Evolution, 44(5), 1990, pp. 1242–1253

POPULATION STRUCTURE AND KINSHIP IN *POLISTES* (HYMENOPTERA, VESPIDAE): AN ANALYSIS USING RIBOSOMAL DNA AND PROTEIN ELECTROPHORESIS

SCOTT K. DAVIS,[1,3] JOAN E. STRASSMANN,[2] COLIN HUGHES,[2]
L. SUSAN PLETSCHER,[1,3] AND ALAN R. TEMPLETON[1]
[1]*Department of Biology, Washington University, St. Louis, MO 63130 USA*
[2]*Department of Biology, Rice University, Houston, TX 77005 USA*

Abstract.—Six variable protein loci and one variable ribosomal DNA restriction site were used for an analysis of population structure in five species of *Polistes* from Texas. A sample-reuse algorithm was developed that estimated F_{ST}, F_{IS}, and $\varnothing$ (the coefficient of kinship) from probabilities of identity. Of the four species analyzed in detail only one, *Polistes exclamans*, had statistically significant values of F_{ST}. These values may reflect natural constraints on successful nesting for migrants of this species. Three of the four species had significant values of F_{IS} and three of the four species had significant values of $\varnothing$. In many cases $\varnothing$ also differed from the expected value under haplodiploidy and random mating. Values of $\varnothing$ did not differ from expectations under haplodiploidy and local inbreeding. These results emphasize that theories of social behavior and evolution based on coefficients of kinship should include some explicit consideration of population structure.

Received January 12, 1989. Accepted December 21, 1989.

The order Hymenoptera, with over 250,000 species of ants, bees, and wasps, is among the largest and most diverse groups in existence today. It is characterized by several attributes, but two of these are of special interest to evolutionary biologists. First, most hymenopterans reproduce by arrhenotokous parthenogenesis, also known as haplodiploidy, in which fertilized eggs develop into diploid females, and unfertilized eggs develop into haploid males. This mode of reproduction has many consequences for the population genetics of hymenopterans, because all loci are sex linked (Crozier, 1977, 1985).

Second, the order Hymenoptera is unique in possessing a large number of species that are social. While eusociality has risen as many as 11 times in the Hymenoptera (Andersson, 1984), it has arisen in only one other insect lineage, the termites. Haplodiploidy and eusociality have been the focus of a great deal of study and particular attention has centered on any possible relationship between the two.

The haplodiploid genetic system of hymenopterans, with its unusual consequence that sisters are more closely related to each other than they are to their mother, led Hamilton (1964*a*, 1964*b*) to use the theory of kin selection to explain the repeated evolution of eusociality in the group. This initial work has sparked a rapidly growing body of literature exploring various mechanisms by which eusociality could have evolved (reviewed by Andersson, 1984). Implicit in all of these theoretical discussions are assumptions about the population genetics of the organisms in question. The postulated population structures range from panmixia to population substructuring with local inbreeding, but interactions between the system of mating, migration, and genetic drift determine the predictive outcome of these models.

Haplodiploidy and eusociality are also frequent topics in discussions of variation in natural populations. A large body of electrophoretic data on many different taxa has led to the widely debated idea that hymenopterans maintain significantly lower levels of electrophoretically detectable variation than other insects. Haplodiploidy, eusociality, inbreeding, and small population sizes have all been advanced as possible contributing factors (Snyder, 1974; Metcalf et al., 1975; Lester and Selander, 1979; Pamilo and Crozier, 1981; Wagner and Briscoe, 1983). The issue was analyzed in detail by Berkelhamer (1983), who exam-

[3] Present address: Department of Animal Science, Texas A&M University, College Station, TX 77843 USA.

ined a large body of literature and concluded that haplodiploidy had a significant effect in reducing variability, and that eusociality may arise as a result of inbreeding. Berkelhamer's conclusions are based on the idea that different levels of eusocial development will result in different population genetic structures. Although many of these conclusions have recently been questioned (Graur, 1985; Owen, 1985; Reeve et al., 1985), several of the critics also make arguments about population structure and the relationship between eusocial status and effective population size. Once again, the conclusions of all of these studies depend critically on the population biology of the organisms in question.

In spite of the importance of population genetic parameters to the evolution of eusociality, there are few studies that address aspects of population structure in eusocial Hymenoptera (e.g., Johnson et al., 1969; Tomaszewski et al., 1973; Metcalf and Whitt, 1977; Metcalf, 1980; Ward, 1980; Lester and Selander, 1981; Pamilo, 1983; Crozier et al., 1984, 1987). Few of these studies focus on primitively eusocial species, in spite of the fact that these species offer valuable insights into the evolution of eusociality (Andersson, 1984).

The present study examines the structure of natural populations of the primitively eusocial wasp genus *Polistes*. In addition to the standard technique of starch gel electrophoresis, we also use restriction mapping of ribosomal DNA repeats as a tool for analyzing population structure. Repeat length heterogeneity has proven to be abundant both within individuals and within populations (Arnheim et al., 1982; Ranzani et al., 1984; Saghai-Maroof, 1984; Williams et al., 1985) and in at least one case has shown potential as a method for analyzing population structure (Learn and Schaal, 1987). Ribosomal DNA restriction site differences have also been used as character states in phylogenetic reconstructions (Wilson et al., 1984; Systma and Schaal, 1985; Hillis and Davis, 1986; Sites and Davis, 1989), and it is of considerable interest to see how these genetic markers behave at the population level. Combining restriction site data with protein data from the same populations of *Polistes* will allow a comparative evaluation of their usefulness in this context and of the relationship between population structure and eusociality.

Materials and Methods

The site chosen for this study was Brazos Bend State Park, a 1983 hectare area located 34 km south of Houston, Texas. The park spans a diverse array of habitats ranging from native shortgrass prairie to oak forest and supports several species of paper wasps, including *Polistes exclamans, P. dorsalis, P. metricus, P. carolinus,* and *P. bellicosus*. Because of the difficulties inherent in locating a suitable series of nests in nature, we chose to augment our sample size by placing out numerous small nest boxes. These boxes were placed in transects and grids throughout the study area in such a way that nests could be compared across a range of physical distances up to 10 kilometers. Five sample sites were established with nest boxes, although natural nests were used whenever possible (Fig. 1).

Wasps were sampled in August 1984 from all known nests. At this time of the year all of the nests were producing reproductive individuals, and the array of genotypes present on a given nest should be representative of the population that overwintered in 1984–1985. This is not true of samples taken earlier in the season, because queen mortality is very high and the dominant egg-laying female in the fall may not be the same individual that founded the nest (Strassmann, 1981, 1985). Similarly, a much larger sample of nests could be taken earlier in the season, but nest mortality is also very high, and the majority of these nests will not contribute to the overwintering population. For example, only 58 of the 178 nests studied at Brazos Bend in the 1982 and 1983 seasons survived to produce reproductives in the fall (Strassmann and Hughes, unpubl. data). Since fertilized females that successfully overwinter are the founders of the next season's nests, this sampling scheme allows us to examine the gene pool for 1985. Two additional samples of *P. exclamans* were collected from a small grassland area on the Rice University campus in Houston (H1 and H2). Further collections of *P. bellicosus* and *P. exclamans* were made from all available nests in March 1985 to expand on the

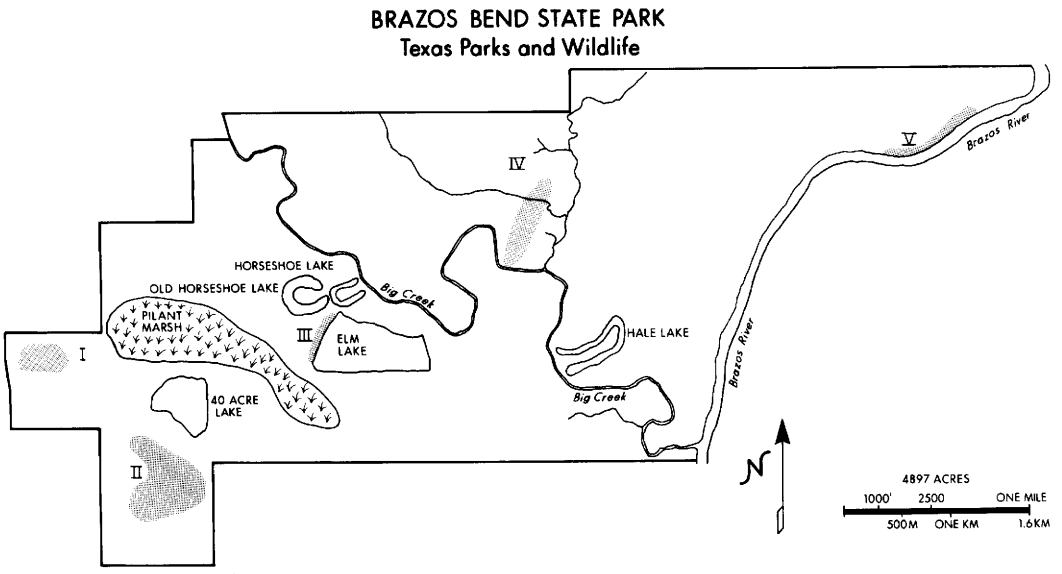

FIG. 1. Map of the Brazos Bend State Park study area. The five collecting localities are indicated by shading.

results seen in 1984. At this early date only foundresses, larvae, and pupae were present on the nests.

Four females were chosen from each nest for genetic analysis. Because of the high incidence of mortality on nests, egg laying by more than one female, multiple mating, and the high rate of queen replacement, nestmates from the fall sample may not be full sibs (Strassmann, 1985), although they are likely to be at least cousins. In the spring 1985 sample, the relatedness of individuals on a nest may be higher.

Previous studies of *Polistes* in Texas have identified 10 highly variable protein coding loci: esterase 1 and 2, phosphoglucomutase 1, 2, and 3, isocitrate dehydrogenase 1 and 2, 6-phosphogluconate dehydrogenase, glucose phosphate isomerase, aconitase, and hexokinase 1 and 2 (Lester and Selander, 1981). All individuals were scored for these systems using standard horizontal starch gel electrophoresis as described by Selander et al. (1971). All protein gels were run using an extract of the gaster, preserving the thorax for DNA analysis. Scoring for esterases was verified using the cellulose acetate gel system of Helena Laboratories (Beaumont, Texas) and methods modified from Clayton and Tretiak (1972). The use of haploid males as internal standards greatly simplified designations of alleles and loci, especially in the complex esterase systems.

Total cellular DNA was extracted from the thorax of each individual using a modification of the proteinase-K protocol described by Hillis and Davis (1986). Depending on the size of the wasp, this procedure yielded from 7 to 15 μg of high-molecular-weight DNA per individual. Seventeen restriction endonucleases, which cut the ribosomal DNA repeating unit of *Polistes* at least once, were used to survey for both length and site heterogeneity. They include *Apa*I, *Bam*HI, *Bcl*I, *Bgl*II, *Bst*EII, *Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Kpn*I, *Nco*I, *Pst*I, *Pvu*II, *Sac*I, *Stu*I, *Xba*I, and *Xmn*I. Electrophoresis of restriction fragments, blotting of gels, hybridization, and autoradiography were carried out exactly as described by Hillis and Davis (1986). The ribosomal DNA probes used were pI19 and p2546, which were gifts from Dr. Norman Arnheim and contain the *Mus musculus* 28 S and 18 S ribosomal genes, respectively.

Once all individuals were scored, the samples were grouped for further analysis. Only those samples with at least five conspecific nests (=40 haploid genomes) were examined further.

Statistical Methods—Estimation

We organized the data in a hierarchical fashion: wasps within nests, nests within subpopulations, and subpopulations within the total population. Corresponding to this hierarchical sampling scheme, we performed a hierarchical F statistic analysis. In particular, we estimated three population genetic parameters: F_{ST} which measures the extent of genetic differentiation between subpopulations; F_{IS}, which measures the extent of local inbreeding; and $\varnothing$, which is the coefficient of kinship between individuals from the same nest relative to the subpopulation. F statistics have been defined both in terms of variances and correlations and in terms of identity-by-descent (Cockerham and Weir, 1987). The identity-by-descent definitions are more flexible (Cockerham and Weir, 1987), particularly when dealing with complex sampling designs. Accordingly, we estimated our F statistics using probabilities of identity determined by a sample-reuse statistical procedure. Similar algorithms have been developed by Pamilo and Crozier (1982) and Pamilo (1984, 1985).

Our sample-reuse algorithm consists of performing many iterations of the following procedure. For every individual in the sample, three additional individuals are drawn at random. One is drawn at random from the remaining individuals in the same nest, the second is drawn from a different nest within the same subpopulation, and the third is drawn from a different subpopulation. An allele is drawn at random for each locus from each of these individuals. Since our sample consisted only of diploid individuals, there are only two alleles to choose from. For each locus, these randomly drawn alleles are then contrasted with one another and a score of "1" is assigned if they are identical and "0" otherwise. The identity of the two alleles borne by the reference individual is also checked for each locus. This is repeated for every individual in the sample, and these scores are summed over individuals for each locus and then divided by the total sample size to estimate the following identity probabilities at each locus: (1) $ID(1, L)$, the probability that the two alleles at locus L from the same individual are identical; (2) $ID(2, L)$, the probability that two randomly drawn alleles from individuals from different subpopulations are identical; (3) $ID(3, L)$, the probability that randomly drawn alleles from individuals from different nests within the same subpopulation are identical; and (4) $ID(4, L)$, the probability that two randomly drawn alleles from two individuals from the same nest are identical. These locus specific identities are then averaged across loci to estimate the overall probability of identity for each of these four contrasts:

$$TID(I) = \frac{\sum_{L=1}^{NL} ID(I, L)}{NL} \quad (1)$$

where NL is the total number of loci used in the analysis.

The above IDs and TIDs represent estimates of probability of identity. However, the F statistics are defined in terms of identity-by-descent (ibd). Unfortunately, it is not generally possible to distinguish between identity-by-descent and identity-by-state. To correct for identity-by-state, the F statistics were estimated as deviations from an appropriate reference population. The probability of identity (not just ibd) of the two alleles drawn from the same individual should be equal to F_{IS} (the probability of ibd relative to the subpopulation) plus the probability that the alleles are identical within the subpopulation given that they are not ibd due to inbreeding within the subpopulation; that is,

$$ID(1, L) = F_{IS} + (1 - F_{IS})ID(3, L) \quad (2)$$

Solving Equation (2) for F_{IS} yields

$$F_{IS} = \frac{[ID(1, L) - ID(3, L)]}{[1 - ID(3, L)]} \quad (3)$$

Similarly, the coefficient of kinship (the probability of ibd between alleles drawn from different individuals from the same nest relative to the subpopulation) is related to the identity of alleles from different nestmates relative to the identities found in the subpopulation as

$$\varnothing = \frac{[ID(4, L) - ID(3, L)]}{[1 - ID(3, L)]} \quad (4)$$

Finally, F_{ST} (the probability of ibd between

alleles drawn from different individuals from the same subpopulation but different nests relative to the total population) is estimated by

$$F_{ST} = \frac{[ID(3, L) - ID(2, L)]}{[1 - ID(2, L)]} \quad (5)$$

Equations (3) through (5) yield locus specific estimators of the F statistics. To obtain an overall estimate of an F statistic, one simply substitutes the corresponding TIDs [Equation (1)] for the IDs found in Equations (3) through (5). An alternative would be to take the average of the locus specific estimators. However, Reynolds et al. (1983) have shown that averaging the variance components across loci and then estimating the F statistics from these average components is much superior to averaging the locus specific F statistic estimators. The distinction between these two estimation procedures becomes very important when some loci have one very common allele such that almost all individuals in the sample are homozygous for this allele. The estimator associated with a very rare allele is unreliable and can yield mathematically meaningless results (Reynolds et al., 1983; Weir and Cockerham, 1984). By averaging the F statistic estimators across loci, the almost meaningless results associated with loci with one very common allele can make a major contribution to the overall estimator. However, by averaging the identities across loci through Equation (1), these rare alleles make only a minor, but informative, contribution to the total identities. Hence, in analogy to Reynolds et al. (1983) and Weir and Cockerham (1984), we will average loci at the identity component level and not at the locus-specific estimator level.

The numerators of Equations (3)–(5) all involve differences between identity probabilities. Thus, it is possible for the estimators to take on negative values. Negative values of $\varnothing$ and F_{ST} are biologically and mathematically meaningless. However, if F_{ST} were truly zero, then the expected value of $TID(2)$ would be the same as $TID(3)$. Thus sampling error would be expected to produce negative values with high probability whenever the true values of F_{ST} (and $\varnothing$) are zero or close to zero.

The definition of F_{IS} that motivated our estimation procedure is based on probabilities of ibd. However, in practice we estimate F_{IS} relative to a reference population. In this case, negative values are not only possible but are biologically meaningful. For example, if mates were avoiding relatives within subpopulations and/or primarily mating with individuals outside their own subpopulation (and with $F_{ST} > 0$), then one would expect $TID(4) < TID(3)$ and $F_{IS} < 0$. Hence F_{IS} could take on negative values that are biologically meaningful.

To average out the stochastic effects of randomly drawn sets of individuals, our final locus-specific and overall estimators are obtained as an average of 1,000 iterations of the above estimation procedure. By repeating the estimation procedure many times, we can also address the issue of hypothesis testing.

Hypothesis Testing and Confidence Intervals

The values of the F statistics estimators are kept for each of the 1,000 iterations of the above estimation procedure. This set of 1,000 estimators generates an empirical probability distribution that reflects the error associated with our resampling algorithm. One thousand iterations is sufficient to ensure that a 95% confidence region can be accurately estimated (Edgington, 1987). These empirical probability distributions are then used to test various hypotheses. To test the hypothesis that a particular F statistic is zero, we see if the empirical 95% confidence interval contains zero. If the 95% confidence interval does not include zero, we reject at the 5% level the null hypothesis that the F statistic is zero.

We examined some additional hypotheses for the coefficient of kinship. Under haplodiploidy and assuming that nestmates are full sisters, the expected coefficient of kinship is 3/8 under random mating. If the 95% confidence interval for the coefficient of kinship does not include 3/8, then we conclude that the estimated coefficient of kinship is significantly different from 3/8 at the 5% level. If the subpopulations are not randomly mating but instead are inbreeding with $F_{IS} > 0$, then the coefficient of kinship be-

TABLE 1. Sample sizes for the five species of *Polistes* surveyed. The upper value refers to the number of nests; the lower value to the total number of individuals.

	Houston		Brazos Bend State Park					
	H1	H2	I	II	III	IV	V	Total
P. exclamans[1]	9/35	8/31	4/15	2/8	5/20	4/16	3/12	35/137
P. exclamans[2]	0/0	0/0	29/111	13/49	0/0	2/8	5/20	49/188
P. metricus	0/0	0/0	1/4	14/50	0/0	5/19	2/8	22/82
P. carolinus	0/0	0/0	0/0	7/28	0/0	6/24	0/0	13/52
P. bellicosus[1]	0/0	0/0	9/36	12/48	0/0	0/0	1/4	22/88
P. bellicosus[2]	0/0	0/0	23/62	15/54	0/0	0/0	0/0	38/116
P. dorsalis	0/0	0/0	2/8	4/16	0/0	1/4	0/0	7/28

[1] Fall of 1984.
[2] Spring of 1985.

tween full sisters under haplodiploidy is no longer 3/8 but instead is

$$\varnothing = 3/8 + (5/8)F_{IS} \quad (6)$$

Equation (6) is not applicable when $F_{IS} < 0$. In that case, the expected coefficient of kinship remains 3/8.

To test the hypothesis that the coefficient of kinship is given by Equation (6), we first note that F_{IS} is estimated with error. By substituting the upper and lower values of the 90% confidence interval for F_{IS} in Equation (6), we can generate the corresponding 90% confidence interval for the expected value of the coefficient of kinship for full sisters under haplodiploidy with local inbreeding. If the 90% confidence interval for the estimated coefficient of kinship did not overlap with the 90% confidence interval defined by Equation (6), we rejected the null hypothesis that Equation (6) gives the true value of the coefficient of kinship at the 5% level. The 90% confidence intervals were used to generate this 5% test because no overlap in the 90% confidence interval means that there is less than 5% overlap in the two distributions being compared since the overlap generally occurs in only one tail.

Additional hypotheses about the coefficient of kinship could be tested with the 1985 data. For *Polistes bellicosus,* there were many nests with multiple foundresses. Hence we could test the hypothesis that the foundresses were full sisters by using the same test described in the preceding paragraph. The 1985 data for both *P. bellicosus* and *P. exclamans* contain many nests with at least one foundress and one or more workers. Thus, we could also estimate the coefficient of kinship between a foundress and workers. The expected coefficient of kinship between a single foundress and her workers under the hypothesis that the workers are the daughters of the foundress and that local inbreeding may be occurring is

$$\varnothing = 1/4 + (3/4)F_{IS} \quad (7)$$

In the case of multiple foundresses and workers on the same nest, the expected coefficient of kinship between foundresses and workers could be closer to that between aunts and nieces. The expected coefficient of kinship between foundresses and workers under the hypothesis that the workers are the nieces of the foundresses and that local inbreeding may be occurring is

$$\varnothing = 3/16 + (13/16)F_{IS} \quad (8)$$

Once again, the hypotheses defined by Equations (7) and (8) are tested at the 5% level by looking for overlap in the appropriate 90% confidence intervals.

RESULTS

Sample sizes and collecting locations are listed in Table 1. Four species were found in sufficient numbers in two or more samples to permit analysis. They are *Polistes metricus, P. carolinus, P. bellicosus,* and *P. exclamans.* Allele frequency data for all variable loci are summarized in Table 2. Esterase 1 and 2 could not be scored with confidence from *P. exclamans* and were not used for that species. The esterase 1 locus was not scorable from pupae and so was not used for the spring, 1985 *P. bellicosus* samples. Heterozygosity was remarkably low, especially since the 12 loci surveyed were chosen because they were known to be polymorphic.

TABLE 2. Allele frequency data for the variable loci.

		P. exclamans							*P. metricus*		*P. carolinus*		*P. bellicosus*	
		H1	H2	I	II	III	IV	V	II	IV	II	IV	I	II
Est-1[1]	Null												0.444	0.417
	90										0.630	0.300		
	100								1.000	1.000	0.370	0.700	0.556	0.583
Est-2	80										0.074			
	90										0.130	0.425		
	100								1.000	1.000	0.796	0.575	0.837	0.858
	110												0.163	0.142
PGM-2	90					0.100				0.067			0.117	0.196
	100	1.000	1.000	1.000	1.000	0.900	1.000	1.000	1.000	0.933	0.778	1.000	0.878	0.794
	110										0.222			0.010
PGM-3	90			0.037			1.000	0.265						
	100	1.000	1.000	0.963	1.000	1.000		0.735	0.902	1.000				
	105								0.098		0.741	0.925	1.000	1.000
	110										0.259	0.075		
6-PGDH	90								0.587	0.233				
	100	0.986	1.000	1.000	1.000	1.000	1.000	0.906	0.370	0.700	1.000	1.000		
	110	0.014						0.094	0.043	0.067			1.000	1.000
PGI	100	1.000	1.000	0.967	1.000	1.000	1.000	1.000	0.239	0.063	0.474	0.528		
	150			0.033					0.761	0.937	0.526	0.472	1.000	1.000
ACON	90										1.000	1.000		
	100	1.000	1.000	0.883	0.939	0.900	1.000	0.953						
	105			0.021										
	110			0.096	0.061	0.100		0.047	0.978	0.933			1.000	0.990
	120								0.022	0.067			1.000	0.010
IDH-2	50								1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	0.987	1.000	1.000	1.000	0.937	1.000						
	150		0.013				0.063							
*Kpn*I	Present							0.750						
	Absent	1.000	1.000	1.000	1.000	1.000	1.000	0.250						

[1] 1984 samples only.

TABLE 3. F_{IS}, F_{ST}, and $\varnothing$ values for *Polistes* females collected in the fall of 1984.

Species	Number of loci	$\varnothing$	Expected $\varnothing$[1]	F_{IS}	F_{ST}
P. bellicosus	3	0.58[2,3]	0.53	0.25[2]	−0.04
P. carolinus	6	0.36	0.38	−0.12	0.03
P. exclamans[4]	6	0.52[2]	0.49	0.19	0.77[2]
P. exclamans[5]	6	0.55[2]	0.38	−0.15	0.81[2]
P. metricus	6	0.66[2,3]	0.51	0.22[2]	−0.04

[1] The expected value of $\varnothing$ is 3/8 + (5/8)F_{IS} (full sibs).
[2] Significantly different from zero at the 1% level.
[3] Significantly different from the expected value for full sibs without inbreeding at the 1% level.
[4] Houston and Brazos Bend populations.
[5] Brazos Bend populations only.

The entire sample was surveyed for restriction site variability, and a total of 14,356 sites were identified and examined in the study as a whole. Virtually no heterogeneity in length or restriction sites was detected among the ribosomal repeats within each species. The only variability observed within a species was a *Kpn*I site gain in the nontranscribed spacer of *P. exclamans*. The site gain was virtually fixed in the 12 individuals from sample site V, although it was not found in any other individuals in the study. The *Kpn*I variant was treated as a one locus, two allele system for statistical analysis. Although this is an obvious oversimplification, and ribosomal genes may be present at several different loci, it should allow a conservative estimate of population subdivision. A limited amount of repeat length heterogeneity was seen in *Polistes dorsalis*, but the small sample sizes for this species precluded analysis.

The results of the F_{ST}, F_{IS}, and $\varnothing$ calculations are presented in Table 3 for the 1984 data and in Table 4 for the 1985 data. F_{ST} values for *Polistes carolinas, P. metricus,* and *P. bellicosus* were not significantly different from zero across the relatively small area encompassed by the study. F_{ST} values for *P. exclamans* were significantly different from zero at the 1% level, both with and without the Houston populations. In fact, the inclusion of the Houston sample did not significantly change the value of F_{ST}, even though the distance between the Houston locality and the remainder of the collecting localities was much greater than any of the distances within Brazos Bend. Also, the F_{ST} values for the Brazos Bend *P. exclamans* samples at the end of 1984 were significantly higher than the F_{ST} values for the same populations in the spring of 1985.

F_{IS} values differed significantly from zero for all species except *Polistes carolinus*. F_{IS} values for *P. bellicosus* were not significantly different from 1984 to 1985, but F_{IS} values for *P. exclamans* were significantly higher in spring 1985 than in fall 1984. The difference stems from the presence of nests homozygous for the PGM-3 90 allele or the *Kpn*I rDNA variant found in sample sites that lacked these markers in 1984. The anomaly could result from the dispersal of mated females bearing the PGM-3 90 allele from sample site IV or the *Kpn*I rDNA variant from sample site V to the other sample sites in the spring of 1985.

Virtually all values of $\varnothing$ were significantly different from zero, and many values were also significantly different from 3/8 (or 1/4 where appropriate). The observed values were as much as twice the values expected under haplodiploidy alone. The one exception to this trend was *P. carolinus,* whose $\varnothing$ value of almost exactly 3/8 was not significantly different from zero. This was also the only species with a nonsignificant value of F_{IS}. None of the observed values of $\varnothing$ differed significantly from the values predicted by Equations (6), (7), and (8), which take both haplodiploidy and local inbreeding into account. These results demonstrate that inbreeding can make a significant contribution to the coefficient of kinship within a nest.

DISCUSSION

Population Subdivision

The values of F_{ST} calculated for *P. exclamans* were much higher than local F_{ST} values reported for other eusocial hymenopterans. In ants, F_{ST} values across areas comparable to this study range from 0.042

TABLE 4. F_{IS}, F_{ST}, and ø values for *Polistes* samples collected in the spring of 1985.

Species	Number of loci	ø	Expected ø[1]	F_{IS}	F_{ST}
P. bellicosus					
Total	3	0.52[2,3]	0.51 (full sibs)	0.21[4]	−0.01
Workers only	3	0.56[2,3]	0.52 (full sibs)	0.19[4]	−0.12
Foundresses only	3	0.47[2]	0.38 (full sibs)	−0.12	−0.02
Foundresses/workers	3	0.11[5]	0.36 (aunt/niece) 0.41 (mother/daughter)		
P. exclamans					
Total	2	0.79[2,6]	0.70	0.52[2]	0.47[2]
Workers only	2	0.81[2,6]	0.68	0.49[2]	0.46[2]
Foundresses only	2	—	0.73	0.57[2]	0.51[2]
Foundresses/workers	2	0.64[2,3]	0.64	—	—

[1] The expected value of ø is 3/8 + (5/8)F_{IS} (full sibs), 1/4 + (3/4)F_{IS} (mother/daughter), or 3/16 + (13/16) F_{IS} (aunt/niece).
[2] Significantly different from zero at the 1% level.
[3] Significantly different from the expected value without inbreeding at the 5% level.
[4] Significantly different from zero at the 5% level.
[5] Significantly different from the expected value for a mother/daughter relationship and inbreeding.
[6] Significantly different from the expected value without inbreeding at the 1% level.

to 0.090 (Johnson et al., 1969; Tomaszewski et al., 1973; Ward, 1980; Pamilo, 1983). The cases known to date where F_{ST} values are as high as those in *P. exclamans* usually involve populations separated by physical factors that prove substantial barriers to gene flow. The isolation is typically exacerbated not only by habitat barriers but also by the low innate vagility of the species involved, such as plethodontid salamanders with an average F_{ST} of 0.53 (Larson et al., 1984). In the present study, a true physical barrier to gene flow does not exist. A typical individual of *P. exclamans* is roughly 2 cm in length and could easily fly across the entire study area in a single day if it chose to do so. The fact that sample site IV lacks an allele present in at least 70% frequency in every other sample site indicates that either individual *P. exclamans* do not disperse as far as they are physically capable or that those individuals that do disperse are not likely to survive and leave offspring. The rDNA site variant unique to sample site V also supports this idea.

Dispersal in *Polistes* occurs at two times in the colony cycle: in the fall when reproductives move to hibernacula to mate and overwinter and in the spring when mated females disperse from the hibernacula to found nests. The data presented here give an indication of the relative importance of these two dispersal times in *Polistes exclamans,* the only species sampled from a large number of sites in the fall and spring. In the fall 1984 sample, the PGM-3 90 allele was limited to sample site IV, while in the spring 1985 samples it was found in localities I and V as well. Three of the four nests with the 90 allele in sites I and V were fixed for it, and this pattern of fixation in the spring 1985 nests caused the increase in F_{IS} values seen for this species. The F_{IS} value for 1984, which was not significantly different from zero, probably represents the normal condition for this species, while the significant, nonzero F_{IS} value from 1985 is due to the Wahlund effect resulting from the migration of individuals into genetically different populations. The spread of the 90 allele in the population also caused the reduction in F_{ST} values for this species in spring 1985. Since the nests in question were founded by homozygotes mated to homozygotes, it appears that dispersal occurred after mating. It is further interesting to note that only one nest of PGM-3 90 homozygotes was found in 29 nests examined from sample site I on the south side of the Big Creek drainage, while two nests of PGM-3 90 homozygotes were found in five nests on the north side of the Big Creek drainage (Fig. 1). It seems likely that gene flow in *Polistes exclamans* is due primarily to the dispersal of mated individuals and that the pattern of their dispersal may be influenced by local topography.

The amount of dispersal seen in *Polistes exclamans* in the spring of 1985 seems large given the high F_{ST} values for this species.

However, the amount of gene flow depends on the reproductive success of the individual migrants. As mentioned previously, nest success in this species is low, and it is estimated that less than half of the individuals that found nests in the spring will actually contribute directly to the fall pool of reproductives (Strassmann, 1981, 1985). This low figure may be exacerbated for migrants by another factor, the absence of siblings to aid in nesting. In many species of *Polistes,* including *P. exclamans,* mated females try to found their own nests independently. Within a month, the majority of these initial nests have failed, and the females from these nests join existing nests. Strassmann and Hughes (unpublished results) have shown that nests with multiple foundresses are much more likely to succeed than nests with a single foundress. Studies have also shown that females preferentially join nests founded by nestmates from the previous season (Klahn, 1979; Gamboa et al., 1986), as was indicated by the relatedness data for *P. bellicosus* foundresses in 1985. Therefore, the ability of females to join nests with their relatives enhances the probability that they will successfully contribute to the next generation. If a foundress disperses, then she cannot simply join another nest of related individuals if her own nest fails, and no relatives will be available to recruit to her own nest if it should succeed. This means that the average reproductive success for a migrant could be much lower than for a nonmigrant and decreases the already low chance that migration will actually lead to gene flow.

Coefficients of Kinship and Local Inbreeding

The presence of local inbreeding in at least two of the four species of *Polistes* in this study had a major impact on local relatedness. A generalized form of Equations (6), (7), and (8) is given by

$$\varnothing_t = \varnothing_p + (1 - \varnothing_p)F_{IS} \qquad (9)$$

where $\varnothing_p$ is the coefficient of kinship of two individuals on a nest based on haplodiploidy alone and $\varnothing_t$ is the overall coefficient of kinship taking into account haplodiploidy and local inbreeding. If F_{IS} is zero, then $\varnothing_t$ for full sisters in a haplodiploid system is 0.375 and that between half sisters with different fathers is 0.125. With a background F_{IS} of 0.20, as was found in some samples of *Polistes bellicosus* and *P. metricus,* $\varnothing_t$ becomes 0.50 and 0.30 for full and half sisters, respectively. Note that this background kinship contributed by inbreeding not only can greatly increase the total coefficient of kinship for all types of relationships, but also tends to reduce the difference in the degree of kinship between different classes of pedigree related individuals. Thus, in the example given above, the coefficient of kinship between full sisters is three times that between half sisters if there were no background relatedness contributed by inbreeding. With only the moderate (for this study) value of F_{IS} seen in *P. bellicosus,* the total coefficient of kinship for full sisters is about 1.5 times that of half sisters. Thus the importance of inbreeding for calculations of inclusive fitness is not only that it increases the importance of coefficient of kinship, but also that distinctions between various types of pedigree relationships become less important. For example, any calculation of inclusive fitness in a species like *P. bellicosus* would reveal only half as much difference between full and half sisters if local inbreeding is taken into account. Alternatively, not considering local inbreeding would result in very misleading inclusive fitness calculations. Hence, theories of social behavior and evolution should not be based on considerations of haplodiploidy alone, but must include some consideration of population structure as well.

Acknowledgments

This research was supported by NIH Grant R01 GM31571 to A.R.T. and NSF Grant BSR 86-050263 to J.E.S.

Literature Cited

Andersson, M. 1984. The evolution of eusociality. Annu. Rev. Ecol. Syst. 15:165–189.

Arnheim, N., D. Treco, B. Taylor, and E. M. Eicher. 1982. Distribution of ribosomal length variants among mouse chromosomes. Proc. Natl. Acad. Sci. U.S.A. 79:4677–4680.

Berkelhamer, R. C. 1983. Intraspecific genetic variation and haplodiploidy, eusociality, and polygyny in the Hymenoptera. Evolution 37:540–545.

Clayton, J., and D. Tretiak. 1972. Amine-citrate buffers for pH control in starch gel electrophoresis. J. Res. Bd. Can. 26:1169–1172.

COCKERHAM, C. C., AND B. S. WEIR. 1987. Correlations, descent measures: Drift with migration and mutation. Proc. Natl. Acad. Sci. U.S.A. 84:8512–8514.

CROZIER, R. H. 1977. Evolutionary genetics of the Hymenoptera. Annu. Rev. Entomol. 22:263–288.

———. 1985. Adaptive consequences of male-haploidy, pp. 201–222. *In* W. Helle and M. Sabelis (eds.), Spider Mites, Their Biology, Natural Enemies and Control, Vol. 1A. Elsevier Science Publishers, Amsterdam.

CROZIER, R. H., P. PAMILO, AND Y. C. CROZIER. 1984. Relatedness and microgeographic genetic variation in *Rhytidoponera mayri,* an Australian arid-zone ant. Behav. Ecol. Sociobiol. 15:143–150.

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.

EDINGTON, E. S. 1987. Randomization Tests. Marcel Dekker, N.Y.

GAMBOA, G. J., H. K. REEVE, AND D. W. PFENNIG. 1986. The evolution and ontogeny of nestmate recognition in social wasps. Annu. Rev. Entomol. 31:431–454.

GRAUR, D. 1985. Gene diversity in Hymenoptera. Evolution 39:190–199.

HAMILTON, W. D. 1964*a*. The genetical evolution of social behavior. I. J. Theoret. Biol. 7:1–16.

———. 1964*b*. The genetical evolution of social behavior. II. J. Theoret. Biol. 7:17–52.

HILLIS, D. M., AND S. K. DAVIS. 1986. Evolution of ribosomal DNA: Fifty million years of recorded history in the frog genus *Rana.* Evolution 40:1275–1288.

JOHNSON, F. M., H. E. SCHAFFER, J. E. GILLASPY, AND E. S. ROCKWOOD. 1969. Isozyme genotype-environment relationships in natural populations of the harvester ant, *Pogonomyrmex barbatus,* from Texas. Biochem. Genet. 3:429–450.

KLAHN, J. E. 1979. Philopatric and nonphilopatric foundress association in the social wasp *Polistes fuscatus.* Behav. Ecol. Sociobiol. 5:417–424.

LARSON, A., D. B. WAKE, AND K. P. YANEV. 1984. Measuring gene flow among populations having high degrees of genetic fragmentation. Genetics 106:293–308.

LEARN, G. H., AND B. A. SCHAAL. 1987. Population subdivision in *Clematis fremontii*: Local differentiation for rDNA repeats. Evolution 41:433–437.

LESTER, L. J., AND R. K. SELANDER. 1979. Population genetics of haplodiploid insects. Genetics 92:1329–1345.

———. 1981. Genetic relatedness and social organization of *Polistes* colonies. Am. Nat. 117:147–166.

METCALF, R. A. 1980. Sex ratios, parent-offspring conflict, and local competition for mates in the social wasps *Polistes metricus* and *Polistes variatus.* Am. Nat. 116:642–654.

METCALF, R. A., AND G. S. WHITT. 1977. Intranest relatedness in the social wasp *Polistes metricus.* Behav. Ecol. Sociobiol. 2:339–351.

METCALF, R. A., J. C. MARLIN, AND G. S. WHITT. 1975. Low levels of genetic heterozygosity in Hymenoptera. Nature (London) 257:792–794.

OWEN, R. E. 1985. Difficulties with the interpretation of patterns of variation in the eusocial Hymenoptera. Evolution 39:201–205.

PAMILO, P. 1983. Genetic differentiation within subdivided populations of *Formica* ants. Evolution 37: 1010–1022.

———. 1984. Genotpyic correlation and regression in social groups: Multiple alleles, multiple loci and subdivided populations. Genetics 107:307–320.

———. 1985. Effect of inbreeding on genetic relatedness. Hereditas 103:195–200.

PAMILO, P., AND R. H. CROZIER. 1981. Genetic variation in male haploids under deterministic selection. Genetics 98:199–214.

———. 1982. Measuring genetic relatedness in natural populations: Methodology. Theoret. Pop. Biol. 21:171–193.

RANZANI, G. N., L. F. BERNINI, AND M. CRIPPA. 1984. Inheritance of rDNA spacer length variants in man. Mol. Gen. Genet. 196:141–145.

REEVE, H. K., J. S. REEVE, AND D. W. PFENNIG. 1985. Eusociality and genetic variability: A reevaluation. Evolution 39:200–201.

REYNOLDS, J., B. S. WEIR, AND C. C. COCKERHAM. 1983. Estimation of the coancestry coefficient: Basis for a short-term genetic distance. Genetics 105:767–779.

SAGHAI-MAROOF, M. S., K. M. SOLIMAN, R. A. JORGENSON, AND R. W. ALLARD. 1984. Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. U.S.A. 81:8014–8018.

SELANDER, R. K., M. SMITH, S. YANG, W. JOHNSON, AND J. GENTRY. 1971. Biochemical polymorphism and systematics in the genus *Peromyscus.* I. Variation in the old-field mouse, *Peromyscus polionotus.* Studies in Genetics VI. Univ. of Texas Publ. 7103, pp. 81–90.

SITES, J. W., AND S. K. DAVIS. 1989. Phylogenetic relationships and molecular variability within and among six chromosome races of *Sceloporus grammicus* (Sauria, Iguanidae) based on nuclear and mitochondrial markers. Evolution 43:63–75.

SNYDER, T. P. 1974. Lack of allozymic variability in three bee speciés. Evolution 28:687–688.

STRASSMANN, J. E. 1981. Evolutionary implications of early male and satellite nest production in *Polistes exclamans* colony cycles. Behav. Ecol. Sociobiol. 8:55–64.

———. 1985. Relatedness of workers to brood in the social wasp *Polistes exclamans* (Hymenoptera: Vespidae). Z. Tierpsychol. 69:141–148.

SYSTMA, K. J., AND B. A. SCHAAL. 1985. Phylogenetics of the *Lisianthus skinneri* (Gentianacae) species complex in Panama utilizing DNA restriction fragment analysis. Evolution 39:594–608.

TOMASZEWSKI, E., H. E. SCHAFFER, AND F. M. JOHNSON. 1973. Isozyme genotype-environment association in natural populations of the harvester ant *Pogonomyrmex badius.* Genetics 75:405–421.

WAGNER, A. E., AND D. A. BRISCOE. 1983. An absence

of enzyme variability within two species of *Trigona* (Hymenoptera). Heredity 50:97–103.

WARD, P. S. 1980. Genetic variation and population differentiation in the *Rhytidoponera impressa* group, a species complex of ants (Hymenoptera: Formicidae). Evolution 34:1060–1076.

WEIR, B. S., AND C. C. COCKERHAM. 1984. Estimating F-statistics for the analysis of population structure. Evolution 34:1060–1076.

WILLIAMS, S. M., R. DESALLE, AND C. STROBECK. 1985. Homogenization of geographical variants at the nontranscribed spacer of ribosomal DNA in *Drosophila mercatorum*. Mol. Biol. Evol. 2:338–346.

WILSON, G. N., M. KNOLLER, L. L. SZURA, AND R. D. SCHMICKEL. 1984. Individual and evolutionary variation of primate ribosomal DNA transcription initiation regions. Mol. Biol. Evol. 1:221–237.

Corresponding Editor: R. L. Honeycutt