

## Multilocus Evolution in Fire Ants: Effects of Selection, Gene Flow and Recombination

Kenneth G. Ross

Department of Entomology, University of Georgia, Athens, Georgia 30602-2603

Manuscript received September 3, 1996

Accepted for publication December 16, 1996

### ABSTRACT

The reproductive success of individual fire ant queens (*Solenopsis invicta*) previously has been shown to be strongly influenced by their genotype at a single enzyme-encoding gene, designated *Pgm-3*. This paper presents evidence that a second, tightly linked gene, designated *Gp-9*, is under similarly strong selection in these ants. Selection appears to act independently on the two genes and is detectable in only one of the two social forms of this species (the "polygyne" social form, in which nests contain multiple fertile queens). Strong directional selection on *Pgm-3* in this form involves worker destruction of all queens with genotype *Pgm-3<sup>AA</sup>* before they reproduce. Selection on *Gp-9* is more complex, involving both lethality of all *Gp-9<sup>b</sup>* females and a strong or even complete survival advantage to reproductive queens with the heterozygous genotype *Gp-9<sup>ab</sup>*. *Pgm-3* and *Gp-9* are tightly linked ( $r_f = 0.0016$ ) and exhibit strong gametic phase disequilibrium in introduced populations in the U.S. This disequilibrium seems not to have stemmed from the founder event associated with the introduction, because the same associations of alleles found in the U.S. apparently occur also in two native populations in Argentina. Rather, selection acting independently on *Pgm-3* and *Gp-9*, in conjunction with gene flow from the alternate, "monogyne" social form (in which nests contain a single fertile queen), may explain the origin of disequilibrium between the two loci in polygyne fire ants.

**E**VOLUTIONARY geneticists seek to understand how the forces of evolution influence distributions of fitness values and genetic variation in wild populations. Understandably, empirical work on the subject is dominated by studies that focus on only one evolutionary force. Relatively few examples exist of more complex systems in which the roles of two or more major forces have been analyzed simultaneously (e.g., McNEILLY 1968; ANTONOVICS *et al.* 1971; KOEHN *et al.* 1983; RIECHERT 1993; CHEVILLON *et al.* 1995; KING and LAWSON 1995; reviewed in HEDRICK 1985; BARTON and CLARK 1990). Fewer still involve analysis of the effects of several evolutionary forces on complex traits that depend on several genetic factors for their expression (e.g., CLEGG *et al.* 1976; JONES *et al.* 1977; TURNER 1977; ENNOS 1981; SILVER 1985; MALLET *et al.* 1990; LITTLE 1991; MORITA *et al.* 1992; HEDRICK 1994; HOLLOCHER and TEMPLETON 1994; PARHAM and OHTA 1996). Interest in such complex systems of multilocus evolution stems in part from concerns over the adequacy of single-locus theory and data for depicting evolutionary processes (LEWONTIN 1974; BIJLSMA 1978; HEDRICK *et al.* 1978; HEDRICK 1985; BERLOCHER 1993). These concerns can be addressed by considering the following issues in particular model organisms. To what extent is variation at different genes associated? What are the

causes of such associations? How do such associations affect distributions of fitness values? How do they influence genetic variation at the population level?

This paper considers these basic issues of multilocus evolution using the fire ant *Solenopsis invicta* as a model. The study builds on recent progress made in understanding how selection and gene flow jointly influence a simple genetic polymorphism in these ants. This genetic polymorphism has evolved in the context of a distinctive social polymorphism involving the number of reproductive (egg-laying) queens per colony. In the "polygyne" social form of *S. invicta*, in which colonies contain multiple reproductive queens, strong directional selection acts on the enzyme-encoding gene *Pgm-3* (*phosphoglucosyltransferase-3*). Nonreproductive females (the sterile workers and young queens) commonly possess a homozygous genotype designated as *Pgm-3<sup>AA</sup>*, but this genotype is never found in reproductive queens (ROSS 1992; ROSS *et al.* 1996a). Laboratory studies have revealed that *Pgm-3<sup>AA</sup>* is absent because workers selectively destroy all queens possessing it at an early age (KELLER and ROSS 1993). Remarkably, this strong selection is restricted to the polygyne form; genotype *Pgm-3<sup>AA</sup>* occurs at high frequencies in the single reproductive queens found in nests of the alternate "monogyne" social form (ROSS 1992; KELLER and ROSS 1993; ROSS *et al.* 1996a). Allele *Pgm-3<sup>A</sup>* is thought to persist at moderate frequencies in the polygyne form in spite of strong negative selection against it because of extensive gene flow from the monogyne form (ROSS 1992; ROSS and

Address for correspondence: Kenneth G. Ross, Department of Entomology, University of Georgia, Athens, GA 30602-2603.  
E-mail: ross@bscr.uga.edu

SHOEMAKER 1993; ROSS and KELLER 1995a). Thus, *Pgm-3* gene frequency dynamics in polygyne *S. invicta* are strongly influenced by the combined effects of both selection and gene flow.

In this paper I show that the influence of these two forces on *Pgm-3* in *S. invicta* is more complicated than previously recognized because of the involvement of a second, closely linked gene. This second gene, designated *Gp-9* (*general protein-9*), appears also to be under strong selection, although this selection is more complex and of a different form than that acting on *Pgm-3*. Significantly, the alleles of *Gp-9* are not randomly associated with those of *Pgm-3* (large linkage disequilibria exist), so that neither the transmission of alleles nor the population gene frequency dynamics are independent for the two loci. The goals of this paper are to provide an initial description of the genetics of this complex system and to use this information to begin to address pertinent issues concerning multilocus evolution in wild populations of these ants.

#### MATERIALS AND METHODS

**Sample collection:** Samples were collected from introduced (U.S.) and native (Argentina) populations of *S. invicta*, and were obtained from nests of both social forms at most sampling sites. Ants were placed in liquid nitrogen or dry ice in the field and were held at  $-80^{\circ}$  in the laboratory pending electrophoresis.

In the United States, samples were obtained mostly from northern Georgia (Putnam and Walton Cos.). Sexu-als [adult males or nonreproductive (winged) queens] and workers were collected from each putative monogyne colony sampled in Georgia. Monogyne initially was inferred in the field using well established criteria that distinguish nests of the two forms (GREENBERG *et al.* 1985). It was confirmed subsequently in most nests by genotyping eight or more sexuals from each nest at four polymorphic allozyme loci to verify that the genotype arrays conformed to those expected for the offspring of a single, monandrous queen (*S. invicta* queens mate once; ROSS and FLETCHER 1985; SHOEMAKER *et al.* 1992; ROSS *et al.* 1993). Polygyne was confirmed for all putative polygyne nests sampled in the United States by capturing several mated reproductive queens per nest. Workers were collected from all polygyne nests sampled in Georgia, and nonreproductive (winged) and reproductive (wingless) queens were collected opportunistically from these same nests.

Additional samples of introduced fire ants were collected from two sites outside of Georgia. Newly mated monogyne queens (which had just completed a mating flight and were preparing to found colonies) were obtained from Leon Co., Florida, a site over 400 km from the Georgia study site. Polygyne has not been observed at this well studied Florida site (PORTER 1992; TSCHINKEL 1993). Reproductive polygyne queens were collected from their nests in Travis Co., Texas, a site over 1400 km from the Georgia site.

Samples from Argentina were collected from two geographic populations centered around the cities of Corrientes (Corrientes Province) and Formosa (Formosa Province). These populations, which are separated by a substantial biogeographic barrier (the Río Paraná), are genetically divergent at several nuclear loci and their mitochondrial DNA (ROSS *et al.* 1996a; K. G. ROSS, unpublished data). Nests of both social forms occur at each site. Workers were collected from all

sampled nests, and queens (nonreproductive and reproductive) were collected opportunistically. Initial determination of the social form of each nest in Argentina was made using field criteria. These determinations were verified for every nest by inspecting the genotype arrays of 12 or more female nestmates at 10 polymorphic allozyme loci (ROSS *et al.* 1993).

**Scoring of genotypes at *Pgm-3* and *Gp-9*:** Electrophoresis was conducted on horizontal gels of 14% starch, with visualization of the banding phenotypes accomplished using specific histochemical staining in the case of *Pgm-3* and nonspecific protein staining in the case of *Gp-9* (SHOEMAKER *et al.* 1992). The products of both loci behave as monomers, with two staining bands of equal intensity in heterozygotes and single bands in homozygotes and (*Pgm-3* only) haploid males. The locus *Pgm-3* is expressed and scorable in all life stages of all individuals. The protein product of *Gp-9* is not detectable in brood or in males of any developmental stage with current protocols. Furthermore, it only becomes detectable in adult females at a minimum posteclosion age of 8–14 days (data not shown).

Progeny studies have confirmed that the electromorphs of *Pgm-3* are inherited in Mendelian fashion (ROSS 1992; SHOEMAKER *et al.* 1992). This locus is biallelic in *S. invicta* in the United States, whereas several additional rare alleles occur in Argentina, the most common of these at a frequency of  $<0.05$  (ROSS *et al.* 1996a). The two common allelic variants are designated here as *Pgm-3<sup>A</sup>* and *Pgm-3<sup>B</sup>* (corresponding to *Pgm-3<sup>a</sup>* and *Pgm-3<sup>b</sup>*, respectively, of earlier studies). Data presented below indicate that the products of *Gp-9* also exhibit Mendelian inheritance. Only two allelic variants have been detected in *S. invicta*; these are designated as *Gp-9<sup>B</sup>* and *Gp-9<sup>C</sup>*.

Both loci were scored in the same individuals in most samples. The exceptions are workers from the Argentina populations, where *Pgm-3* was scored in worker pupae and *Gp-9* in worker adults. Thus, two-locus genotypes are not available for these individuals.

**Estimation of genotype and allele frequencies:** Nests of fire ants comprise more or less complex families, so that multiple genotypes scored from a given nest often are not independent. To obtain unbiased genotype and allele frequencies using all collected data, a resampling procedure was employed whereby a single one- or two-locus genotype was drawn at random from each nest in a sample, yielding a distribution of independent genotypes. Genotype and allele frequencies were estimated as the arithmetic mean frequencies for 1000 such resampled genotype distributions from each original sample. The 95% confidence intervals (CIs) about the frequency estimates for *Gp-9* were obtained by dropping the 25 highest and 25 lowest frequency values derived from the 1000 resampled distributions.

**Family studies:** Sixty mated queens originating from 12 polygyne nests from Walton Co., Georgia were held individually in small rearing chambers in the laboratory with several hundred worker adults and brood originating from the same parent nest (*e.g.*, ROSS and FLETCHER 1985). These rearing units were provided abundant food and water over a period of 8 weeks, at which time all worker pupae present were known to be the offspring of the resident queen (based on reported brood developmental periods; O'NEAL and MARKIN 1975). Several dozen worker pupae and their mother then were removed from each unit and transferred into secondary isolation units. After all workers had eclosed, each secondary unit was held for an additional week to ensure that the workers were sufficiently mature that *Gp-9* was expressed. The *Pgm-3* and *Gp-9* genotypes of all 60 queens and of  $\sim 30$  worker offspring of each queen were determined.

**Estimation of disequilibrium:** Values of the Hardy-Weinberg disequilibrium coefficient ( $D_A$ ), which measures depar-

tures of genotype proportions observed at single loci from those expected under Hardy-Weinberg equilibrium (HWE), were estimated for *Gp-9* for all classes of females in all study populations using Equation 3.1 of WEIR (1996). Statistical significance of the estimates was determined by calculating the chi-squared test statistic (WEIR 1996, Equation 3.6).

Linkage disequilibrium between *Pgm-3* and *Gp-9* was measured for females in each population when data for both loci were available from the same individuals. Values of the gametic disequilibrium coefficient ( $D_{AB}$ ), which measures the association between alleles at the two loci on the same chromosome, were estimated for polygyne queens, their mates, and their worker offspring from the family studies using the procedure of WEIR (1996, p. 113). This was possible because monandry, coupled with male haploidy, permits complete specification of all parental and offspring haplotypes when crossing over is infrequent (as it is between *Pgm-3* and *Gp-9*, see below). Furthermore, because double heterozygote queens and workers in the families almost invariably exhibited the *cis* (coupling) linkage phase (see below), two-locus haplotypes could be inferred directly from two-locus genotypes in the Georgia polygyne population by assuming that *trans* (repulsion) double heterozygotes are absent in the wild. Thus, gametic disequilibrium could be estimated directly from genotypic data for this population. Estimates of  $D_{AB}$  could not be obtained directly from genotypic data in other populations because the requisite information on the linkage phase of double heterozygotes is lacking. Nor could this coefficient be estimated indirectly using the maximum likelihood approach of HILL (1974), because departures from HWE observed at both loci violate the assumptions of this method (WEIR and COCKERHAM 1979).

Values of the composite genotypic disequilibrium coefficient ( $\Delta_{AB}$ ), which measures the summed effects of both the gametic and nongametic components of linkage disequilibrium, were estimated from genotypic data for females in all populations (WEIR 1996, p. 126). In the case of the Georgia polygyne population, where  $D_{AB}$  values were available, the nongametic disequilibrium coefficient ( $D_{A/B}$ ) also could be estimated, using the relationship  $D_{A/B} = \Delta_{AB} - D_{AB}$ . This nongametic coefficient measures the nonrandom association between alleles at the two loci on alternate homologous chromosomes within diploid individuals. Significance of all three linkage disequilibrium coefficients was evaluated by calculating the appropriate chi-squared test statistics (WEIR 1996, pp. 113–114, 125–127). Values of the disequilibrium coefficients were normalized according to the specific allele frequencies for that class of individuals (HEDRICK 1987; WEIR 1996), yielding standardized coefficients designated as  $\Delta'_{AB}$ ,  $D'_{AB}$ , and  $D'_{A/B}$  that were employed for comparisons across life stages and populations. The rare alleles at *Pgm-3* in the Argentina populations were pooled with allele *Pgm-3<sup>a</sup>* for all of the disequilibrium analyses.

The potential nonindependence of genotypes scored from single nests creates a problem in specifying sample sizes for calculating the disequilibrium test statistics (e.g., CROZIER *et al.* 1987). The appropriate sample size varies between the total number of genotypes scored (when nestmate relatedness is low) and the number of nests from which these genotypes were obtained (when nestmate relatedness is high). A reasonable approximation of the number of independent diploid genomes sampled ( $n$ ) can be obtained for female social Hymenoptera using the formula  $n = (3/4r) \cdot (3N/2)$ , where  $r$  is the population-wide average nestmate relatedness for the relevant class of individuals and  $N$  is the number of nests sampled (PAMILO 1983; SUNDRÖM 1993). The upper limit for this adjusted sample size is taken to be the total number of genotypes sampled. Estimates of nestmate relatedness, pre-

viously obtained using allozyme markers for all classes of individuals in all populations studied here (ROSS and FLETCHER 1985; ROSS 1993; ROSS *et al.* 1996b), were used to calculate adjusted sample sizes for evaluating disequilibrium.

## RESULTS

**Genotype and allele frequencies at *Gp-9*:** Estimates of the frequencies of *Gp-9* genotypes and alleles in *S. invicta* are presented in Tables 1 and 2 for the monogyne and polygyne social forms, respectively. In the monogyne form, the allele *Gp-9<sup>B</sup>* is fixed in nearly all study populations (Table 1). The exception is the Corrientes population, where one of the 37 monogyne colonies surveyed contained approximately equal numbers of females with the genotypes *Gp-9<sup>BB</sup>* and *Gp-9<sup>Bb</sup>*, suggesting that the mother queen was a heterozygote.

Genotype and allele frequencies at *Gp-9* in the polygyne form differ strikingly from those in the monogyne form (Table 2). The locus exhibits substantial polymorphism in polygyne populations, with the frequency of allele *Gp-9<sup>b</sup>* (which is essentially absent in the monogyne form) ranging from 0.10 in nonreproductive queens from Corrientes to 0.50 in reproductive queens from the United States. Two remarkable patterns are evident in the *Gp-9* genotype distributions in the polygyne form. First, the genotype *Gp-9<sup>bb</sup>* is completely absent in the Argentine populations and virtually so in the United States. Given the large sample sizes and substantial frequencies of allele *Gp-9<sup>b</sup>* in these populations, the near absence of *Gp-9<sup>bb</sup>* females cannot be attributed to sampling error. Indeed, the binomial probability of sampling as few *Gp-9<sup>bb</sup>* females as were found in the United States is  $<0.0001$  if this genotype actually occurs at the frequency expected in the absence of selection (predicted from the family data to be 0.05; see below). The virtual absence of genotype *Gp-9<sup>bb</sup>* in the wild thus suggests the presence of strong viability selection against it.

The second notable pattern in the *Gp-9* data from polygyne populations is the pronounced and consistent elevation in frequency of genotype *Gp-9<sup>Bb</sup>* in reproductive queens compared to nonreproductive queens and, especially, workers (Table 2). These heterozygotes are represented in reproductive queens at about twice the frequency as in workers in all polygyne populations. In the United States, essentially all reproductive queens are heterozygotes. This elevated frequency of genotype *Gp-9<sup>Bb</sup>* in reproductive queens corresponds to substantially higher frequencies of allele *Gp-9<sup>b</sup>* in these queens than in nonreproductive females (Table 2). The persistent elevation in heterozygosity in polygyne reproductive queens is consistent with some advantage of genotype *Gp-9<sup>Bb</sup>* over *Gp-9<sup>BB</sup>* in queens attempting to make the transition from nonreproductive to reproductive status. Thus, locus *Gp-9* appears to be subject to overdominant selection in polygyne *S. invicta* colonies; viability selection removes all *Gp-9<sup>bb</sup>* females at some time during development while an unknown mechanism

**TABLE 1**  
**Genotype and allele frequencies at *Gp-9* in the monogyne social form of *S. invicta* in the United States and Argentina**

	Genotype frequencies			Allele frequencies	
	<i>BB</i>	<i>Bb</i>	<i>bb</i>	<i>B</i>	<i>b</i>
Georgia, U.S.					
Workers [ <i>N</i> = 82, <i>n</i> = 82]	1.0	0	0	1.0	0
Nonreproductive queens [ <i>N</i> = 466, <i>n</i> = 695]	1.0	0	0	1.0	0
Florida, U.S.					
Nonreproductive queens [ <i>n</i> = 130] <sup>a</sup>	1.0	0	0	1.0	0
Corrientes, Argentina					
Workers [ <i>N</i> = 37, <i>n</i> = 443]	0.981 (0.935–1.0)	0.019 (0–0.065)	0	0.990 (0.973–1.0)	0.010 (0–0.027)
Nonreproductive queens [ <i>N</i> = 24, <i>n</i> = 134]	0.958 (0.875–1.0)	0.042 (0–0.125)	0	0.979 (0.937–1.0)	0.021 (0–0.063)
Formosa, Argentina					
Workers [ <i>N</i> = 35, <i>n</i> = 417]	1.0	0	0	1.0	0
Nonreproductive queens [ <i>N</i> = 28, <i>n</i> = 152]	1.0	0	0	1.0	0

Sample sizes (in brackets) are for the numbers of nests (*N*) and individuals (*n*) from which genetic data were obtained. The 95% confidence intervals are in parentheses.

<sup>a</sup> Individuals were not collected in association with nests, so only *n* is reported.

hinders or precludes *Gp-9<sup>BB</sup>* queens from becoming egg layers.

A telltale effect of this proposed selection on *Gp-9* in polygyne *S. invicta* is the deviations of the observed genotype proportions from those expected under Hardy-Weinberg equilibrium (HWE). Estimates of the Hardy-Weinberg disequilibrium coefficient (*D<sub>A</sub>*) for *Gp-*

*9* were significantly negative (indicating excess heterozygosity) for reproductive queens from all populations, as well as for the other classes of females from all populations except Corrientes ( $\chi^2$  tests, all *P* < 0.025). Consistent excess heterozygosity is predicted because of selection against one or both homozygotes at *Gp-9*.

**Inheritance of *Gp-9*:** The unusual distributions of *Gp-*

**TABLE 2**  
**Genotype and allele frequencies at *Gp-9* in the polygyne social form of *S. invicta* in the United States and Argentina**

	Genotype frequencies			Allele frequencies	
	<i>BB</i>	<i>Bb</i>	<i>bb</i>	<i>B</i>	<i>b</i>
Georgia, U.S.					
Workers [ <i>N</i> = 181, <i>n</i> = 406]	0.407 (0.346–0.460)	0.591 (0.540–0.654)	0.002 (0–0.002)	0.702 (0.673–0.730)	0.298 (0.270–0.327)
Nonreproductive queens [ <i>N</i> = 95, <i>n</i> = 279]	0.202 (0.158–0.252)	0.798 (0.748–0.842)	0	0.601 (0.579–0.626)	0.399 (0.374–0.421)
Reproductive queens [ <i>N</i> = 215, <i>n</i> = 1535] <sup>a</sup>	0	0.999	0.001	0.500	0.500
Texas, U.S.					
Reproductive queens [ <i>N</i> = 30, <i>n</i> = 451]	0	0.994 (0.966–1.0)	0.006 (0–0.034)	0.497 (0.483–0.500)	0.503 (0.500–0.517)
Corrientes, Argentina					
Workers [ <i>N</i> = 44, <i>n</i> = 514]	0.777 (0.660–0.886)	0.223 (0.114–0.340)	0	0.889 (0.830–0.943)	0.111 (0.057–0.170)
Nonreproductive queens [ <i>N</i> = 19, <i>n</i> = 175]	0.792 (0.684–0.894)	0.208 (0.106–0.316)	0	0.896 (0.842–0.947)	0.104 (0.053–0.158)
Reproductive queens [ <i>N</i> = 28, <i>n</i> = 187]	0.530 (0.500–0.536)	0.470 (0.464–0.500)	0	0.765 (0.750–0.768)	0.235 (0.232–0.250)
Formosa, Argentina					
Workers [ <i>N</i> = 35, <i>n</i> = 399]	0.682 (0.572–0.800)	0.318 (0.200–0.428)	0	0.841 (0.786–0.900)	0.159 (0.100–0.214)
Nonreproductive queens [ <i>N</i> = 23, <i>n</i> = 171]	0.628 (0.522–0.740)	0.372 (0.260–0.478)	0	0.814 (0.761–0.870)	0.186 (0.130–0.239)
Reproductive queens [ <i>N</i> = 26, <i>n</i> = 121]	0.385 (0.200–0.577)	0.615 (0.423–0.800)	0	0.692 (0.600–0.789)	0.308 (0.211–0.400)

Sample sizes (in brackets) are for the numbers of nests (*N*) and individuals (*n*) from which genetic data were obtained. The 95% confidence intervals are in parentheses.

<sup>a</sup> Informative 95% confidence intervals cannot be obtained for these individuals.

9 genotypes raise concerns about whether the variants scored are actually the products of a single codominant Mendelian locus. One line of evidence often marshalled to support this model in Hymenoptera, the presence of single staining bands in haploid males, is not available for *Gp-9* because no gene product is detectable in males. Nonetheless, several other lines of evidence suggest that this model of inheritance is valid. For instance, all three putative genotypes have been found in the wild following an intense sampling effort. Also, the occurrence of rare triploid workers in the Georgia polygyne population, indicated by uneven intensities of the staining bands at several monomeric allozyme loci, is revealed also by uneven band intensities for the products of *Gp-9* (unpublished data).

The most compelling evidence for Mendelian inheritance of *Gp-9* comes from the family studies. Because all 60 polygyne mother queens heading these families were *Gp-9<sup>bb</sup>* heterozygotes, the diploid offspring of each queen should comprise an even mix of heterozygotes and of homozygotes of one class. All 60 progenies indeed were mixes of heterozygotes and individuals with a single homozygous genotype. Furthermore, the two genotypic classes occurred in ratios indistinguishable from a 1:1 ratio in all but five progenies (binomial tests with probability values corrected using the sequential Bonferroni procedure; HOCHBERG 1988). Significantly, the genotype underrepresented in each of these five progenies was *Gp-9<sup>bb</sup>* (the segregation ratio in a sixth family possessing *Gp-9<sup>bb</sup>* and *Gp-9<sup>Bb</sup>* offspring did not differ significantly from 0.50). Thus, data from the family studies not only are consistent with *Gp-9* representing a single codominant Mendelian locus, they also reveal the low viability of genotype *Gp-9<sup>bb</sup>* under controlled conditions [although the survival of this genotype in laboratory families was substantially higher than in the wild (0.018 *vs.* 0.002)].

**Linkage of *Pgm-3* and *Gp-9*:** The association of *Gp-9* genotype with reproductive role in polygyne *S. invicta* bears some similarities to the previously reported influence of *Pgm-3* genotype. Thus, the relationship between the two genes, in terms of both their frequency of recombination (linkage) and patterns of association between alleles and genotypes in the wild (linkage disequilibrium), is of interest.

The family data permit specification of the two-locus haplotype composition of parents and offspring, and therefore estimation of the recombination frequency, by the following means. The two-locus genotypes of the polygyne mother queens and their daughters were obtained for each family. Because each queen mates with a single haploid male, the array of genotypes in her daughters can be compared with the mother's genotype to infer the sole two-locus haplotype of the father's gametes. The father's haplotypic contribution then can be subtracted from each daughter's two-locus genotype to infer the mother's haplotypic contribution to each

offspring. Recombinant offspring of double heterozygote mothers stand out in a large progeny when recombination frequencies are substantially <0.5 because of their unusual maternal haplotypic endowment (see EPERSON and ALLARD 1987; NEJATI-JAVAREMI and SMITH 1996).

Forty-five of the 60 polygyne queens heading families were double heterozygotes. Among the 1290 worker offspring of these queens successfully reared to maturity and genotyped, only two were inferred to be recombinants. The frequency of recombination between *Gp-9* and *Pgm-3* in queens ( $r_f$ ) thus is estimated to be 0.16%. Because the two genes are not inherited independently, the associations of their genotypes and alleles in the wild were next examined.

**Two-locus genotype frequencies:** Observed two-locus genotype frequencies for nonreproductive females from the polygyne populations are presented in Table 3 along with the frequencies expected if *Pgm-3* and *Gp-9* were independent (the products of the observed frequencies at each locus). The "observed" frequencies for zygotes in the Georgia population are calculated from the genotype distributions in the 60 families by assuming that *Gp-9<sup>Bb</sup>* and *Gp-9<sup>bb</sup>* workers actually segregate in a 1:1 ratio in the families in which they occur. Thus, these represent projected frequencies in zygotes before viability selection acts against *Gp-9<sup>bb</sup>*. The most prominent departures from expected frequencies in these zygotes involve excesses of genotypes *Pgm-3<sup>AA</sup> Gp-9<sup>BB</sup>* and *Pgm-3<sup>Aa</sup> Gp-9<sup>Bb</sup>*, as well as deficiencies of genotypes *Pgm-3<sup>AA</sup> Gp-9<sup>bb</sup>* and *Pgm-3<sup>Aa</sup> Gp-9<sup>Bb</sup>*. These patterns are suggestive of strong positive gametic disequilibrium, as confirmed below, because genotypes that can be composed of haplotypes *A B* or *a b* are in excess, whereas those that require haplotypes *A b* or *a B* are in deficit. This same association of genotypes at the two loci persists in the adult workers and nonreproductive queens, that is, after selection against *Gp-9<sup>bb</sup>* has occurred and selection against *Pgm-3<sup>AA</sup>* and *Gp-9<sup>BB</sup>* in queens has begun. The association is evident in the native (Argentina) as well as the introduced (U.S.) populations.

The observed frequencies of the two-locus genotypes are depicted graphically for polygyne females from the introduced populations in Figure 1 and from the native populations in Figure 2. Comparison of zygotes with the other classes from Georgia (Figure 1) reveals the effects of selection acting jointly on *Gp-9* and *Pgm-3* (the zygote data represent the inferred preselection condition). Inviability of *Gp-9<sup>bb</sup>* adults is evident in the drop in frequency of this genotype from 0.05 in zygotes to essentially zero in adults of all classes (see also Table 2). The invariant destruction of queens with genotype *Pgm-3<sup>AA</sup>* as they mature sexually, coupled with an all but absolute advantage to genotype *Gp-9<sup>bb</sup>* in queens attempting to become egg layers, results in a winnowing of the six genotypic classes present in zygotes to essen-

TABLE 3

Observed and expected two-locus genotype frequencies for *Pgm-3* and *Gp-9* in nonreproductive females of polygyne *S. invicta* in the United States and Argentina

<i>Pgm-3</i> :	AA			Aa			aa		
	<i>Gp-9</i> : BB	Bb	bb	BB	Bb	bb	BB	Bb	bb
Georgia, U.S.									
Zygotes	<b>0.283</b> (0.127)	<b>0</b> (0.142)	0 (0.014)	<b>0.133</b> (0.240)	<b>0.400</b> (0.267)	0 (0.027)	0.033 (0.083)	0.100 (0.092)	0.050 (0.009)
Adult workers	<b>0.256</b> (0.090)	<b>0</b> (0.137)	0 (0.001)	<b>0.154</b> (0.244)	<b>0.437</b> (0.370)	0 (0.002)	0.006 (0.062)	0.145 (0.094)	0.002 (0)
Nonreproductive queens	<b>0.133</b> (0.027)	<b>0</b> (0.106)	0 (0)	<b>0.067</b> (0.139)	<b>0.621</b> (0.548)	0 (0)	0.003 (0.036)	0.176 (0.143)	0 (0)
Corrientes, Argentina									
Nonreproductive queens	<b>0.490</b> (0.409)	<b>0.022</b> (0.103)	0 (0)	<b>0.305</b> (0.387)	<b>0.179</b> (0.097)	0 (0)	0.004 (0.003)	0 (0.001)	0 (0)
Formosa, Argentina									
Nonreproductive queens	<b>0.239</b> (0.151)	<b>0</b> (0.089)	0 (0)	<b>0.363</b> (0.450)	<b>0.351</b> (0.264)	0 (0)	0.027 (0.029)	0.019 (0.017)	0 (0)

The expected values (in parentheses) are the products of the genotype frequencies observed at each locus for that class of individuals. Two-locus genotype frequencies exhibiting consistent strong discrepancies between the observed and expected values are shown in bold.

tially two in the reproductive queens. This selective process that culminates in reproductive queens is evident already in young nonreproductive queens, in which frequencies of the two-locus genotypes involving *Pgm-3*<sup>AA</sup> or *Gp-9*<sup>bb</sup> are intermediate between those of workers and of reproductive queens. The two-locus genotype frequencies in reproductive queens from Texas are almost identical to those seen in Georgia reproductive queens (Figure 1), suggesting that similar processes of selection involving *Gp-9* and *Pgm-3* occur in association with polygyny in the two areas.

Several features of the data from Argentina suggest similarities to the United States, including the lack both of genotype *Gp-9*<sup>bb</sup> among adult females and of *Pgm-3*<sup>AA</sup> among reproductive queens (Figure 2). However, because the apparent selective advantage of *Gp-9*<sup>bb</sup> over *Gp-9*<sup>Bb</sup> is weaker in reproductive queens in Argentina than in the United States, a greater diversity of common two-locus genotypes is retained in such queens in the native than in the introduced populations. The magnitude of the heterozygote advantage at *Gp-9* in native queens can be quantified by computing the increase in frequency of *Gp-9*<sup>bb</sup> relative to *Gp-9*<sup>Bb</sup> between nonreproductive and reproductive queens (after omitting from consideration queens with genotype *Pgm-3*<sup>AA</sup>). Heterozygous reproductive queens are overrepresented by a factor of 1.53 in Corrientes and 1.69 in Formosa compared to heterozygous nonreproductive queens from the same sites.

**Linkage disequilibrium:** The ability to specify two-locus haplotype compositions of parents and offspring in the 60 families from the Georgia polygyne population means that values for the gametic disequilibrium coefficients ( $D_{AB}$ ,  $D'_{AB}$ ) can be obtained directly from these individuals. Estimates for these coefficients are

presented along with the observed and expected frequencies of the four haplotypes in the top part of Table 4. The haplotypes *A B* and *a b* are far more common than expected under random association of the alleles in the offspring zygotes and their parents, yielding highly significant positive values of  $D_{AB}$  near the maximum possible for the observed allele frequencies (indicated by values of  $D'_{AB}$  close to 1.0).

All double heterozygote queens and daughters from the family studies ( $n = 636$ ) exhibited the *cis* linkage phase (*A B/a b*), with the exception of the two recombinant daughters. Thus, *trans* double heterozygotes can be assumed to be essentially absent in the Georgia polygyne population and the haplotype composition of all genotypes in the wild can be inferred. Gametic disequilibrium values estimated in this way from the population genotypic data are presented in the bottom part of Table 4. For all classes of females, the haplotypes *A B* and *a b* again are far more common than expected if the alleles at *Pgm-3* and *Gp-9* were randomly associated, yielding highly significant positive values of  $D_{AB}$  that approach the maximum possible values.

The data in Table 4 suggest that haplotype *A b* is rare or nonexistent in the Georgia polygyne population. Consistent with this conclusion, the three genotypes that require this haplotype (*Pgm-3*<sup>AA</sup> *Gp-9*<sup>bb</sup>, *Pgm-3*<sup>AA</sup> *Gp-9*<sup>Bb</sup>, *Pgm-3*<sup>Aa</sup> *Gp-9*<sup>bb</sup>) essentially were absent in the families and in the wild. The absence of one of these, *Pgm-3*<sup>AA</sup> *Gp-9*<sup>Bb</sup>, in the 338 workers from the wild (Figure 1) lends particularly strong support to this conclusion, as there is no evidence of selection against either *Pgm-3*<sup>AA</sup> or *Gp-9*<sup>bb</sup> in workers. Gametic disequilibrium is expected to attain its maximal positive value when one of the *trans* haplotypes is missing (HEDRICK 1985, Equation 10.4), as was found for this population. Significantly,

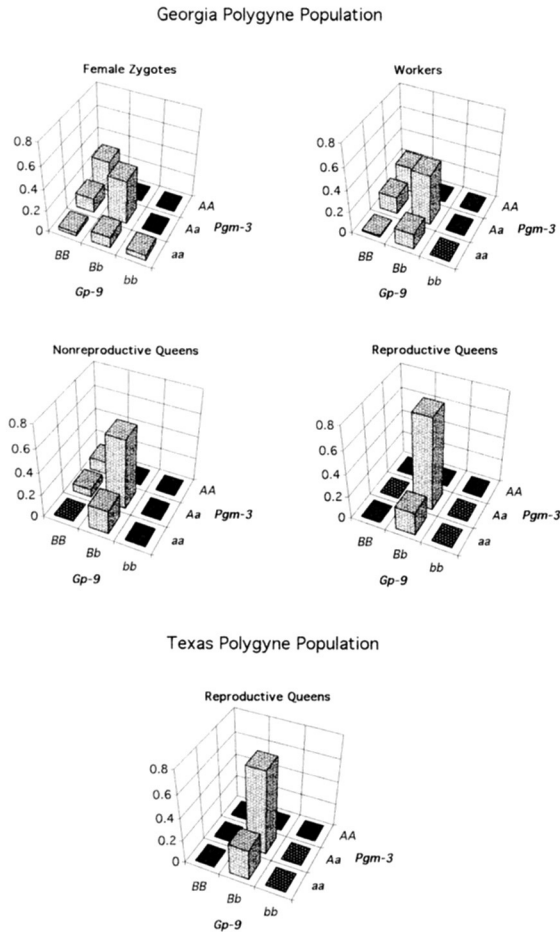


FIGURE 1.—Observed two-locus genotype frequencies for *Pgm-3* and *Gp-9* in various classes of females from introduced populations of polygyne *S. invicta*. The frequencies for female zygotes from Georgia are derived from family studies and represent projected values before selection on *Gp-9* occurs. Darkly stippled patches represent genotypes present at frequencies < 0.01; black patches indicate genotypes that were not found.

the absence of genotype  $Pgm-3^{AA} Gp-9^{bb}$  in nonreproductive queens from the Formosa polygyne population (Figure 2) suggests that haplotype  $A b$  may be missing in this native population as well.

The nongametic component of linkage disequilibrium ( $D_{A/B}$ ) was estimated in the Georgia polygyne population by calculating the composite genotypic disequilibrium ( $\Delta_{AB}$ ) and employing the formula  $D_{A/B} = \Delta_{AB} - D_{AB}$ . For zygotes, the genotype frequencies are those expected if  $Gp-9^{bb}$  and  $Gp-9^{bb}$  segregated in 1:1 ratios in the families in which they occurred; that is, they represent the preselection frequencies.

Estimates of the magnitude of all three forms of disequilibrium in the Georgia polygyne population are presented as the standardized measures  $\Delta'_{AB}$ ,  $D'_{AB}$ , and  $D'_{A/B}$  in Figure 3. Values for  $D'_{AB}$  consistently are near their positive maxima, and changes in  $\Delta'_{AB}$  from one class of individuals to another correspond to changes of opposite sign in  $D'_{A/B}$ . In reproductive queens, the

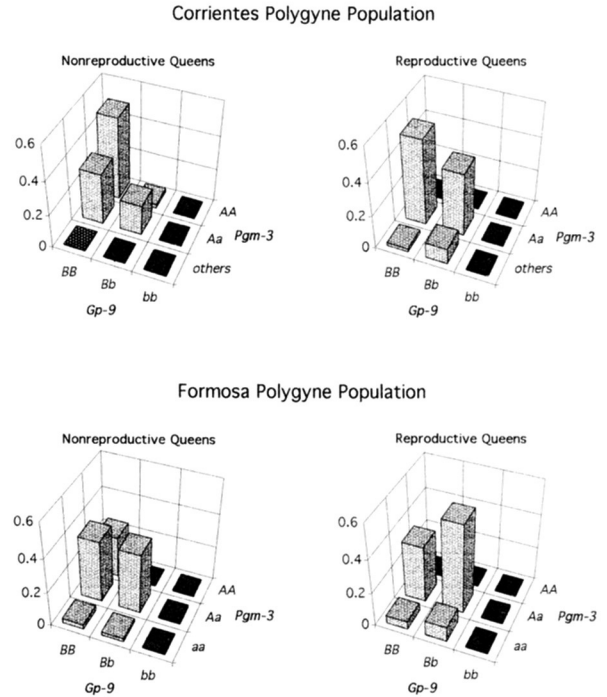


FIGURE 2.—Observed two-locus genotype frequencies for *Pgm-3* and *Gp-9* in nonreproductive and reproductive queens from native populations of polygyne *S. invicta*. The class “others” for *Pgm-3* in Corrientes includes the genotype  $Pgm-3^{aa}$  pooled with all genotypes at this locus carrying rare alleles. Darkly stippled patches represent genotypes present at frequencies < 0.01; black patches indicate genotypes that were not found.

estimated value of zero for the composite measure ( $\Delta'_{AB}$ ) corresponds to a large negative value for nongametic disequilibrium ( $D'_{A/B}$ ) that offsets the large positive gametic disequilibrium ( $D'_{AB}$ ). Thus, in reproductive queens, allele  $A$  tends to be strongly associated with allele  $B$  (and  $a$  with  $b$ ) within chromosomes, but the converse associations ( $A$  with  $b$  and  $a$  with  $B$ ) exist between alleles at the two loci on alternate homologous chromosomes. In the zygotes produced by these queens,  $D'_{A/B}$  increases to nonsignificance, resulting in a significant positive value for the composite measure. This dramatic change in pattern of nongametic disequilibrium between zygotes and their mothers corresponds with an episode of gene flow from the monogyne form (see below). In adult workers,  $D'_{A/B}$  becomes increasingly (and significantly) negative relative to zygotes, indicating a buildup of negative nongametic allelic associations following selection against genotype  $Gp-9^{bb}$ . The trend toward greater negative values of  $D'_{A/B}$  continues in nonreproductive queens, in which selection against genotypes  $Pgm-3^{AA}$  and  $Gp-9^{BB}$  begins. It culminates as these queens are recruited as new reproductives, following complete elimination of any remaining individuals with either genotype. The pattern of disequilibrium in the Georgia polygyne population thus can be summarized as invariant large positive gametic disequilibrium

TABLE 4

Observed and expected two-locus haplotype frequencies and estimates of gametic disequilibrium ( $D_{AB}$ ,  $D'_{AB}$ ) for *Pgm-3* and *Gp-9* in polygyne *S. invicta* from Georgia

	Haplotypes				$D_{AB}$	$D'_{AB}$	$\chi^2$
	<i>A B</i>	<i>A b</i>	<i>a B</i>	<i>a b</i>			
Laboratory families							
Zygotes [ <i>n</i> = 240]	0.550 (0.385)	0 (0.165)	0.150 (0.315)	0.300 (0.135)	0.165	0.996	125.7 <i>P</i> < 0.005
Mother queens [ <i>n</i> = 120]	0.383 (0.192)	0 (0.192)	0.117 (0.308)	0.500 (0.308)	0.192	0.992	74.6 <i>P</i> < 0.005
Fathers [ <i>n</i> = 60]	0.717 (0.645)	0 (0.072)	0.183 (0.255)	0.100 (0.028)	0.072	0.983	16.9 <i>P</i> < 0.005
Population samples							
Workers [ <i>n</i> = 676]	0.536 (0.373)	0 (0.162)	0.161 (0.324)	0.303 (0.141)	0.162	0.998	339.2 <i>P</i> < 0.005
Nonreproductive queens [ <i>n</i> = 558]	0.450 (0.266)	0 (0.184)	0.141 (0.325)	0.409 (0.225)	0.184	0.998	315.2 <i>P</i> < 0.005
Reproductive queens [ <i>n</i> = 2914]	0.394 (0.197)	<0.001 <sup>a</sup> (0.197)	0.106 (0.303)	0.500 (0.303)	0.196	0.999	1785.9 <i>P</i> < 0.005

The sample sizes (in brackets) are total counts of haplotypes for each class of individuals. Expected haplotype frequencies (in parentheses) are the products of the allele frequencies. Probability values from the chi-squared test statistics indicate the probabilities that the alleles *Pgm-3*<sup>A</sup> and *Gp-9*<sup>B</sup> (or *Pgm-3*<sup>a</sup> and *Gp-9*<sup>b</sup>) are randomly associated on chromosomes.

<sup>a</sup> A single haplotype of this kind was identified (frequency = 0.00034).

in all classes of individuals, coupled with nongametic disequilibrium that becomes increasingly negative following each bout of selection and that is reset to almost zero following gene flow from the genetically differentiated monogyne form.

The linkage phase of double heterozygotes is not known for the polygyne ants from Texas or Argentina, so only composite disequilibrium coefficients can be estimated from the genotypic data. These estimates parallel the results obtained from the Georgia population,

with small and nonsignificant estimates of  $\Delta'_{AB}$  in reproductive queens and moderately positive but significant estimates of this coefficient in nonreproductive queens. If haplotype *A b* is assumed to be missing in Formosa (see above), then patterns for all three disequilibrium coefficients for nonreproductive and reproductive queens are identical to those found in Georgia.

**Gene flow:** Earlier work has suggested that gene flow into the polygyne population from the monogyne form is extensive in Georgia, and that the major route for this interform gene flow is via polygyne queens mating with monogyne males (ROSS 1992; ROSS and SHOEMAKER 1993; ROSS and KELLER 1995a). These studies have made use of the large differences in *Pgm-3* allele frequencies that result from differential selection between the social forms to calibrate the magnitude of gene flow occurring via this route. Because allele frequencies at *Gp-9* also differ substantially between the forms, gene flow from the monogyne form is certain to have an impact on allele and genotype frequencies at *Gp-9* in the polygyne form. Moreover, this locus provides an additional marker to infer levels of such gene flow.

A previous study employing genotype frequency data for *Pgm-3* in Georgia polygyne queens and workers estimated that monogyne males accounted for 84% of matings of polygyne queens (ROSS and KELLER 1995a). A separate study of 69 laboratory families derived from Georgia polygyne queens indicated that 90% were sired by monogyne males, based on comparison of the inferred proportion of fathers possessing genotype *Pgm-3*<sup>A</sup> with the known frequencies of this genotype in wild-caught males of the two forms (ROSS and SHOEMAKER

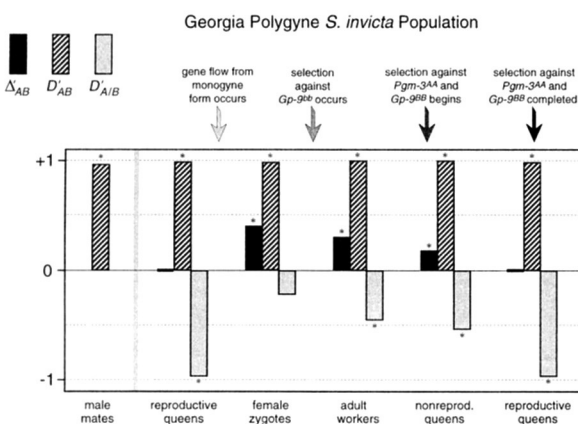


FIGURE 3.—Estimates of linkage disequilibrium in various classes of females from the Georgia polygyne *S. invicta* population and in the mates of the reproductive queens. Values are for the normalized disequilibrium coefficients  $\Delta'_{AB}$  (composite linkage disequilibrium),  $D'_{AB}$  (gametic disequilibrium), and  $D'_{A/B}$  (nongametic disequilibrium), which are scaled by local allele frequencies to vary between  $-1$  and  $+1$ . Values that differ significantly from zero are indicated by \*. The major bouts of gene flow and selection are indicated in approximate relation to the life stages or castes that they affect.



1993). This latter method can be used in the present study to infer the proportion of the 60 progenies of polygyne queens sired by monogyne males. Using *Pgm-3* data, the estimate is 77%. Using data from *Gp-9*, the figure is 80%. These estimates presumably are not independent because of disequilibrium between the two genes. Nonetheless, the data provide further evidence for the substantial gene flow into the polygyne form that must be considered in attempts to understand the two-locus frequency dynamics.

## DISCUSSION

**Overview of effects of multiple evolutionary forces in polygyne *S. invicta*:** This study suggests a general scheme for how complex selection acting on two linked genes, in conjunction with gene flow from a genetically differentiated population, influences gene frequencies and disequilibrium in polygyne *S. invicta* in Georgia. In the diagram presented as Figure 4, this scheme is superimposed on the generational cycle of the ants. Observed two-locus genotype and haplotype frequencies are presented for the different life stages and castes and compared to predicted frequencies when appropriate. Disequilibrium values estimated for each of these classes are also shown.

Reproductive queens effectively possess only two different two-locus genotypes. Because most of these queens are *cis* double heterozygotes and crossing over is rare, the large majority of gametes produced (90%) are in the *cis* linkage phase [either both "wild-type" (*A, B*) or both "mutant" (*a, b*) alleles are present on the same chromosome]. Among these *cis* gametes, over half are haplotype *a b*. The virtual absence of the *trans* haplotype *A b* generates a maximal positive value for gametic disequilibrium ( $D'_{AB}$ ) in reproductive queens (as well as other polygyne females). The large negative value for nongametic disequilibrium ( $D'_{A/B}$ ) in these queens stems from the preponderance of *cis* double heterozygotes, in which the wild-type allele at one locus necessarily is paired with the mutant allele at the second locus on the homologous chromosome (Figure 4).

*Gp-9* cannot be scored in males, so two-locus haplotype frequencies cannot be obtained directly for these individuals in the polygyne social form. Therefore, the frequencies for such males are assumed to be the same as for their mothers, a reasonable assumption given the low rate of crossing over and the fact that males develop from unfertilized eggs (but see PAMILO (1993) for counterexamples from the ant *Formica aquilonia*). Two-locus haplotype frequencies for monogyne males can be obtained directly from estimates of *Pgm-3* allele frequencies (ROSS 1992) because *Gp-9* is monomorphic in this form. Data from the family studies suggest that monogyne males account for ~80% of successful inseminations of polygyne queens. The composite male gamete pool formed by mixing the male gamete pools of the

two forms in an 80:20 ratio (monogyne:polygyne) is lacking haplotype *A b* (thus yielding the high  $D'_{AB}$  value), and haplotype *A B* predominates among the remaining gametes because of the proportionately high representation of monogyne males (Figure 4). Mating and fertilization in this polygyne population thus involves the union of mostly *cis* gametes, but the proportions of the two *cis* haplotypes differ markedly between queens and their mates.

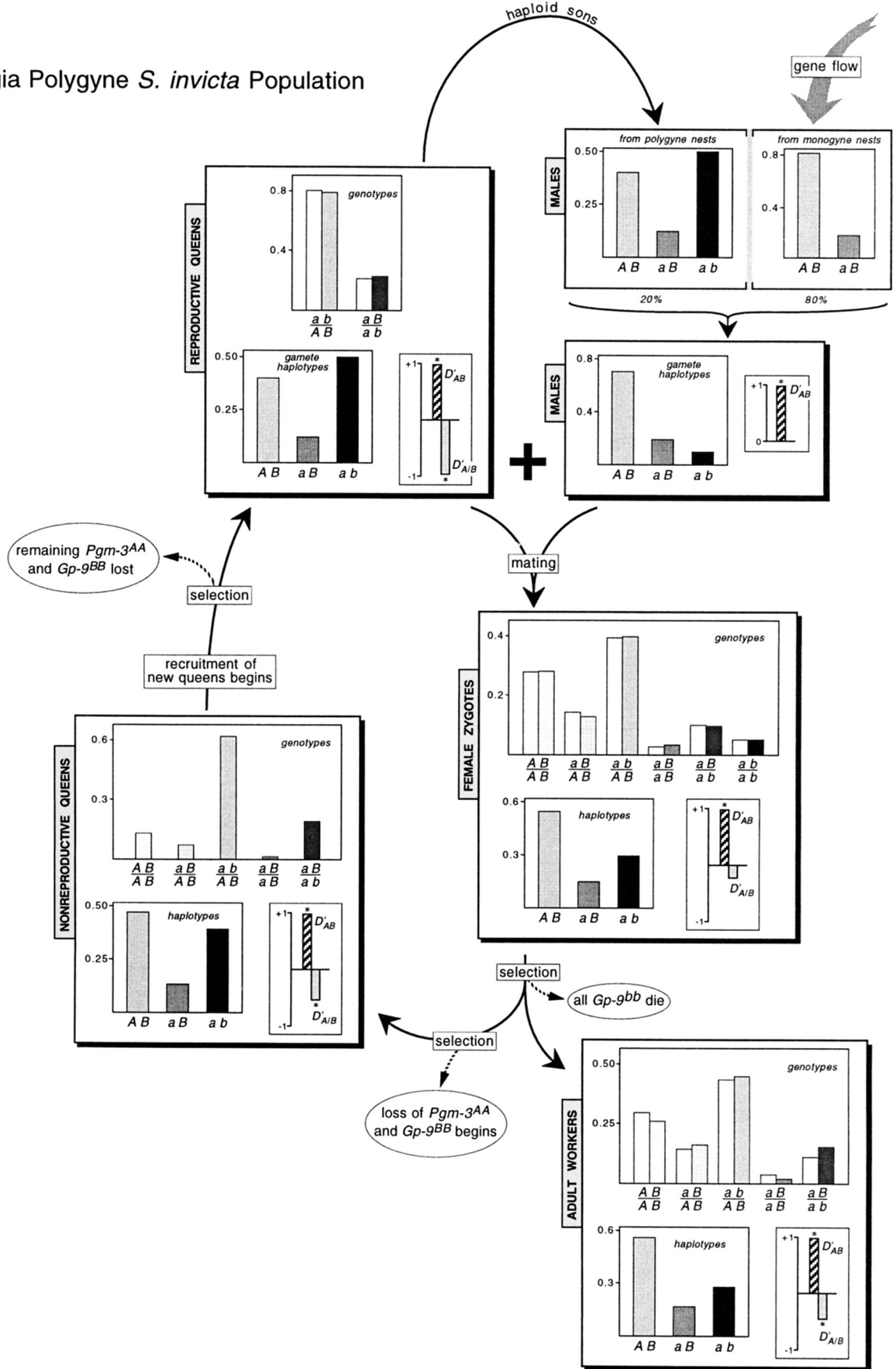
The observed distribution of genotypes in female zygotes, assuming no diminished survival of genotype *Gp-9<sup>bb</sup>*, is similar to that expected with random union of gametes from the female and composite male gamete pools (Figure 4). Genotypic diversity increases substantially in zygotes (six genotypes) compared to their mothers (two genotypes), with about one-third of zygotes possessing the double homozygous genotypes *Pgm-3<sup>AA</sup> Gp-9<sup>BB</sup>* and *Pgm-3<sup>aa</sup> Gp-9<sup>bb</sup>*, which are absent in their mothers. The proportion of double heterozygotes (all *cis*) correspondingly drops from 80% in the mothers to 40% in zygotes. The dramatic reduction of this class, coupled with the appearance of double homozygotes in which the two wild-type or two mutant alleles are paired between the two loci on alternate homologues, explains the sharp positive increase in nongametic disequilibrium ( $D'_{A/B}$ ) in zygotes relative to their mothers.

Viability selection against *Gp-9<sup>bb</sup>* occurs during postzygotic development, resulting in the loss in adults of the sole two-locus class involving this genotype (*Pgm-3<sup>aa</sup> Gp-9<sup>bb</sup>*). This genotype, expected at a frequency of 0.05 in zygotes, has the two mutant alleles paired at the two loci on alternate homologues. Thus, its loss explains the more negative value of  $D'_{A/B}$  in adult workers from the wild than in zygotes. The five genotypes remaining in workers are observed at frequencies similar to those expected after accounting for the loss of genotype *Gp-9<sup>bb</sup>* (Figure 4).

In the mature nonreproductive queens sampled for this study, selection against genotypes *Pgm-3<sup>AA</sup>* and *Gp-9<sup>BB</sup>* so evident in reproductive queens apparently already has begun (above). The major class affected is genotype *Pgm-3<sup>AA</sup> Gp-9<sup>BB</sup>*, in which the two wild-type alleles are paired at the two loci on alternate homologues. Thus, the onset of such selection on queens has the effect of driving down further the negative value of  $D'_{A/B}$  initially generated by selection against *Gp-9<sup>bb</sup>*.

Selection on queens culminates in the complete loss of the three classes involving *Pgm-3<sup>AA</sup>* or *Gp-9<sup>BB</sup>* in reproductive queens, with the loss of the previously common genotype *Pgm-3<sup>AA</sup> Gp-9<sup>BB</sup>* producing the final large negative increase in  $D'_{A/B}$ . The observed frequencies of the two genotypes remaining in reproductive queens are in good agreement with the frequencies expected given complete loss of all genotypes involving *Gp-9<sup>bb</sup>*, *Gp-9<sup>BB</sup>*, or *Pgm-3<sup>AA</sup>* from the original genotype pool inferred for zygotes (Figure 4).

Georgia Polygyne *S. invicta* Population



**Number and identity of genes under selection:** That two linked loci under similar strong selection pressures should be discovered by chance in a genome with  $2N = 32$  chromosomes (GLANCEY *et al.* 1976) prompts consideration of several points. Foremost among these is whether the genotypes scored at either locus could be artifactual or whether the two putative loci could be markers of a single locus that is the sole target of selection. Extensive earlier studies of the inheritance of *Pgm-3* (ROSS 1992; SHOEMAKER *et al.* 1992) and those reported here for *Gp-9* indicate that these behave as co-dominant Mendelian genes, and the contrasting patterns of expression of the banding phenotypes assigned to *Pgm-3* and *Gp-9* in the different sexes and life stages indicate that the products of a single gene were not misinterpreted as the products of two genes. Moreover, the close fit of genotype frequencies in the United States to those predicted by a model incorporating different forms of selection on two genes lends credence to the assumption that two separate but linked Mendelian genes are involved in this system, as does the ability of a two-locus model employing gene flow to account for the disequilibrium in native populations (see below). Indeed, the apparent failure of any single-locus model to explain the persistent disequilibrium observed between *Pgm-3* and *Gp-9* in Argentina implicates multiple genes under selection, although theoretical studies suggest that such genes must interact epistatically in their effects on fitness rather than additively, as assumed here (*e.g.*, HARTL and CLARK 1989, pg. 215). Additional population and molecular data are required to definitively resolve the issue of the number of genes under selection.

Assuming that multiple genes are under selection, these may actually be *Pgm-3* and *Gp-9*, or one or both of these loci may mark other genes that are the true targets of selection (*e.g.*, KOJIMA and SCHAFFER 1967; HEDRICK 1982). Differences in the physiological phenotypes of queens with different *Pgm-3* genotypes are such as might be expected from variation in activity or stability of an enzyme involved in glycogen metabolism, such as phosphoglucosmutase, a fact that has been used to argue that *Pgm-3* is directly responsible for the phenotypic variation detected by selection (ROSS *et al.* 1996a). The coincidence of patterns of expression and of selection at *Gp-9* suggests that this gene also may be a real target of selection. Expression of *Gp-9* is age-dependent, the protein product being detectable only in adults over 1 week of age. Workers sampled in the field were older workers, recognizable by their dark cuticles. Workers sampled in the family studies were known to

be no older than 10 days, and appeared considerably younger than the field workers based on cuticle color. Although genotype *Gp-9<sup>bb</sup>* was consistently underrepresented among workers from the laboratory families, it was nonetheless far more common in these families than in the field, suggesting that the lethal effects of the genotype are age-dependent and parallel the time course of expression of *Gp-9*. Significantly, *Gp-9* is not lethal in hemizygous (haploid) males; 10% of the fathers of the laboratory families were inferred to possess this haplotype. The fact that *Gp-9* appears not to be expressed in males of any age accords with this finding and further supports the idea that this gene may be an actual target of viability selection.

A related issue is whether selection on *Gp-9* alone can explain the changes in *Pgm-3* genotype frequencies previously attributed to direct selection on *Pgm-3*. For instance, apparent discrimination against *Pgm-3<sup>AA</sup>* queens by workers in the United States could be due instead to discrimination against queens bearing *Gp-9<sup>BB</sup>*, because *Pgm-3<sup>AA</sup>* only occurs on the background of this single *Gp-9* genotype (other possible two-locus genotypes involving *Pgm-3<sup>AA</sup>* do not arise because haplotype *A b* is absent). Evidence that this is not the case and that selection acts independently at the two loci, at least in native polygyne populations, comes from the fact that genotype *Gp-9<sup>BB</sup>* occurs commonly in reproductive queens in Argentina but *Pgm-3<sup>AA</sup>* nonetheless is completely absent in these queens. The difference in the apparent strength of selection against *Gp-9<sup>BB</sup>* queens in the native and introduced populations may be due to differences in selective regimes between the two environments, to loss of variation in the United States at regulatory loci determining the magnitude of the phenotypic effects of each locus, or to loss of cryptic variation at either *Pgm-3* or *Gp-9*.

The existence of two tightly linked genes under strong differential selection in the two social forms raises the question of whether they constitute a supergene or coadapted gene complex involved in the origin of polygyny in *S. invicta* (monogyny is presumed to be ancestral; ROSS and KELLER 1995b). On the one hand, appropriate variation at these genes seems to be a genetic prerequisite to polygyny, because the absence of allele *Pgm-3<sup>a</sup>* or of genotype *Gp-9<sup>bb</sup>*, both of which are strongly favored in reproductive queens, presumably would preclude the formation of stable polygyne societies. On the other hand, selection apparently has not produced harmoniously interacting alleles that are the hallmark of epistasis (WADE 1992; WHITLOCK *et al.* 1995), but instead seems to act in opposition at the two

FIGURE 4.—Overview of multilocus selection and gene flow in the Georgia polygyne *S. invicta* population. Observed two-locus genotype and haplotype frequencies (shaded bars) are presented for the different sexes, life stages, and castes (genotypes and haplotypes present at frequencies < 0.01 are not shown); when available, expected genotype frequencies also are shown (□). Estimates of gametic and nongametic disequilibrium are shown for each class of individuals (\* indicates values significantly different from zero; see also Figure 3). Data for female zygotes are derived from family studies and represent projected values before selection occurs.

loci in the polygyne form. For example, overdominance at *Gp-9* in reproductive queens opposes directional selection against *Pgm-3<sup>A</sup>* driven by worker destruction of *Pgm-3<sup>AA</sup>* queens. Thus, the two genes appear not to have evolved as a single integrated genomic unit carrying particular allele combinations favorable to polygyny. Nevertheless, *Pgm-3* and *Gp-9* may mark a region of reduced recombination with a concentration of genes involved in the determination of social organization.

A recessive lethal expressed in reproductive females, such as *Gp-9<sup>l</sup>*, is subject to strong negative selection that should drive it to low frequency and, perhaps, eventual loss via drift. Positive selection on allele *Pgm-3<sup>l</sup>*, mediated by worker execution of all *Pgm-3<sup>AA</sup>* queens, apparently plays some role in counteracting the negative selection on allele *Gp-9<sup>l</sup>* because of strong gametic disequilibrium. However, the advantage of genotype *Gp-9<sup>bb</sup>* over *Gp-9<sup>Bb</sup>* in polygyne queens presumably is the major force opposing selective elimination of *Gp-9<sup>l</sup>* and thereby maintaining polymorphism at the locus. In the monogyne form, any such advantage to heterozygous queens must be outweighed by disadvantages associated with the presence of the recessive lethal allele, leading to its virtual absence in populations of this form. Indeed, the distribution of *Gp-9<sup>l</sup>* in the two forms may be attributable to differences in their breeding behavior and associated differences in the "cost" to queens of producing *Gp-9<sup>bb</sup>* offspring. Monogyne queens found new colonies independently, without the assistance of workers (MARKIN *et al.* 1972). Success in this enterprise requires production of a large number of offspring over a brief period, so that the incipient colony can grow, feed its queen, and defend its territory against other colonies (MARKIN *et al.* 1973). Any heterozygous monogyne queen mated to a *Gp-9<sup>l</sup>* male would lose half of her new worker force at the most critical point of colony development, resulting in certain death of the colony and queen. Polygyne queens rarely practice independent colony founding. Instead, they remain in their parent nest or enter another polygyne nest to initiate reproduction (GLANCEY and LOFGREN 1988; PORTER 1991; ROSS *et al.* 1996b). Because a large worker force already exists for such queens, there may be little or no fitness penalty to individual polygyne queens that lose half of their offspring due to the lethality of genotype *Gp-9<sup>l</sup>*.

**Origin of gametic disequilibrium:** An important goal of studies of multilocus evolution is to identify the forces responsible for generating gametic disequilibrium and for maintaining it against the randomizing force of recombination (HEDRICK *et al.* 1978; ASMUSSEN and CLEGG 1981; HEDRICK 1982; HASTINGS 1985). Determination of the type and amount of selection required to produce disequilibrium is of particular interest, because broad claims for the importance of epistatic selection in evolution are bolstered by evidence that epistasis actually causes disequilibrium in wild popula-

tions (BARKER 1979; HEDRICK 1985). Any of the forces capable of generating disequilibrium (genetic drift, gene flow between differentiated populations, selection) are more effective when coupled with linkage or inbreeding (HEDRICK *et al.* 1978; LANGLEY *et al.* 1978; EPERSON and ALLARD 1987; ZAPATA and ALVAREZ 1992). Thus, the tight linkage of *Pgm-3* and *Gp-9* clearly is important in *maintaining* strong disequilibrium in polygyne *S. invicta* once it is established. [Inbreeding apparently is not important in this regard, as it has not been detected in extensive electrophoretic surveys of native and introduced *S. invicta* (ROSS and FLETCHER 1985; ROSS 1993; ROSS *et al.* 1993, 1996a)].

With respect to the *origin* of gametic disequilibrium, it is tempting to ascribe it to drift associated with the pronounced population bottleneck experienced by fire ants during their recent colonization of the United States (*e.g.*, BERLOCHER 1984; ROSS *et al.* 1993). However, data from the native polygyne populations suggest that the allelic associations responsible for disequilibrium were established before the introduction. Patterns of excesses and deficiencies of two-locus genotypes in both Argentine study populations parallel those observed in Georgia (Table 3), suggesting that *cis* haplotypes are more common and *trans* haplotypes less common than expected by chance in the native as well as introduced ants. This is suggested further by composite disequilibrium values for the Argentine ants that mirror those in the United States. Moreover, the haplotype *A b* appears to be effectively absent in the Formosa population, as it is in Georgia. Independent bottlenecks in the native populations due to historical restrictions of effective population size could explain the origin of disequilibrium, although the specific allelic associations are not necessarily expected to be the same in the two native populations under this scenario.

Gene flow between the social forms, which are strongly differentiated at *Pgm-3* and *Gp-9*, apparently also does not generate significant disequilibrium in the introduced ants. The virtual absence of gamete *A b* in the Georgia polygyne population leads to near-maximal values of  $D_{AB}$ , so there is little opportunity for gene flow from the monogyne form to further increase disequilibrium over the short term. Furthermore, results of a model suggest that extensive variation can be maintained at *Pgm-3* in the Georgia population in the absence of interform gene flow when there is strong overdominance at *Gp-9* in queens (unpublished data). Thus, it is not obvious that cessation of gene flow in the introduced ants would lead to eventual loss of variability at *Pgm-3* and the elimination of disequilibrium by this means. In the native populations, on the other hand, interform gene flow does seem to be necessary to prevent erosion of variability at *Pgm-3* (unpublished data) and thus may contribute significantly in this way to the observed disequilibrium.

Turning finally to the role of selection, consideration

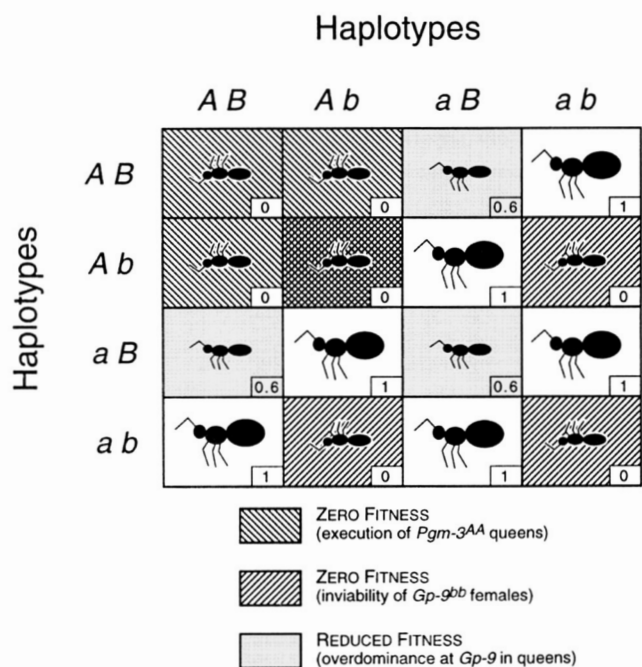


FIGURE 5.—Estimated fitnesses of two-locus genotypes of *Pgm-3* and *Gp-9* in native polygyne populations of *S. invicta*.

of the conditions responsible for the rarity of *Ab* gametes in the native ants suggests that selection acting independently on *Pgm-3* and *Gp-9* (or other genes that they mark) may play a role in the establishment of disequilibrium. The relative fitnesses of the four haplotypes can be judged from the fitnesses of the two-locus genotypes that they compose (Figure 5). Genotypes including *Pgm-3<sup>AA</sup>* or *Gp-9<sup>bb</sup>* have zero fitness because they are effectively lethal in reproductive females. Among the remainder, those with *Gp-9<sup>bb</sup>* can be assigned a fitness of 1.0 and those with *Gp-9<sup>BB</sup>* a fitness of 0.6, reflecting the ~1.6-fold survival advantage of reproductive queens that possess the former genotype in Argentina (above). Haplotype *Ab* can give rise to only one viable diploid genotype, the fewest among the four haplotypes. Haplotypes *aB* and *ab* appear to have unconditionally higher fitness than *Ab*, whereas *AB* may have lower fitness than *Ab* when haplotype *ab* is relatively uncommon. However, *AB* predominates in the gametes introduced via gene flow from the monogyne form, suggesting that it may be buffered against large decreases in frequency when its overall fitness is low. Thus, selection acting independently on *Pgm-3* and *Gp-9*, combined with gene flow from the monogyne form, may contribute directly to the rarity or absence of haplotype *Ab* in native polygyne fire ant populations under a variety of initial frequencies of the haplotypes. Epistatic selection is not involved in the consequent appearance of strong disequilibrium in this scenario.

I thank MARJORIE ASMUSSEN, JOHN AVISE, MIKE GOODISMAN, LAURENT KELLER, and ED VARGO for comments on a previous version of the paper, and LAURENT KELLER, JAMES TRAGER, WALTER TSCHINKEL, and ED VARGO for help in obtaining specimens. Partial funding

for this study was provided by the National Geographic Society and by the Georgia Agricultural Experiment Stations, University of Georgia.

## LITERATURE CITED

- ANTONOVICS, J., A. D. BRADSHAW and R. G. TURNER, 1971 Heavy metal tolerance in plants. *Adv. Ecol. Res.* **7**: 1–85.
- ASMUSSEN, M. A., and M. T. CLEGG, 1981 Dynamics of the linkage disequilibrium function under models of gene-frequency hitchhiking. *Genetics* **99**: 337–356.
- BARKER, J. S. F., 1979 Inter-locus interactions: a review of experimental evidence. *Theor. Pop. Biol.* **16**: 323–346.
- BARTON, N., and A. CLARK, 1990 Population structure and process in evolution, pp. 115–173 in *Population Biology: Ecological and Evolutionary Viewpoints*, edited by K. WÖHRMANN and S. K. JAIN. Springer, Berlin.
- BERLOCHER, S. H., 1984 Genetic changes coinciding with the colonization of California by the walnut husk fly *Rhagoletis completa*. *Evolution* **38**: 906–918.
- BERLOCHER, S. H., 1993 Gametic disequilibrium between allozyme loci and sex chromosomes in the genus *Rhagoletis*. *J. Hered.* **84**: 431–437.
- BIJLSMA, R., 1978 Polymorphism at the G6PD and 6PGD loci in *Drosophila melanogaster*. II. Evidence for interaction in fitness. *Genet. Res.* **31**: 227–237.
- CHEVILLON, C., N. PASTEUR, M. MARQUINE, D. HEYSE and M. RAYMOND, 1995 Population structure and dynamics of selected genes in the mosquito *Culex pipiens*. *Evolution* **49**: 997–1007.
- CLEGG, M. T., J. F. KIDWELL, M. G. KIDWELL and N. J. DANIEL, 1976 Dynamics of correlated genetic systems. I. Selection in the region of the Glued locus of *Drosophila melanogaster*. *Genetics* **83**: 793–810.
- CROZIER, R. H., B. H. SMITH and Y. C. CROZIER, 1987 Relatedness and population structure of the primitively eusocial bee *Lasioglossum zephyrum* (Hymenoptera: Halictidae) in Kansas. *Evolution* **41**: 902–910.
- ENOS, R. A., 1981 Manifold effects of the cyanogenic loci in white clover. *Heredity* **46**: 127–132.
- EPPELSON, B. K., and R. W. ALLARD, 1987 Linkage disequilibrium between allozymes in natural populations of lodgepole pine. *Genetics* **115**: 341–352.
- GLANCEY, B. M., and C. S. LOFGREN, 1988 Adoption of newly-mated queens: a mechanism for proliferation and perpetuation of polygynous red imported fire ants, *Solenopsis invicta* Buren. *Florida Entomol.* **71**: 581–587.
- GLANCEY, B. M., M. K. ST. ROMAIN and R. H. CROZIER, 1976 Chromosome numbers of the red and the black imported fire ants, *Solenopsis invicta* and *S. richteri*. *Ann. Entomol. Soc. Am.* **69**: 469–470.
- GREENBERG, L., D. J. C. FLETCHER and S. B. VINSON, 1985 Differences in worker size and mound distribution in monogynous and polygynous colonies of the fire ant *Solenopsis invicta* Buren. *J. Kansas Entomol. Soc.* **58**: 9–18.
- HARTL, D. L., and A. G. CLARK, 1989 *Principles of Population Genetics*. Sinauer, Sunderland, MA.
- HASTINGS, A., 1985 Multilocus population genetics with weak epistasis. I. Equilibrium properties of two-locus two-allele models. *Genetics* **109**: 799–812.
- HEDRICK, P. W., 1982 Genetic hitchhiking: a new factor in evolution? *Bioscience* **32**: 845–853.
- HEDRICK, P. W., 1985 *Genetics of Populations*. Jones and Bartlett, Boston.
- HEDRICK, P. W., 1987 Gametic disequilibrium measures: proceed with caution. *Genetics* **117**: 331–341.
- HEDRICK, P. W., 1994 Evolutionary genetics of the major histocompatibility complex. *Am. Nat.* **143**: 945–964.
- HEDRICK, P. W., S. JAIN and L. HOLDEN, 1978 Multilocus systems in evolution. *Evol. Biol.* **11**: 101–182.
- HILL, W. G., 1974 Estimation of linkage disequilibrium in randomly mating populations. *Heredity* **33**: 229–239.
- HOCHBERG, Y., 1988 A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* **75**: 800–802.
- HOLLOCHER, H., and A. R. TEMPLETON, 1994 The molecular through ecological genetics of *abnormal abdomen* in *Drosophila mercatorum*. VI. The non-neutrality of the Y chromosome rDNA polymorphism. *Genetics* **136**: 1373–1384.

- JONES, J. S., B. H. LEITH and P. RAWLINGS, 1977 Polymorphism in *Cepaea*: a problem with too many solutions? *Annu. Rev. Ecol. Syst.* **8**: 109–143.
- KELLER, L., and K. G. ROSS, 1993 Phenotypic basis of reproductive success in a social insect: genetic and social determinants. *Science* **260**: 1107–1110.
- KING, R. B., and R. LAWSON, 1995 Color-pattern variation in Lake Erie water snakes: the role of gene flow. *Evolution* **49**: 885–896.
- KOEHN, R. K., A. J. ZERA and J. G. HALL, 1983 Enzyme polymorphism and natural selection, pp. 115–136 in *Evolution of Genes and Proteins*, edited by M. NEI and R. K. KOEHN. Sinauer, Sunderland, MA.
- KOJIMA, K., and H. E. SCHAFER, 1967 Survival processes of linked mutant genes. *Evolution* **21**: 518–531.
- LANGLEY, C. H., D. B. SMITH and F. M. JOHNSON, 1978 Analysis of linkage disequilibria between allozyme loci in natural populations of *Drosophila melanogaster*. *Genet. Res.* **32**: 215–229.
- LEWONTIN, R. C., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.
- LYTTLE, T. W., 1991 Segregation distorters. *Annu. Rev. Genet.* **25**: 511–557.
- MALLET, J., N. BARTON, G. LAMAS, J. SANTISTEBAN, M. MUEDAS *et al.*, 1990 Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. *Genetics* **124**: 921–936.
- MARKIN, G. P., H. L. COLLINS and J. H. DILLIER, 1972 Colony founding by queens of the red imported fire ant, *Solenopsis invicta*. *Ann. Entomol. Soc. Am.* **65**: 1053–1058.
- MARKIN, G. P., J. H. DILLIER and H. L. COLLINS, 1973 Growth and development of colonies of the red imported fire ant, *Solenopsis invicta*. *Ann. Entomol. Soc. Am.* **66**: 803–808.
- MCNEILLY, T., 1968 Evolution in closely adjacent plant populations. III. *Agrostis tenuis* on a small copper mine. *Heredity* **23**: 99–108.
- MORITA, T., H. KUBOTA, K. MURATA, M. NOZAKI, C. DELARBRE *et al.*, 1992 Evolution of the mouse *t* haplotype: recent and worldwide introgression to *Mus musculus*. *Proc. Natl. Acad. Sci. USA* **89**: 6851–6855.
- NEJATI-JAVAREMI, A., and C. SMITH, 1996 Assigning linkage haplotypes from parent and progeny genotypes. *Genetics* **142**: 1363–1367.
- O'NEAL, J., and G. P. MARKIN, 1975 Brood development of the various castes of the imported fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae). *J. Kansas Entomol. Soc.* **48**: 152–159.
- PAMILO, P., 1983 Genetic differentiation within subdivided populations of *Formica* ants. *Evolution* **37**: 1010–1022.
- PAMILO, P., 1993 Polyandry and allele frequency differences between the sexes in the ant *Formica aquilonia*. *Heredity* **70**: 472–480.
- PARHAM, P., and T. OHTA, 1996 Population biology of antigen presentation by MHC Class I molecules. *Science* **272**: 67–74.
- PORTER, S. D., 1991 Origins of new queens in polygynous red imported fire ant colonies (Hymenoptera: Formicidae). *J. Entomol. Sci.* **26**: 474–478.
- PORTER, S. D., 1992 Frequency and distribution of polygynous fire ants (Hymenoptera: Formicidae) in Florida. *Florida Entomol.* **75**: 248–257.
- RIECHERT, S. E., 1993 Investigation of potential gene flow limitation of behavioral adaptation in an aridlands spider. *Behav. Ecol. Sociobiol.* **32**: 355–363.
- ROSS, K. G., 1992 Strong selection on a gene that influences reproductive competition in a social insect. *Nature* **355**: 347–349.
- ROSS, K. G., 1993 The breeding system of the fire ant *Solenopsis invicta*: effects on colony genetic structure. *Am. Nat.* **141**: 554–576.
- ROSS, K. G., and D. J. C. FLETCHER, 1985 Comparative study of genetic and social structure in two forms of the fire ant, *Solenopsis invicta* (Hymenoptera: Formicidae). *Behav. Ecol. Sociobiol.* **17**: 349–356.
- ROSS, K. G., and L. KELLER, 1995a Joint influence of gene flow and selection on a reproductively important genetic polymorphism in the fire ant *Solenopsis invicta*. *Am. Nat.* **146**: 325–348.
- ROSS, K. G., and L. KELLER, 1995b Ecology and evolution of social organization: insights from fire ants and other highly eusocial insects. *Annu. Rev. Ecol. Syst.* **26**: 631–656.
- ROSS, K. G., and D. D. SHOEMAKER, 1993 An unusual pattern of gene flow between the two social forms of the fire ant *Solenopsis invicta*. *Evolution* **47**: 1595–1605.
- ROSS, K. G., E. L. VARGO, L. KELLER and J. C. TRAGER, 1993 Effect of a founder event on variation in the genetic sex-determining system of the fire ant *Solenopsis invicta*. *Genetics* **135**: 843–854.
- ROSS, K. G., E. L. VARGO and L. KELLER, 1996a Simple genetic basis for important social traits in the fire ant *Solenopsis invicta*. *Evolution* **50**: 2387–2399.
- ROSS, K. G., E. L. VARGO and L. KELLER, 1996b Social evolution in a new environment: the case of introduced fire ants. *Proc. Natl. Acad. Sci. USA* **93**: 3021–3025.
- SHOEMAKER, D. D., J. T. COSTA and K. G. ROSS, 1992 Estimates of heterozygosity in two social insects using a large number of electrophoretic markers. *Heredity* **69**: 573–582.
- SILVER, L. M., 1985 Mouse *t* haplotypes. *Annu. Rev. Genet.* **19**: 179–208.
- SUNDSTRÖM, L., 1993 Genetic population structure and sociogenetic organization in *Formica truncorum* (Hymenoptera: Formicidae). *Behav. Ecol. Sociobiol.* **33**: 345–354.
- TSCHINKEL, W. R., 1993 Resource allocation, brood production and cannibalism during colony founding in the fire ant, *Solenopsis invicta*. *Behav. Ecol. Sociobiol.* **33**: 209–223.
- TURNER, J. R. G., 1977 Butterfly mimicry: the genetical evolution of an adaptation. *Evol. Biol.* **10**: 163–206.
- WADE, M. J., 1992 Epistasis, pp. 87–91 in *Keywords in Evolutionary Biology*, edited by E. F. KELLER and E. A. LLOYD. Harvard University Press, Cambridge, MA.
- WEIR, B. S., 1996 *Genetic Data Analysis II*. Sinauer, Sunderland, MA.
- WEIR, B. S., and C. C. COCKERHAM, 1979 Estimation of linkage disequilibrium in randomly mating populations. *Heredity* **42**: 105–111.
- WHITLOCK, M. C., P. C. PHILLIPS, F. B.-G. MOORE and S. J. TONSOR, 1995 Multiple fitness peaks and epistasis. *Annu. Rev. Ecol. Syst.* **26**: 601–629.
- ZAPATA, C., and G. ALVAREZ, 1992 The detection of gametic disequilibrium between allozyme loci in natural populations of *Drosophila*. *Evolution* **46**: 1900–1917.

Communicating editor: G. B. GOLDING