GA × G cross: In Figure 2, A and B, the results of CIM of sperm development in GA × G males are represented. In this cross the effect of introgressed *arabiensis* alleles in a *gambiae* background was examined. Due to the lack of recombination, the *X* chromosome acts as a single marker, and only the second and third chromosomes are represented. In Table 4, the additive effect and the percentage of the phenotypic variance explained by each QTL, as well as their interaction calcu-

lated by MIM, are represented. The total amount of phenotypic variance explained by the QTL is 80.5%. The *X* chromosome explains 39.5% of the phenotypic variance, excluding epistatic interactions. The LR score of the *X* chromosome is 212.2. This is highly significant, since the 5% experiment-wide significance threshold is at an LR score of 9.7.

Two significant QTL were found on each autosome, explaining 4.2 and 7.9% of the variance. The interpretation of what these QTL signify is not straightforward, since their contribution to the phenotypic variance can be the result of at least two effects. First, at the QTL position, one or more dominant genes can exist in the "foreign" species that cause incompatibility with the genetic background of the "native" species. Under this scenario, the presence of a foreign allele at the QTL position reduces the fertility of the hybrid. In the second scenario, a recessive native gene exists at the QTL location that is incompatible with some of the introgressed foreign alleles. Under this scenario, the presence of a foreign allele at the QTL position leads to a decrease in sterility. These two scenarios can therefore be distinguished by the direction of their effect.

The phenotypic mean of the GA × G males carrying an X_A chromosome is 5.9. The mean of the individuals carrying X_G chromosomes is 4.2. Therefore, if an X_A chromosome is present, the fertility decreases. The phenotypic mean of individuals that are heterozygous for either the second or the third chromosome, but otherwise native homozygous, is 3.5 and 4.1, respectively. The phenotypic mean of individuals that are all native homozygous is 1.5. Furthermore, since all the effects of the QTL in this cross are in the same direction, the QTL indicate a situation described above under the first scenario; *i.e.*, at the QTL position the *arabiensis* genome contains one or more dominant genes that are incompatible with the *gambiae* background. Note also that the LR statistic is extremely low for substantial lengths along the two autosomes.

The GA × G data set was divided into two data sets, one of which contained all individuals carrying an X_G chromosome ($n = 199$) and the other contained individuals carrying an X_A chromosome ($n = 231$). This division served several purposes. First, additional small QTL, whose effects were previously obscured by the large effect of the X_A chromosome, may be detected. Second, autosomal QTL whose effects were derived solely or mostly from incompatibilities with the X_G chromosomes are detectable only in the X_G data set, whereas QTL that act mostly with the autosomal background may be detected in both data sets. Finally, recessive *gambiae* autosomal regions that are incompatible with the X_A may be detected in the X_A data set. As noted previously, these regions contribute to the phenotypic variance in backcrosses and may be detectable in a QTL analysis. Conceivably, their effect can be obscured by the presence of dominant *arabiensis* sterility factors on the autosomes. The X_A data set provides increased power to detect these recessive factors, especially if some of the effect of the obscuring *arabiensis* sterility factors is due to incompatibility with the X_G chromosome.

In Figure 3, A and B, the results of the QTL analysis of the X_G data set are presented. The additive effects and the percentage of the phenotypic variance explained by the QTL are provided in Table 5. Four significant QTL were detected using CIM, explaining 6.5–24.4% of the phenotypic variance. Using MIM, two additional QTL were detected, although they are here considered tentative. The total amount of the phenotypic variance explained by all the QTL is 75.7%. A comparison of Figure 3, A and B, to Figure 2, A and B, showed that, using the X_G data set, an additional small QTL (no. 14) at the tip of the third chromosome and a putative QTL (no. 10) on the second chromosome were detected. QTL 1, 3, and 4 from the total GA × G data set are all present in the X_G data set analyses. This is expected if their action is not contingent upon the X_A chromosome. However, QTL 2 from the total data set has disappeared from this analysis. Its location is close to that of putative QTL 11, and they may be identical. Also, under this analysis some regions that have very low LR scores remain.

In Figure 3, C and D, the LR scores of the CIM analysis of the X_A data set are represented. The additive effects and percentages of explained variance are given in Table 5. One QTL was detected on each autosome. The total amount of variance explained was 16.7%. The direction of the effect of both QTL was opposite that of the QTL detected in any of the previous analyses. In this case, if individuals were homozygous for the background *gambiae* alleles at the QTL position, sterility was increased. Therefore, these QTL identified *gambiae* chromosomal regions that contain (partially) recessive sterility factors whose interaction was entirely or mostly with the X_A chromosome. Only two very small regions were found at which (partially) recessive sterility factors were present. If the QTL were not completely recessive, they would be detected in the analyses of the GA × A cross as well. The LR score for almost the entire second and third chromosomes was very low.

QTL mapping of male hybrid sterility factors in the GA × A cross: In Figure 2, C and D, the results of CIM of hybrid sterility factors in the GA × A cross are represented. In this cross, the effect of *gambiae* chromosomal segments when introgressed into an *arabiensis* background was assessed. In Table 4, the additive effects and the percentage of the phenotypic variance explained by the QTL are provided. The total amount of variation explained by all QTL and the *X* in this cross was 56.5%. This is considerably lower than that in the other cross, although this may be a consequence of the low frequency of the X_G chromosome in this cross. If the *X* chromosome has a large effect on sterility compared to the autosomes, its absence will reduce the amount of explained phenotypic variance. The amount of variance

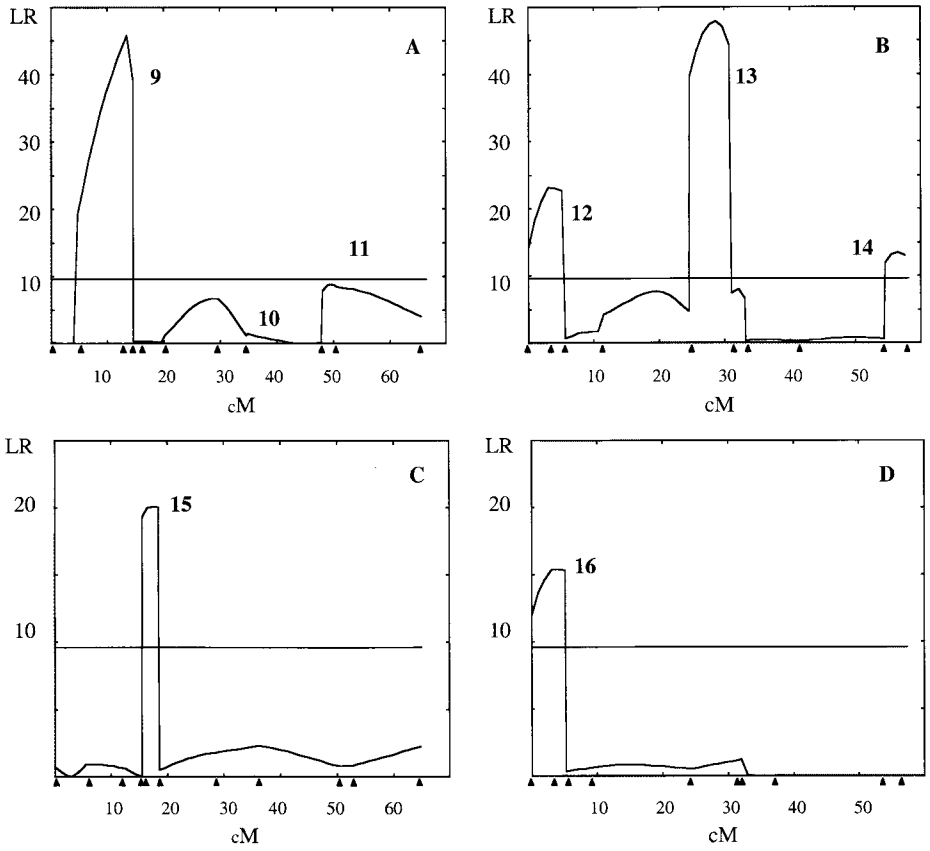

FIGURE 3.—QTL affecting male hybrid sterility in GA × G backcross using partial data set analysis. Notations and axes are as in Figure 2. (A and B) Chromosome *2* and *3* using data set containing individuals with the *gambiae X* chromosome. (C and D) Chromosome *2* and *3* using data set containing individuals with the *arabiensis X* chromosome.

explained by the *X* chromosome is 4.7%. The LR score of the *X* chromosome is 12.7. The 5% experiment-wide significance threshold for this data set is 9.9. The X_G chromosome is present only at very low frequency in this cross, and it is in the presence of autosomes that are largely heterozygous (see Table 2). Consequently, a small-to-moderate effect of the X_G chromosome becomes difficult to detect. The fact that a significant effect of the *X* chromosome was detected indicates that the X_G chromosome has a large effect on sterility, but that effect is obscured by inviability.

Four QTL were detected on the autosomes. Two QTL were located on each autosome. The amount of variance explained by these QTL ranged from 3.4 to 20.3%. The phenotypic means of individuals carrying either the *gambiae* second or third chromosome, compared to the individuals that are all homozygous, are 4.0 and 3.7 *vs.* 2.0, respectively. Considering that all detected QTL have effects in the same direction, the detected QTL indicate the location of *gambiae* sterility alleles. The percentage of explained phenotypic variance, as well as the size of the additive effect by three of the QTL, is high compared to that of QTL in the other cross (see Tables 4 and 5). Again, large regions of the autosomes have very low LR scores.

TABLE 5

Additive effects and percentage of variance explained by QTL and their epistatic interactions in partial GA × G data sets

X_G data set			X_A data set		
QTL	Additive[a]	%[a]	QTL	Additive[a]	%[a]
9	1.32	17.8	15	−0.16	6.4
10	0.60	8.2	16	−0.13	4.5
11	0.52	6.5	15–16	0.26	5.8
12	0.81	11.0			
13	1.42	24.4			
14	0.62	7.8			

[a] As calculated by multiple interval mapping implemented in QTL cartographer.

If the results from this analysis are compared with those presented in Figure 3, C and D, some interesting observations emerge. First, QTL 5 is in exactly the same location as QTL 15 in Figure 3C. QTL 15 in Figure 3C represents the effect of one or more *gambiae* factors when homozygous. In the GA × A cross, only the effect of heterozygous *gambiae* sterility factors is detected. Therefore, at the positions of QTL 5 and 15, sterility factors that have an effect when both homozygous and heterozygous are present. The heterozygous effect is represented by QTL 5, and the homozygous effect by QTL 15. Alternatively, multiple sterility factors could be present at this location, some of which are mostly dominant, and others of which are mostly recessive. The fact that QTL 16 from Figure 3D is undetected in the analysis

of the GA × A cross indicates that the sterility factor(s) located at this position is entirely or mostly recessive.

Epistatic interactions: One of the advantages of MIM is that it allows for the identification of epistatic interactions between QTL; *i.e.*, the effect of the combined QTL is larger or smaller than the sum of the individual effects. The MIM analyses indicated that epistatic interactions exist between all autosomal loci and the *X* chromosome in the GA × G cross. However, the specific interactions between the QTL are not easily discernible on the basis of the output of QTL cartographer. For example, epistasis detected by MIM could result from incompatibilities being complex, such that multiple loci need to carry a foreign allele in order for incompatibilities to be expressed. Alternatively, the detected epistasis could be interspecific; *i.e.*, Dobzhansky-Muller incompatibilities between the different genomes could be responsible for the epistasis detected by MIM.

In the GA × G cross, MIM detected epistatic interactions between the *X* chromosome and all autosomal QTL. However, when *arabiensis* alleles are present at both the QTL and the *X* chromosome, the phenotypic mean in all four cases is less than when the X_A is present by itself. The difference is very small, however, and the X_A causes complete sterility when in the presence of *gambiae* homozygous autosomes. This indicates that the detected epistatic interactions were not the result of complex incompatibilities. A similar comparison regarding the interaction between QTL 3 and 4, 7 and 8, as well as 15 and 16, leads to the same conclusion. The combined effect of these QTL pairs is less than the sum of their individual effect. Therefore, none of the epistatic interactions detected by MIM were the result of complex incompatibilities.

It has been suggested that epistasis can be a by-product of the manner in which sterility is measured (Palopoli and Wu 1994). Treating sterility as a binary trait, which has been a custom in studies of hybrid sterility, obscures effects of sterility factors that interact additively, but do not reach the threshold of complete sterility by themselves. Our conclusions on the lack of conspecific epistasis did not change, however, after the data were reanalyzed using a binary coding scheme, taking the percentage of completely sterile individuals as our phenotypic measure (results not presented here).

DISCUSSION

Inviability factors: We have identified incompatibilities between the *gambiae X* chromosome and at least one region of each *arabiensis* autosome, causing complete inviability when they are hemi- or homozygous. GA × A males heterozygous for the *arabiensis* inviability loci show an effect as well, but this is incomplete. This effect is also expected in GA × G males, but we did not find a significant effect in this cross. Perhaps interactions between the X_G and the *arabiensis Y* chromosome play a role as well.

The results of this study concur to a certain extent with what is known about hybrid inviability in Drosophila. The fact that the *arabiensis* autosomal factors are closely linked to one or two markers indicates that few inviability genes exist on the *arabiensis* autosomes. Furthermore, no inviability effects were found on the X_A, nor were any dominant inviability effects detected on the *gambiae* autosomes. Several studies have found that only a few regions of the genome cause inviability when introgressed into closely related species (Carvajal *et al.* 1996; Hollocher and Wu 1996; True *et al.* 1996). The fact that we have found only two autosomal regions that cause complete hybrid inviability indicates that not many inviability factors are present in the genome of *gambiae* and *arabiensis*. Our data also show that our strains were originally polymorphic for these inviability factors.

The dominance theory states that incompatibility factors tend to be recessive (Muller 1942). This has been found in several cases of Drosophila hybrid inviability. For example, Presgraves (2003) has shown that recessive inviable incompatibilities far outnumber dominant ones in *D. melanogaster* and *D. simulans* hybrids. Introgression studies have also shown that inviability factors tend to act recessively (Carvajal *et al.* 1996; Hollocher and Wu 1996; True *et al.* 1996). In hybrids between the haplodiploid wasp species *Nasonia vitripennis* and *N. giraulti*, the haploid males are much more inviable than the diploid females (Breeuwer and Werren 1995). Our data suggest that the inviability factors on the X_G chromosome are mostly recessive. The data on the *arabiensis* factors are inconclusive. The GA × A cross suggests that these factors could be largely dominant, but the GA × G cross does not support this.

Sex-ratio distortion and strain changes: We observed a very strong bias in sex ratio in our G × A cross, while previous crosses with the same strains produced large numbers of both males and females. Several possible reasons could explain this observation. One possibility is that our strains were contaminated with a bacterial symbiont, such as Wolbachia, which can cause male killing, feminization, or prevent male sperm from entering the egg. We did not detect any Wolbachia in our crosses, although other species of symbiont might be present. However, we think this is unlikely because in the backcross progeny of the sex-biased cross both males and females were plentiful. This would not be expected if a bacterial symbiont was responsible, unless the symbiont causes an effect only in combination with a hybrid genotype. However, in our GA × A backcross, males carrying the F_1 hybrid genotype were present. Meiotic drive might be responsible as well, but similarly, it is expected to cause problems in backcross progeny with the F_1 hybrid genotype. Additionally, meiotic drive has never been reported in these mosquitoes.

Another possibility is that rearing conditions varied

between the different crosses. For other taxa, it is well known that hybrid inviability is sensitive to environmental conditions (Bordenstein and Drapeau 2001). All rearing parameters were kept constant in our experiment, but small fluctuations in temperature or humidity cannot be ruled out. Finally, the most intriguing possibility is that a hybrid incompatibility factor became (almost) fixed in one of our colonies between 2000 and 2001. On the basis of the large numbers of both males and females that we observed in our backcross progeny, this would likely involve a *gambiae* maternal factor that acts recessively.

The Armor strain experienced a bottleneck before the cross with the deficiency of males was performed. This could have led to the fixation of previously rare alleles. The strains were derived from multiple females, and the microsatellite screening showed that even after several years in the insectary, approximately two-thirds of the microsatellite loci were polymorphic. Furthermore, Davidson (1964) reported that variation in the sex ratio is often observed in crosses of *An. gambiae s.l.* This led him to conclude that some of these species are variable for incompatibility factors. Our data have confirmed that *An. gambiae* and/or *An. arabiensis* are polymorphic for several inviability factors.

Hybrid sterility QTL: Most pairs of Drosophila species investigated rigorously have diverged some time ago. A recent study by Orr and Irving (2001) investigated hybrids between two recently diverged subspecies of *D. pseudoobscura* (Nei's $D = 0.194$). These authors estimated that ~15 sterility factors separate these subspecies. *An. gambiae* and *An. arabiensis* should be very useful taxa for studies into the genetics of hybrid sterility, in that they are very closely related. A Nei's genetic distance of $D = 0.15$, averaged over 30 loci, has been reported for these species (Bullini and Coluzzi 1978). Therefore, these species are more closely related than the species of Drosophila on which detailed studies on hybrid sterility or inviability have been performed. This is important since the number of sterility factors separating two species is expected to increase exponentially with divergence (Orr 1995).

Interpretation of a QTL mapping study in terms of the genetic architecture of the trait of interest is not straightforward. First, if power is too low, even QTL of moderate effect may not be found, and the effect of the identified QTL may be severely overestimated (Beavis 1994). Our sample sizes are 430 and 453, and 80% of the phenotypic variance in the GA × G cross and 56% in the GA × A cross were explained. The phenotypic values of individuals were determined by estimation, which must introduce some experimental variance. Furthermore, a certain amount of environmental variance is expected. Therefore, we are confident that much of the genetic variance was identified in this study. That is, it is not likely that many QTL of large-to-moderate effect were missed.

The two main issues in dealing with our data are how to interpret the identified QTL and how to interpret the regions of the chromosomes with very low LR scores. A QTL can indicate a single gene of moderate-to-large effect or, alternatively, numerous genes may be located at a QTL position. This issue is strongly affected by the relative recombination rate at the QTL location. If genes are randomly distributed across the genome, the probability that sterility factors will cluster is not very high. However, Noor *et al.* (2001b) pointed out that variation in recombination rates may affect the detection of QTL. On the basis of a simulation these authors showed that QTL of substantial effect can be found in regions of low recombination, even if only randomly distributed genes of small effect are present. An analysis of recombination rates in QTL regions *vs.* non-QTL regions based on our markers and on the *An. gambiae* genome (results not presented) suggests that our QTL are located in regions of low recombination. This complicates the interpretation of these results. However, our QTL explain a large proportion of the phenotypic variance and this could indicate that other regions of the genome contribute little to sterility. Until more accurate data on recombination rates in *An. gambiae* are available, and the performance of QTL mapping procedures is better documented, our results are consistent with both the presence of a small number of genes of moderate-to-large effect and the presence of numerous genes of small effect. Furthermore, although it is unlikely that regions with very low LR scores harbor QTL of moderate or large effect, sterility factors whose contribution is very small may be present in these regions.

Although it is not possible to put an upper limit on the number genes involved in male hybrid sterility in these species, at least five or six sterility factors were detected in each of the two species. None of the autosomal QTL causes complete sterility by itself. If a single foreign autosome of either species is made heterozyous in an otherwise homozygous background, sterility occurs in only a small proportion of the individuals. When both autosomes are heterozygous, full sterility occurs in all individuals. Due to the lack of recombination between the *X* chromosomes, it was not possible to analyze the *X* chromosome in any detail, and it is not known if all regions of the *X* contribute to hybrid sterility. However, on the basis of our analyses of the autosomes, we conclude that at least three to four sterility factors need to be present for full sterility.

The sterility QTL are not in the same positions in the two species; *i.e.*, *gambiae* sterility regions are different from the *arabiensis* sterility regions. This is expected under the Dobzhansky-Muller model, and similar results have been obtained in studies of Drosophila species (*e.g.*, Wu and Beckenbach 1983; Orr and Coyne 1989).

The absence of any recombination between the *X* chromosome of the two species corroborates the results of Curtis and Chalkey (1979). Using a single *X* chromo-

some marker and polytene chromosome analyses, these authors concluded that it is unlikely that recombination on the *X* chromosome takes place in hybrids.

A large effect of the *X* chromosome on male sterility was found in both crosses. This large *X* effect has been described previously for many different Drosophila species (reviewed in Coyne and Orr 1989), as well as for *An. gambiae* and *An. arabiensis* hybrids (Curtis 1982). Charlesworth *et al.* (1987) argued that genes that are recessive and advantageous evolve faster when located on the *X* chromosome. This could be responsible for this large *X* effect. Alternatively, the hemizygosity of the *X* chromosome could be responsible, but unfortunately a backcross analysis does not allow a distinction between the two scenarios.

Our search for recessive sterility factors in the X_A data set detected only two regions with a recessive effect. One of these is in the same location as a large QTL with a substantial dominant effect, leaving only a single QTL that acts mostly recessively. However, the analyses of the X_A data set cannot detect recessive-recessive autosomal incompatibilities, but only interactions of the recessive *gambiae* autosomes with the X_A. The analyses of the GA × G cross have shown that the effect of all autosomal QTL is dependent on the presence of the X_G chromosome. This we derive from the fact that these QTL went undetected in the X_A data set analysis. A possible complication for a comparison of the number of recessive and dominant sterility factors on the autosomes could be a difference in the power of the X_G and the X_A data set. Even though the sample sizes are similar between the two data sets, the phenotypic variance is not, and conceivably, autosomal effects are harder to detect in the X_A data set.

The Dobzhansky-Muller model in its simplest form concerns negative epistasis between two genes. More complex forms are possible, and Orr (1995) suggested that it is expected that incompatibilities underlying sterility or inviability involve numerous genes. This is because the proportion of evolutionary pathways that are not subject to purifying selection is higher when incompatibilities are more complex. The MIM analyses indicated epistatic interactions both between autosomal QTL and the *X* chromosome in the GA × G cross and between two autosomal QTL in both crosses. However, none of these cases of epistasis is indicative of complexity in Dobzhansky-Muller incompatibilities. In every case the presence of foreign alleles at both the QTL and *X* chromosome leads to less sterility than expected under additivity.

Several studies have detected epistasis between closely linked conspecific genes, such that multiple genes need to be co-introgressed to cause sterility (Cabot *et al.* 1994; Palopoli and Wu 1994; Perez and Wu 1995). Palopoli and Wu (1994) suggested that epistasis is expected to be common between closely linked loci if the epistasis played a role in the fixation of the sterility factors that act epistatically. We do not know if our QTL contain multiple sterility factors whose co-introgression is necessary for sterility, so this hypothesis cannot be tested. However, our results contrast with a study by Orr and Irving (2001) and other studies (Muller 1942; Orr and Coyne 1989; Davis *et al.* 1994) that have reported epistatic interactions between conspecific sterility loci on different chromosomes.

A possible explanation for the lack of conspecific epistatic interactions in these data could be that a QTL mapping analysis is not particularly good at detecting QTL whose interaction is contingent upon the presence of co-introgressed alleles at many different loci, especially if these loci are unlinked. For example, if four or five factors are spread throughout the genome and all are required for the expression of sterility, only a few individuals would be sterile due to the action of these loci. Their contribution to the phenotypic variance will be small and very hard to detect. This will be especially true if other QTL that overshadow regions of smaller effect are present. However, a comparison of the effect of entire chromosomes in our backcrosses does not indicate that the autosomes harbor many such genes whose interaction is with genes on the same autosomes. In short, it is likely that sterility factors that are part of very complex incompatibilities are extremely hard to detect; however, several QTL whose effect does not depend on other regions of the genome were identified.

Isolation factors and inversions: Inversions are thought to play a crucial role in the evolution of the *An. gambiae* complex (Coluzzi *et al.* 1979; Coluzzi 1982). Both shared and nonshared inversions are present in these species. Shared inversions imply little selection against introgression (della Torre *et al.* 1997). It has been suggested that sterility or inviability factors should be preferentially present in nonshared inversions (Rieseberg *et al.* 1999; Noor *et al.* 2001c). Therefore, it is of interest to examine whether the identified isolation factors map to the same regions as the inversions.

The inviability regions on the *arabiensis* autosomes are all located outside the inversions. Sterility QTL 1 and 2 are located in the same regions as shared inversions. However, the *arabiensis* strain used in this study is polymorphic for these inversions and the frequency of them in the backcross is not known. Therefore, we do not know if these QTL map to the inverted or noninverted chromosome.

QTL 2 is linked most closely to markers 787 and 637. Both these markers are located within the *2La* inversion, which is fixed in both strains. *An. gambiae* is thought to have acquired this inversion from *An. arabiensis* (Caccone *et al.* 1998), and it introgresses readily between the two species in laboratory colonies (della Torre *et al.* 1997). QTL 2 indicates that this inversion contains *arabiensis* sterility factors that affect male fertility when introgressed into *An. gambiae*. It is possible that these sterility factors evolved after *An. gambiae* acquired the inversion and that no or little introgression has occurred

since. The effect of this QTL may not be strong enough to prevent introgression in a laboratory colony, but selection against it in nature may be strong enough to prevent their exchange.

Several authors have suggested that inversions are more likely to contain isolation factors, because inversions preserve linkage groups and can prevent homogenization of newly isolated taxa (Rieseberg *et al.* 1999; Noor *et al.* 2001c; Navarro and Barton 2003). These theories are similar to one offered by Coluzzi (1982) to provide a model for speciation in the *An. gambiae* complex, although this author envisioned the role of inversions in environmental adaptation as the key factor. This study did not show a preferential presence of inviability or sterility factors in nonshared inversions. If the speciation processes of the *An. gambiae* complex were largely driven by the adaptation to marginal environments during population expansions, as envisioned by Coluzzi (1982), the accumulation of sterility and inviability factors may have played a negligible role. Their preferential presence in nonshared inversions would not be expected under such a scenario.

Synopsis: Our understanding of the genetic basis of reproductive isolation is hampered by the lack of variety of taxa for which detailed information is available. In this study a genome-wide survey of inviability and male sterility factors in hybrids between *An. gambiae* and *An arabiensis*, two very closely related species of mosquito, was performed. In addition, we reported the possible fixation of male hybrid inviability in a laboratory colony. We have identified previously unknown lethal male hybrid genotypes and an incompatibility of mostly recessive *X* chromosomal factor(s) with autosomal factors, causing inviability. Furthermore, regions that contribute to male hybrid sterility were identified. The amount of phenotypic variation explained by our QTL suggests that large regions of the autosomes contribute little to hybrid sterility. Additionally, at least three to four loci are required for complete sterility, and epistatic interactions between conspecific genes may not be the rule. The results presented here will be important for a comparison with the data of an accompanying study of female sterility factors (M. Slotman, A. della Torre and J. R. Powell, unpublished results) and the introgression capacity of chromosomal regions between these species (M. Slotman, A. della Torre, M. Calzetta and J. R. Powell, unpublished results).

We are very grateful to M. Coluzzi for general support of this work and to G. Petrangeli and M. Calzetta, whose technical support was instrumental. Furthermore, we thank H. Zhao and J. Gleason for assistance with the QTL mapping analyses, R. Doerge for a discussion of our results, and D. A. Elnaiem for statistical advice. We are also grateful to L. Zheng for kindly providing microsatellite primer sequences prior to publication. Two reviewers and the associate editor, M. A. F. Noor, provided many insightful comments on an earlier version of this article that greatly helped us improve the manuscript. J.R.P. and M.S. were supported by National Institutes of Health grant R01 46018 and M.S. was also supported by the Centers for Disease Control and Prevention Fellowship Training Program in Vector-Borne Infectious Diseases T01/CCT122306. A.T. was supported by the United Nations Development Programme/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases (Tropical Disease Research) and the Ministero dell'Istruzione, dell'Università e della Ricerca/COFIN funds.

LITERATURE CITED

Basten, C., B. S. Weir and Z-B. Zeng, 2001 *QTL Cartographer: A Reference Manual and Tutorial for QTL Mapping*. Department of Statistics, North Carolina State University, Raleigh, NC.

Beavis, W. D., 1994 The power and deceit of QTL experiments: lessons from comparative QTL studies, pp. 250–266 in *Proceedings of the 49th Annual Corn and Sorghum Industry Research Conference*. American Seed Trade Association, Washington, DC.

Bordenstein, S. R., and M. D. Drapeau, 2001 Genotype-by-environment interaction and the Dobzhansky-Muller model of postzygotic isolation. J. Evol. Biol. **14:** 490–501.

Breeuwer, J. A. J., and J. H. Werren, 1995 Hybrid breakdown between two haplo-diploid species: the role of nuclear and cytoplasmic genes. Evolution **49:** 705–717.

Bullini, L., and M. Coluzzi, 1978 Applied and theoretical significance of electrophoretic studies in mosquitoes (Diptera: Culicidae). Parassitologia **20:** 7–21.

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.

Caccone, A., G. S. Min and J. R. Powell, 1998 Multiple origins of cytologically identical chromosome inversions in the *Anopheles gambiae* complex. Genetics **150:** 807–814.

Carvajal, A. R., M. R. Gandarela and H. F. Naveira, 1996 A three-locus system of interspecific incompatibility underlies male inviability in hybrids between *Drosophila buzzatii* and *D. koepferae*. Genetica **98:** 1–19.

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

Chase, K., F. R. Adler and K. G. Lark, 1997 Epistat: a computer program for identifying and testing interactions between pairs of quantitative trait loci. Theor. Appl. Genet. **94:** 724–730.

Coluzzi, M., 1982 Spatial distribution of chromosomal inversions and speciation in Anopheline mosquitoes, pp. 113–153 in *Mechanisms of Speciation*, edited by C. Barigozzi. Alan R. Liss, New York.

Coluzzi, M., A. Sabatini, V. Petrarca and M. A. Dideco, 1979 Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. Trans. R. Soc. Trop. Med. Hyg. **73:** 483–497.

Coyne, J. A., 1984 The genetic basis of male sterility in hybrids between two closely related species of Drosophila. Proc. Natl. Acad. Sci. USA **81:** 4444–4447.

Coyne, J. A., and H. A. Orr, 1989 Two rules of speciation, pp. 180–207 in *Speciation and Its Consequences*, edited by D. Otte and J. A. Endler. Sinauer Associates, Sunderland, MA.

Coyne, J. A., S. Simeonidis and P. Rooney, 1998 Relative paucity of genes causing inviability in hybrids between *Drosophila melanogaster* and *D. simulans*. Genetics **150:** 1091–1103.

Curtis, C. J., 1982 The mechanism of hybrid male sterility from crosses in the *Anopheles gambiae* and *Glossina morsitans* complexes, pp. 290–312 in *Recent Developments in the Genetics of Disease Vectors*, edited by W. M. Steiner. Stipes Publishing, Champaign, IL.

Curtis, C. J., and J. Chalkey, 1979 Lack of recombination between the *X* chromosomes of different members of the *Anopheles gambiae* complex. Heredity **42:** 323–326.

Davidson, G., 1964 The five mating types of the *Anopheles gambiae* complex. Rivista di Malariologia **13:** 167–183.

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.

Davis, A. W., E. G. Noonburg and C.-I Wu, 1994 Evidence for complex genic interactions between conspecific chromosomes

underlying hybrid female sterility in the *Drosophila simulans* clade. Genetics **137:** 191–199.

della Torre, A., L. Merzagora, J. R. Powell and M. Coluzzi, 1997 Selective introgression of paracentric inversions between two sibling species of the *Anopheles gambiae* complex. Genetics **146:** 239–244.

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

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

Doerge, R. W., and G. A. Churchill, 1996 Permutation tests for multiple loci affecting a quantitative character. Genetics **142:** 285–294.

Haldane, J. B. S., 1922 Sex-ratio and unisexual sterility in hybrid animals. J. Genet. **12:** 101–109.

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.

Holt, R. A., G. M. Subramanian, A. Halpern, G. G. Sutton, R. Charlab *et al.*, 2002 The genome sequence of the malaria mosquito *Anopheles gambiae*. Science **298:** 129–149.

Hutter, P., 1997 Genetics of hybrid inviability in Drosophila. Adv. Genet. **36:** 157–185.

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

Lander, E., J. Abrahamson, A. Barlow, M. Daly, S. Lincoln *et al.*, 1987 Mapmaker: a computer package for constructing genetic linkage maps. Cytogenet. Cell Genet. **46:** 642.

Marchand, R. P., 1984 Field observations on swarming and mating in *Anopheles gambiae* mosquitoes in Tanzania. Netherlands J. Zool. **34:** 367–387.

Muller, H. J., 1940 Bearing of the Drosophila work on systematics, pp. 185–268 in *The New Systematics*, edited by J. Huxley. Oxford University Press, Oxford.

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

Navarro, A., and N. H. Barton, 2003 Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. Evolution **57:** 447–459.

Noor, M. A. F., K. L. Grams, L. A. Bertucci, Y. Almendarez, J. Reiland *et al.*, 2001a The genetics of reproductive isolation and the potential for gene exchange between *Drosophila pseudoobscura* and *D. persimilis* via backcross hybrid males. Evolution **55:** 512–521.

Noor, M. A. F., A. L. Cunningham and J. C. Larkin, 2001b Consequences of recombination rate variation on quantitative trait locus mapping studies: simulations based on the *Drosophila melanogaster* genome. Genetics **159:** 581–588.

Noor, M. A. F., K. L. Grams, L. A. Bertucci and J. Reiland, 2001c Chromosomal inversions and the reproductive isolation of species. Proc. Natl. Acad. Sci. USA **98:** 12084–12088.

Orr, H. A., 1992 Mapping and characterization of a speciation gene in Drosophila. Genet. Res. **59:** 73–80.

Orr, H. A., 1995 The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics **139:** 1805–1813.

Orr, H. A., and J. A. Coyne, 1989 The genetics of postzygotic isolation in the *Drosophila virilis* group. Genetics **121:** 527–537.

Orr, H. A., and S. Irving, 2001 Complex epistasis and the genetic basis of hybrid sterility in the *Drosophila pseudoobscura* Bogota-USA hybridization. Genetics **158:** 1089–1100.

Palopoli, M. F., and C.-I Wu, 1994 The genetics of hybrid male sterility between Drosophila sibling species: a complex web of epistasis is revealed in interspecific studies. Genetics **138:** 329–341.

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

Petrarca, V., J. C. Beier, F. Onyango, J. Koros, C. Asiago *et al.*, 1991 Species composition of the *Anopheles gambiae* complex (Diptera, Culicidae) at two sites in western Kenya. J. Med. Entomol. **28:** 307–313.

Presgraves, D. C., 2003 A fine-scale genetic analysis of hybrid incompatibilities in Drosophila. Genetics **163:** 955–972.

Rieseberg, L. H., J. Whitton and K. Gardner, 1999 Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics **152:** 713–727.

Tao, Y., S. Chen, D. L. Hartl and C. C. Laurie, 2003a 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.

Tao, Y., Z-B. Zeng, J. Li, D. L. Hartl and C. C. Laurie, 2003b Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. II. Mapping hybrid male sterility loci on the third chromosome. Genetics **164:** 1399–1418.

Touré, Y. T., V. Petrarca, S. Traoré, A. Coulibaly, H. M. Maiga *et al.*, 1998 The distribution and inversion polymorphism of chromosomally recognized taxa of the *Anopheles gambiae* complex in Mali, West-Africa. Parassitologia **40:** 477–511.

Tripet, F., Y. T. Toure, C. E. Taylor, D. E. Norris, G. Dolo *et al.*, 2001 DNA analysis of transferred sperm reveals significant levels of gene flow between molecular forms of Anopheles gambiae. Mol. Ecol. **10:** 1725–1732.

True, J. R., B. S. Weir and C. C. Laurie, 1996 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 A. T. Beckenbach, 1983 Evidence for extensive genetic differentiation between the sex-ratio and the standard arrangement of *Drosophila pseudoobscura* and *Drosophila persimilis* and identification of hybrid sterility factors. Genetics **105:** 71–86.

Wu, C.-I, N. A. 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.

Yen, J. H., and A. R. Barr, 1971 New hypothesis of the cause of cytoplasmic incompatibility in *Culex pipiens* L. Nature **232:** 657–658.

Yen, J. H., and A. R. Barr, 1973 The etiological agent of cytoplasmic incompatibility in *Culex pipiens*. J. Invertebr. Pathol. **22:** 242–250.

Zeng, Z-B., 1994 Precision mapping of quantitative trait loci. Genetics **136:** 1457–1468.

Zheng, L. B., M. O. Benedict, A. J. Cornel, F. H. Collins and F. C. Kafatos, 1996 An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. Genetics **143:** 941–952.

Zhou, W. G., F. Rousset and S. O'Neill, 1998 Phylogeny and PCR based classification of Wolbachia strains using wsp gene sequences. Proc. R. Soc. Lond. Ser. B Biol. Sci. **265:** 509–515.

Zouros, E., K. Lofdahl and P. A. Martin, 1988 Male hybrid sterility in Drosophila: interactions between autosomes and sex-chromosomes in crosses of *Drosophila mojavensis* and *Drosophila arizonensis*. Evolution **42:** 1321–1331.

Communicating editor: M. A. F. Noor