Evolution, 48(4), 1994, pp. 1158–1167

HIERARCHICAL GENETIC STRUCTURE AND GENE FLOW IN MACROGEOGRAPHIC POPULATIONS OF THE EASTERN TENT CATERPILLAR (*MALACOSOMA AMERICANUM*)

JAMES T. COSTA III[1] AND KENNETH G. ROSS
Department of Entomology, University of Georgia, Athens, Georgia 30602-2603

Abstract.—Genetic structure and inferred rates of gene flow in macrogeographic populations of the eastern tent caterpillar *Malacosoma americanum* were analyzed at two hierarchical scales: local demes and regional subpopulations. Wright's F-statistics were used to estimate population genetic structure using multilocus genotypic data generated electrophoretically. Estimated values of F_{ST} and the distribution of private alleles were then used to obtain indirect estimates of gene flow. We found modest, though significant, genetic structure at both spatial scales, a pattern consistent with high rates of gene flow over the large distances involved. Modest values obtained for Nei's genetic distance also suggested high levels of gene flow across the range of this species, although some gene-flow restriction resulting from isolation by distance was suggested by a positive regression of genetic distance on geographic distance. The observed homogeneity at enzyme loci across the range of *M. americanum* parallels the reported uniformity in morphology, suggesting a general absence of local genetic differentiation in this widely distributed species. The genetic homogeneity observed in this wide-ranging insect is discussed in terms of organism-specific environmental experience at different spatial scales. Some organisms occupying apparently heterogeneous environments may ameliorate unsuitable local conditions through microhabitat selection or behavioral modification of their microenvironment. This may be accomplished in *M. americanum* through group shelter construction and behavioral thermoregulation, closely tying thermoregulation to social biology in this species. If in this way the tent helps produce an effectively homogeneous environment for this species across its extensive range, this system may provide a unique example of how social behavior can influence the distribution of genetic variation in a population.

Key words.—Gene flow, hierarchical genetic structure, Lasiocampidae, Lepidoptera, *Malacosoma americanum,* private alleles, tent caterpillars, Wright's F-statistics.

Received March 9, 1993. Accepted October 20, 1993.

Studies of population genetic structure seek to assess patterns in the distribution and movement of genetic variants in populations. These patterns are tested against the interrelated null hypotheses of an absence of genetic structure and high rates of gene flow (panmixis) as a first step in making inferences about the evolutionary processes at work in a population. Analyses of population differentiation may permit the assessment of the relative importance of drift, selection, and gene flow in determining present patterns (with respect to the markers used), as well as prediction of the likely "evolutionary trajectories" of populations (e.g., cohesion vs. divergence). In broad terms, the analysis of population genetic structure and its causes has been integral to a wide variety of evolutionary studies. The significance of genetic structure lies at the heart of controversies over likely modes of speciation (Mayr 1970; Futuyma and Mayer 1980; Lande 1982; Futuyma 1986a; Diehl and Bush 1989), host-race formation (Smith 1988; Carroll and Boyd 1992), and spatial patterns of adaptation (Taylor 1976; Rausher 1982; Hedrick 1986; Rank 1992).

There are several ways to gauge the genetic structuring of populations. In general, the focus can be on individuals and their movement patterns or on the effect of these movement patterns on the distribution of genetic variation in a population. An example of the former might be Wright's (1946) genetic neighborhoods, which are defined in terms of the mean dispersal distance of progeny. Examples of the latter include spatial autocorrelation (Sokal and Wartenberg 1983; Slatkin and Arter 1991) and Wright's (1951) F-statistics. The F-statistics, which are perhaps the most widely used measure of population differentiation, are based on the partitioning of genetic variance within and among subdivisions of the population, measuring the degree of inbreeding and the inbreedinglike effects of isolation in terms of the distribution of genetic variation.

While it is possible for populations to become

[1] Present address: Museum of Comparative Zoology, 26 Oxford St., Harvard University, Cambridge, Massachusetts 02138. E-mail: j_costa @ nocmsmgw.harvard.edu.

© 1994 The Society for the Study of Evolution. All rights reserved.

genetically differentiated in sympatry (for example, by behaviorally mediated assortative mating), spatial separation, even in continuously distributed populations, is probably the most common cause of genetic structure. The role of physical separation is not as simple as it may first appear, in that there exists a continuum in degree of separation. At one extreme is the isolation imposed by microhabitat or behavioral preference at small scales, while at another is isolation by large geographic distances. The latter follows from the relatively limited dispersal abilities of many organisms relative to their species' range and is manifested as a decreased probability of mating as the distance between birthplaces of individuals increases.

It has long been thought that isolation by distance promotes genetic differentiation of populations that span heterogeneous environmental regimes (Futuyma 1986b; Hartl and Clark 1989). This differentiation is a product of the joint effects of drift and adaptation to local conditions, the importance of both being inversely related to the degree of gene flow between subpopulations. Gene flow, meaning the dispersal of individuals or gametes, is therefore important in maintaining genetic cohesion and similarity among populations (Slatkin 1987). Isolation by geographic distance may result in a restriction of gene flow, thereby permitting population differentiation (especially if selection related to local ecological differences is acting as well).

Eastern tent caterpillar (*Malacosoma americanum*) populations offer a unique opportunity to study the ways in which genetic structure and gene flow patterns may change over increasingly greater distances and, perhaps more significantly, over increasingly heterogeneous environmental regimes (McCauley and Eanes 1987; Rank 1992). This social lepidopteran occupies a vast range, covering extensive latitudinal and altitudinal gradients, from southern Canada to northern Florida and from the Atlantic seaboard to west of the Mississippi River (Stehr and Cook 1968). Eastern tent caterpillars are of further interest in that they are largely specific to a patchily distributed host, black cherry trees (*Prunus serotina*). The host population thus imposes a microgeographic physical structure on the caterpillar populations, since the caterpillar is closely tied to its host's phenology, and there is significant geographic variation in seasonal timing of host tree development. The close interrelationship between these caterpillars and their host might be expected to facilitate large-scale genetic differentiation of caterpillar populations, whereas significant movement of adults would oppose the development of such genetic structure.

Hierarchical lower-level genetic structure of eastern tent caterpillar populations was studied previously at scales including (1) the colony, (2) the host tree, (3) the host tree patch, and (4) local demes consisting of several host-tree patches (Costa and Ross 1993). Significant genetic structure was found at the colony level (individuals within colonies are closely related), but only modest genetic structure was observed at the higher spatial scales. The present study extends this analysis to macrogeographic scales, including a larger number of local demes (which correspond to the highest scale analyzed in the previous study), as well as regional subpopulations distributed along the east coast of the United States. Thus, *M. americanum* represents one of the few cases in which relatively complete population genetic structure has been specified, from individual colonies to a substantial portion of the species range. A multilevel analysis such as this can be especially informative in that gene flow patterns may vary at different spatial scales, and comparisons across levels aid in the identification of salient biological characteristics shaping the distribution of genetic variation at these different scales.

Methods

Specimen Collection

Forty caterpillars were collected from each of six demes nested within each of three regional subpopulations, yielding a total of 720 sampled individuals (see fig. 1). Sampled demes were approximately 1 km in diameter and were separated from one another by at least 100 km. The closest demes sampled in different regional subpopulations were separated by at least 200 km. A single caterpillar per colony was collected to avoid using nonindependent genotypes in the analysis (see Costa and Ross 1993).

Caterpillars were frozen immediately upon collection in the field in portable cryogenic containers containing liquid nitrogen. Specimens were later transferred to an ultralow-temperature freezer in the laboratory and maintained at −70°C until electrophoresis.

Electrophoretic Loci

Analyses of population genetic structure and gene flow are based on genotypic data derived

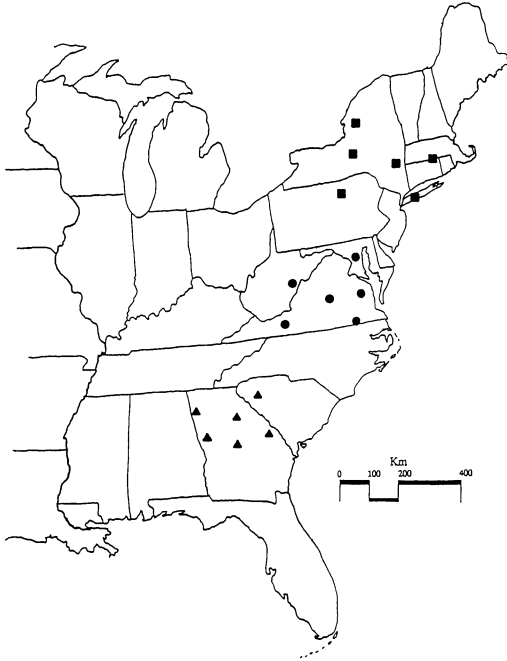

FIG. 1. Locations of sampled demes for studying macrogeographic population genetic structure in *Malacosoma americanum* in the eastern United States. One caterpillar was collected from each of 40 colonies in each deme. Each of three regional subpopulations is composed of six demes (demes in different subpopulations are distinguished by the different symbols).

from eleven electrophoretic loci that are polymorphic at the 99% level. These loci include: *Aat-1, 2* (EC 2.6.1.1), *β-Gala* (EC 3.2.1.53), *α-Glus-1* (EC 3.2.1.20), *Gpi* (EC 5.3.1.9), *Hbdh* (EC 1.1.1.30), *Idh* (EC 1.1.1.42), *Mdh-2* (EC 1.1.1.37), *Pep(pap)* (EC 3.4.11.–), and *Pgm-1, 2* (EC 5.4.2.2). The mean number of alleles per locus is 8.4 (range = 3–16) and the mean effective number of alleles is 1.6 (range = 1.0–2.3). Staining procedures and data confirming Mendelian inheritance for all of the loci are presented in Shoemaker et al. (1992).

Data Analyses

Hardy-Weinberg Tests.—Genotypic data for each locus were tested for conformation to Hardy-Weinberg expectations at each population level using chi-square goodness-of-fit tests (Lessios 1992) with Emigh's (1980) correction for continuity, and by adjusting significance levels using the sequential Bonferroni procedure (Hochberg 1988; Rice 1989). Adjustment of significance levels when performing multiple tests is important in order to minimize type I errors, as the chance of such errors increases with number of tests performed. The sequential Bonferroni procedure of Hochberg (1988) orders tests by probability value; the highest probability is compared to the initial significance level (α) and subsequent probabilities are compared to an adjusted significance value (α_{i+1}), where $\alpha_{i+1} = \alpha(1 + i)$ and i = the number of tests already performed (Lessios 1992). Rare alleles were pooled so that all expected genotypic frequencies were greater than one, and more than 80% of the frequencies were greater than five for each locus (Cochran 1954; Lessios 1992).

Population-Genetic Structure.—Population genetic structure was analyzed using the procedure of Weir and Cockerham (1984) to calculate Wright's (1951) F-statistics for two levels of population subdivision: (1) demes (lower-level) and (2) subpopulations (higher-level), as defined above under *Specimen Collection.* The F-statistics, or inbreeding coefficients, are hierarchically related descriptors of the distribution of genetic variation in populations. The total inbreeding coefficient (F_{IT}) can be decomposed into contributions from actual inbreeding (F_{IS}) and the inbreedinglike effects of population genetic structure (F_{ST}) (Wright 1951). We are especially interested in F_{ST}, also called the fixation index or standardized allele-frequency variance, as a measure of genetically effective movement patterns in the population.

The procedure of Weir and Cockerham (1984) for estimating F-statistics corrects lower-level genetic structure for the effects of higher-level structure, thereby providing unbiased estimates of the F-statistics at both levels. A jackknife procedure was employed to obtain estimates of the mean and variance for all F-statistics following the method of Weir and Cockerham (1984). The jackknife variance estimation was performed in two contexts: (1) over the alleles of each locus independently (yielding the variance among alleles at single loci), and (2) over the loci (yielding the variance among loci). The 95% confidence intervals (CI) were obtained from the standard errors by assuming the t distribution, and these were used to judge whether the mean values differed significantly from zero.

Gene Flow.—Gene flow, or the genetically effective migration rate (Nm), was estimated in two indirect ways: (1) using Wright's F_{ST} and (2) using the distribution of private alleles. Wright (1951) showed that the fixation index F_{ST} can be

used to estimate the effective migration rate using the formula

$$Nm = (1 - F_{ST})/4F_{ST},$$

if an infinite island model of population structure and gene flow is assumed. Although few populations actually conform to the assumptions of this model, primarily because of neglect of the effects of isolation by distance, it is useful as a first approximation of the relative magnitude of gene flow. We calculated Nm from our estimated values of F_{ST} for both levels of population subdivision for each locus individually and all loci collectively. The sole negative estimate of F_{ST} was taken as zero for the calculation of Nm. Slatkin and Barton (1989) point out that while small negative F_{ST} values are acceptable for describing the distribution of genetic variation in a population, these values are problematic for estimating gene flow. These authors interpret such negative values as meaning population structure is too small to be detected; thus, for convenience we take such a value to be zero. This truncation is further justified by the fact that the 95% CI for the negative value overlaps zero (see table 1).

The private alleles method was introduced by Slatkin (Slatkin 1985; Barton and Slatkin 1986; Slatkin and Barton 1989) and calculates Nm using the conditional average frequency of "private" alleles, which is simply the average frequency of alleles unique to a given deme or subpopulation. We employed the private alleles method to estimate Nm for all loci possible at both levels of population subdivision using the reference values based on $N_{Ref} = 50$ ($a = -0.612$, $b = -1.21$) for the linear regression equation (Slatkin and Barton 1989). The Nm estimates were adjusted for sample size using the procedure given by Slatkin and Barton (1989) and were jackknifed over loci to obtain means and their variances for the entire data set (summary values).

To determine the extent to which summary F_{ST} and Nm estimates were influenced by "aberrant" single-locus values, we employed box-and-whisker plots (Velleman and Hoaglin 1981; Hoaglin et al. 1983). This procedure divides the range of data values on the basis of spread about the median so we could, if necessary, eliminate outlying values from subsequent analyses. The "box" is simply the range of values from the first and third quartiles, and the "whiskers" are lines drawn to two classes of values ("outlying" and "extreme") lying outside this range. Outlying and

TABLE 1. Estimates of Wright's F-statistics (means ± SE) for two spatial scales in *Malacosoma americanum* populations. Single-locus values were obtained by jackknifing over alleles. Summary values were obtained by jackknifing over loci.

Locus	F_{IT}	Deme		Subpopulation	
		F_{ST}	F_{IS}	F_{ST}	F_{IS}
Aat-1	0.133 ± 0.061*	0.053 ± 0.014*	0.083 ± 0.042	0.048 ± 0.020*	0.091 ± 0.042*
Aat-2	0.004 ± 0.017	0.027 ± 0.013*	−0.023 ± 0.030	0.017 ± 0.008*	−0.012 ± 0.025
β-Gala	0.040 ± 0.029	0.014 ± 0.004*	0.027 ± 0.027	0.002 ± 0.002	0.038 ± 0.030
α-Glus-1	0.217 ± 0.058*	0.152 ± 0.025*	0.079 ± 0.088	0.031 ± 0.015*	0.194 ± 0.071*
Gpi	0.058 ± 0.021*	0.029 ± 0.007*	0.030 ± 0.027	0.009 ± 0.008	0.050 ± 0.025
Hbdh	−0.046 ± 0.082	0.018 ± 0.008*	−0.067 ± 0.076	0.007 ± 0.010	−0.056 ± 0.074
Idh	0.152 ± 0.119	0.015 ± 0.004*	0.139 ± 0.118	0.000 ± 0.002	0.153 ± 0.119
Mdh-2	0.126 ± 0.055*	0.006 ± 0.002*	0.120 ± 0.054*	0.002 ± 0.000	0.124 ± 0.055*
Pep(pap)	0.109 ± 0.015*	0.022 ± 0.004*	0.089 ± 0.018*	0.005 ± 0.007	0.105 ± 0.021*
Pgm-1	0.119 ± 0.055*	−0.004 ± 0.014	0.124 ± 0.069	0.000 ± 0.005	0.119 ± 0.060
Pgm-2	0.027 ± 0.030	0.020 ± 0.004*	0.007 ± 0.029	0.013 ± 0.007	0.015 ± 0.026
Summary	0.086 ± 0.031*	0.041 ± 0.018*	0.047 ± 0.019*	0.014 ± 0.003*	0.073 ± 0.031*

* Values differ significantly from zero.

extreme values are defined as those values within or beyond 1.5 interquartile distances from the first and third quartiles, respectively (Hoaglin et al. 1983).

Genetic Distance.—The relationship between genetic differentiation of demes and the geographic distance between them was assessed using Spearman's rank correlation test on Nei's (1978, 1987) genetic distance (D) between all pairs of demes. Nei's D was calculated using the program of Sattler and Hilburn (1985); our estimates of D are maximum estimates due to the exclusion of completely monomorphic loci from this study. Evidence for clinal distributions of alleles was sought by estimating Spearman's rank-correlation coefficient for latitudinal displacement and frequency of each of the alleles. Latitudinal displacement for each deme was defined relative to the southernmost deme. The sequential Bonferroni procedure was used to adjust significance levels in the rank correlation analyses.

Results

All 11 enzyme loci displayed genotypic proportions consistent with Hardy-Weinberg expectations at both levels of population spatial structure (all $P > 0.05$ using adjusted alpha levels from the sequential Bonferroni procedure). It should be emphasized that tests for deviation from Hardy-Weinberg expectations are generally sensitive only to gross departures from genotype equilibrium and are therefore used only as general tests of the suitability of the enzyme loci for population-genetic analyses relying on the assumption of neutrality.

Individual-locus and summary estimates for the F-statistics (means ± SE) are presented for demes and subpopulations in table 1. Summary estimates are, again, generated by summing variance components over all alleles and loci in the estimation procedure, and are not simple averages (Weir and Cockerham 1984). Population genetic structure was statistically detectable but modest in extent at both spatial levels of the population (summary values of F_{ST} were significantly greater than zero). Using box-and-whisker plots, we identified *Aat-1* and *α-Glus-1* as outlying and extreme loci, respectively, relative to the median for deme-level F_{ST} values. Both of these loci were outliers in terms of the F_{ST} values estimated at the subpopulation level. Elimination of allele frequency data for only *α-Glus-1* reduced the summary deme-level F_{ST} from 0.041 ± 0.018 to 0.024 ± 0.005 (means ± SE). Elimination of data for both *Aat-1* and *α-Glus-1* further reduced this estimate to 0.020 ± 0.004 and reduced the summary subpopulation-level F_{ST} from 0.014 ± 0.003 to 0.009 ± 0.002.

Gene flow (Nm) estimates obtained for each locus using the F_{ST} and private alleles methods are presented in figure 2. These estimates are generally concordant between the two methods of estimation; in 6 of 8 deme-level comparisons and 7 of 12 subpopulation-level comparisons the private alleles Nm estimates lie within the 95% CIs of the F_{ST}-derived Nm estimates. Perhaps more importantly, virtually all values are well above the theoretical threshold level at which gene flow is sufficient to homogenize populations genetically in the absence of selection ($Nm = 1$; Wright 1931; Slatkin 1987). Summary values for the lower (deme) population level are $Nm = 5.85$ (95% CI: 3.01–46.5) using F_{ST} and $Nm = 13.5$ (9.98–17.0) using private alleles; summary values for the higher (subpopulation) level are $Nm = 17.7$ (11.9–33.2) and $Nm = 13.3$ (3.26–23.4) using F_{ST} and private alleles, respectively. (Summary Nm values are presented with the 95% CI because the standard errors for Nm are asymmetrical due to the nonlinear relationship between F_{ST} and Nm).

To see to what extent outlying and extreme values identified with the box-and-whisker plots influence the summary Nm estimates, we recalculated Nm from both F_{ST} and private alleles without these values, as described above. At the deme level, this increases F_{ST}-derived Nm from 5.85 to 10.4 (7.39–17.2) (*α-Glus-1* excluded) or to 12.3 (8.93–19.7) (*α-Glus-1* and *Aat-1* excluded). For the private-alleles gene-flow estimates at the deme level, *Gpi* was an outlier, and *Aat-1* was extreme at the low end of the distribution. Removal of these loci from the jackknife procedure increases the Nm value from 13.5 to 14.9 (12.3–17.4) (*Aat-1* excluded), or to 16.2 (15.9–16.4) (*Aat-1* and *Gpi* excluded). The locus *Pgm-1* was an outlier at the high end of the distribution. Elimination of this datum decreases Nm from 13.5 to 13.0 (9.12–16.9). At the level of subpopulations, outlying and extreme private-alleles values were found at the high end of the distribution only; *β-Gala* is an outlier and *Idh* an extreme value. Removal of these data decreases Nm from 13.3 to 8.97 (3.53–14.4) (*Idh* excluded), or to 6.72 (3.30–10.1) (*Idh* and *β-Gala* excluded).

To summarize, elimination of aberrant single-locus Nm values changes the overall estimates

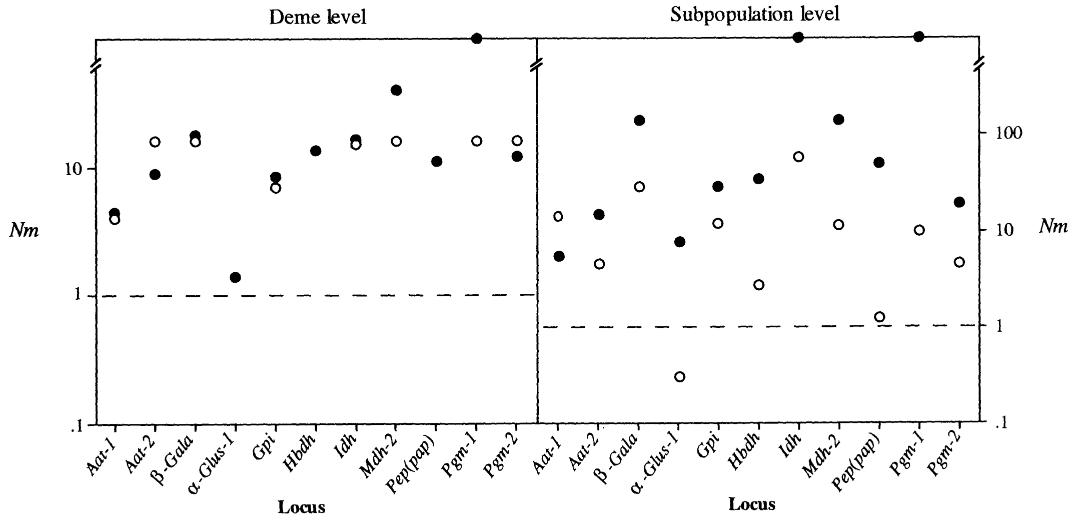

FIG. 2. Individual-locus gene-flow estimates for two levels of structure in *Malacosoma americanum* macrogeographic populations. Shaded circles, estimates of Nm based on F_{ST} (those straddling the upper bar correspond to F_{ST} values of zero and are therefore given as indeterminately high). Open circles, estimates of Nm based on private alleles; estimates are unavailable for three loci at the deme level due to the absence of private alleles for these markers. The dashed lines in each panel are placed at $Nm = 1.0$, the threshold value below which populations may diverge genetically by the action of drift alone (Wright 1931).

little. Both methods of analysis indicate modest genetic structure and high rates of gene flow at macrogeographic scales in eastern tent caterpillar populations for virtually all of the loci analyzed.

Estimates of F_{IS} at both scales studied here were only slightly significantly greater than zero (table 1); no outlying or extreme individual-locus F_{IS} values were found at either population level. At the deme level, F_{IS} and F_{ST} appear to contribute equally to total inbreeding (F_{IT}), while at the subpopulation level F_{IS} is the more important factor. Since total inbreeding is a net consequence of both actual inbreeding (F_{IS}) and population structure (F_{ST}), for a given level of total inbreeding a positive value of one component such as F_{IS} means that the other component (in this case F_{ST}) contributes less. Thus, the positive F_{IS} values we found indicate that the significantly positive value of F_{IT} is not attributable solely to population subdivision (F_{ST}) at the spatial scales studied here, but rather receives a contribution from inbreeding among individuals.

The modest degree of genetic structure inferred from estimates of the F-statistics is further evidenced by the fact that average genetic distance (D) between populations is very low, 0.0077 ± 0.0108 (SD). Moreover, this low value is a maximum estimate because completely monomorphic loci were not used in the estimation. Genetic distance increased significantly with geographic distance between pairs of populations (see fig. 3; Spearman rank correlation: $P = 0.0001$) indicating greater divergence at the loci we studied between more widely separated populations. Analyses of individual allele frequencies along the latitudinal gradient of the sampled demes failed to show any clinal genetic structure at any of the loci (Spearman rank correlations: $P > 0.05$ for all 92 tests using Bonferroni-adjusted α levels).

DISCUSSION

This study of eastern tent caterpillar macrogeographic population structure, together with earlier work on the microgeographic population structure of this species (Costa and Ross 1993), provides one of the most comprehensive analyses of hierarchical population-genetic structure for an insect, encompassing spatial levels from family units to a large portion of the species' range. Costa and Ross (1993) found that the only pronounced genetic structure at the lowest levels of spatial subdivision occurs at the level of the colony. This genetic structure arises from kin association; that is, colonies comprise families. Caterpillars living in the same host tree or group of trees are only very distantly related. The present study increases the scale of population-ge-

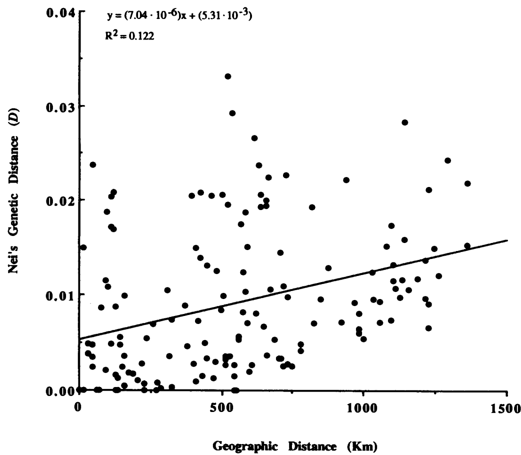

FIG. 3. Relationship between geographic distance and Nei's (1978, 1987) genetic distance (D) for pairs of *Malacosoma americanum* populations sampled from 18 demes (see fig. 1). D was calculated using data from polymorphic loci only. The drawn regression line provides a significant fit to the plotted points ($P = 0.0001$).

netic analysis and reveals, through estimates of Wright's F_{ST} and Nei's D, only modest genetic differentiation over moderate and large geographic distances. While genetic-distance values are quite low between all sampled demes, they do exhibit an increase with geographic distance. This pattern can only arise if there are some constraints to gene flow, perhaps coupled with selection associated with regional habitats acting on some loci. In general, however, dispersal and gene flow are inferred to be quite high over a range of microgeographic and macrogeographic spatial scales in this species.

To put these observations into perspective, it is useful to consider the broader context of what different patterns of population-genetic structure and gene flow may mean for patterns of selection and adaptation in populations. Population-genetic structure in insects such as *Malacosoma americanum* arises from a complex interplay of dispersal, mating patterns, host use patterns, selection, and genetic drift. Gene flow, or the lack thereof, plays a major role in shaping population genetic structure (Slatkin 1987) and influences possible patterns of adaptation in populations by determining levels of exchange of genetic material within and among subpopulations. In nonmigratory herbivorous species with ranges the size of that of *M. americanum,* the apparently large ecological gradients presented to the species (latitudinally and altitudinally) might be expected to result in patterns of local adaptation that are reflected in population differentiation at some elements of the genome, especially if the host plant is patchily distributed and the herbivore exhibits host-plant phenological dependence. This in turn might be expected to be reflected in differentiation at some "marker loci," if there are many such loci spread over the genome and some are linked to selected genomic elements (Hedrick 1986; Rank 1992). Contrary to these expectations, we observe only modest genetic differentiation of *M. americanum* populations over a wide geographic area at the genetic loci that we analyzed.

These expectations make the observed lack of structure (and inferred high rates of gene flow) in eastern tent caterpillar populations all the more striking and raise questions bearing on the experiential relationship between these organisms and their environment. Environmental regimes (in terms of the complement of biotic and abiotic factors perceived by the organism) may be more homogeneous or heterogeneous than they appear to human observers, even at macrogeographic spatial scales. Thus, while a species may exhibit high levels of gene flow over what appears to be a highly heterogeneous range, it is possible that the ecological regime actually experienced by individuals may differ trivially or not at all. Furthermore, because environmental homogeneity is a relative quality, species could be plastic with respect to some environmental variables and have narrow requirements with respect to others. This pattern of adaptation may produce clines or genetic structure with respect to some loci, while other loci indicate unrestricted rates of gene flow (Watt 1977; Power et al. 1983; Watt et al. 1983, 1985; Koehn and Hilbish 1987; DiMichele et al. 1991; Crawford and Powers 1992; Rank 1992).

Experientially, habitat similarity (homogeneity) in wide-ranging species might be achieved in two ways. First, since the environment at macrogeographic scales is a microenvironmental mosaic, organisms with narrow requirements may simply seek out the appropriate microhabitat so that they experience virtually identical conditions across the species range. Second, environmental homogeneity may be achieved through the action of the organism itself, by creation of suitable microenvironmental conditions. This might be accomplished through, for example, shelter construction, behavioral thermoregulation, or both, the net effect of which is to buffer

the organism from unsuitable environmental conditions.

Applying these ideas to eastern tent caterpillars, there are several possible explanations for the observed lack of genetic differentiation. First, it is possible that these caterpillars are rather plastic organisms, capable of dealing effectively with the range of ecological conditions that populations experience with one or a few general, all-purpose physiological genotypes (Baker 1965; Parsons 1983; Nevo et al. 1984). In this scenario, caterpillars could accommodate, for example, the differences in temperature range characterizing the northern and southern extremes of their range without specialized genotypes. All-purpose genotypes might be more efficient than narrowly specialized genotypes or coadapted gene complexes since the observed high rates of gene flow in *M. americanum* presumably would continually disrupt gene complexes underlying local adaptation by recombination with genetic material from other regions.

Second, it is possible that there are locally adapted genotypes or gene complexes, but that none of our markers are linked to them (with the possible exception of *α-Glus-1*). Presumably, these selected genomic elements would be relatively few in number and are not tying up large portions of the genome. Gene flow at these parts of the genome may be significantly retarded relative to the gene flow we have measured (which may represent the flux of neutral genomic elements). Our gene-flow estimates must be considered upper limits in this scenario.

Third, and perhaps most interesting, it is possible that the environmental regimes experienced by eastern tent caterpillar populations are not as heterogeneous as we perceive them to be. There are several potentially important mechanisms that would tend to homogenize the microenvironment occupied by these caterpillars, and these include both microhabitat seeking and microhabitat creation. One mechanism may be host specificity; eastern tent caterpillars have a relatively narrow host-plant range, feeding primarily on black cherry (*Prunus serotina*) and, to a lesser extent, on a few other rosaceous plants (Stehr and Cook 1968; Peterson 1986). Populations of these caterpillars from different parts of their range could in effect be experiencing similar microenvironmental conditions, despite the imposed physical geographic structure, in that they feed on a common host plant with common secondary compounds that is generally found in similar, disturbed, habitats across its range (black cherry is a weedy or pioneer species).

Another potentially important buffering mechanism may lie in the overwintering adaptations of eastern tent caterpillar egg masses. Carmona and Barbosa (1983) suggest that the spumaline coating of the egg mass acts to ameliorate low overwintering temperatures through its hygroscopic and UV absorptive properties; these authors reported that spumaline-covered egg masses in sunlight were significantly warmer than ambient temperatures under a wide range of naturally occurring temperature regimes. Caterpillars of this species eclose and are active during periods of suitable temperature range and are in diapause during unsuitable periods. The egg-mass adaptations may buffer the temperature and humidity extremes experienced during quiescence.

The social behaviors of *M. americanum* may represent another potentially important means of homogenizing the microenvironment across the species' range. These caterpillars exhibit what is perhaps the most complex social organization known among noneusocial insects, exhibiting coordinated tent (nest) construction and cooperative foraging by means of a trail-based pheromonal recruitment system (Fitzgerald 1976; Fitzgerald and Peterson 1983; Fitzgerald and Willer 1983). Sociality may provide a unique form of microenvironmental buffering by virtue of the regulation of temperature and humidity conditions in the nest. Based on the principles of the niche-variation hypothesis, this social buffering of environment has been hypothesized to play a role in the reduced levels of genetic variation reported in many social insect species (Powell 1971; Snyder 1974; Berkelhamer 1983; Shoemaker et al. 1992; but see Graur 1985; Crespi 1991). The social behaviors of eastern tent caterpillars appear to permit effective thermal regulation of their microenvironment, despite the temperature extremes experienced by these caterpillars in different parts of their range. Casey et al. (1988) suggest that these caterpillars thermoregulate through a set of basking and foraging behaviors centered around the microenvironment created by the group-constructed tent. Knapp and Casey (1986) showed that these caterpillars experience temperatures 8°–13°C above ambient temperature as a consequence of aggregating in the insulated tent. If social buffering in this species leads to environmental homogenization, then the niche-variation hypothesis pre-

dicts minimal population-genetic structure, which we have found, as well as relatively low overall levels of genetic variation for the species. Based on an earlier survey of 47 enzyme loci, *M. americanum* does exhibit low levels of genetic variation relative to other Lepidoptera and nonsocial insects (Shoemaker et al. 1992).

In conclusion, minimal genetic structure and correspondingly high levels of gene flow occur at macrogeographic spatial scales in *M. americanum* populations. The genetic homogeneity we observed among different geographic populations of *M. americanum* parallels the morphological and behavioral uniformity of this species across its range (Stehr and Cook 1968), and contributes to an emerging view of this organism as a relatively monomorphic species with similar morphology, habits, host-plant use, and social behaviors across an extensive geographic range. This conclusion is important because it suggests that the results of an increasing number of physiological, ecological, and behavioral studies on this economically important insect may be quite general for the species. Furthermore, this genetic homogeneity may reflect the effective experience of environmental homogeneity on the part of *M. americanum* over an apparently heterogeneous geographic range. This underscores the importance of recognizing microenvironmental mosaicism and the organism-specific nature of environmental experience in interpreting patterns of genetic variation within and among natural populations.

Acknowledgments

We thank T. Fitzgerald, A. Greene, M. Costa, and especially L. Costa for their assistance in the field over the course of this study. We also thank J. Cook and B. Flaherty for their invaluable laboratory assistance, and J. Arnold for assistance with the University of Georgia BioScience Computing Resource and advice on data analysis. The manuscript benefited greatly from the comments of J. Hamrick, an anonymous reviewer, and especially W. Watt. This research was supported by National Science Foundation grant BSR-8901373 to K.G.R. and J.T.C., Sigma Xi Grants-in-Aid of Research, and a Sigma Chemical Company. Cooperative Allowance Grant to J.T.C., and the Georgia Agricultural Experiment Stations, University of Georgia.

Literature Cited

Baker, H. G. 1965. Characteristics and modes of origins of weeds. Pp. 147–168 *in* H. G. Baker and G. L. Stebbins, eds. The genetics of colonizing species. Academic Press, New York.

Barton, N. H., and M. Slatkin. 1986. A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. Heredity 56:409–415.

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

Carmona, A. S., and P. Barbosa. 1983. Overwintering egg mass adaptations of the eastern tent caterpillar, *Malacosoma americanum* (Fab.) (Lepidoptera: Lasiocampidae). Journal of the New York Entomological Society 91:68–74.

Carroll, S. P., and C. Boyd. 1992. Host race radiation in the soapberry bug: natural history with the history. Evolution 46:1052–1069.

Casey, T. M., B. Joos, T. D. Fitzgerald, M. E. Yurlina, and P. A. Young. 1988. Synchronized group foraging, thermoregulation, and growth of eastern tent caterpillars in relation to microclimate. Physiological Zoology 61:372–377.

Cochran, W. G. 1954. Some methods for strengthening the common χ^2 tests. Biometrics 10:417–451.

Costa, J. T. III, and K. G. Ross. 1993. Seasonal decline in intracolony genetic relatedness in eastern tent caterpillars: implications for social evolution. Behavioral Ecology and Sociobiology 32:47–54.

Crawford, D. L., and D. A. Powers. 1992. Evolutionary adaptation to different thermal environments via transcriptional regulation. Molecular Biology and Evolution 9:806–813.

Crespi, B. J. 1991. Heterozygosity in the haplodiploid Thysanoptera. Evolution 45:458–464.

Diehl, S. R., and G. L. Bush. 1989. The role of habitat preference in adaptation and speciation. Pp. 345–365 *in* D. Otte and J. A. Endler, eds. Speciation and its consequences. Sinauer. Sunderland, Mass.

DiMichele, L., K. T. Paynter, and D. A. Powers. 1991. Evidence of lactate dehydrogenase-B allozyme effects in the teleost, *Fundulus heteroclitus.* Science 253:898–900.

Emigh, T. H. 1980. A comparison of tests for Hardy-Weinberg equilibrium. Biometrics 36:627–642.

Fitzgerald, T. D. 1976. Trail marking by larvae of the eastern tent caterpillar. Science 194:961–963.

Fitzgerald, T. D., and S. C. Peterson. 1983. Elective recruitment communication by the eastern tent caterpillar (*Malacosoma americanum*). Animal Behaviour 31:417–423.

Fitzgerald, T. D., and D. E. Willer. 1983. Tent-building behavior of the eastern tent caterpillar *Malacosoma americanum* (Lepidoptera: Lasiocampidae). Journal of the Kansas Entomological Society 56:20–31.

Futuyma, D. J. 1986a. The role of behavior in host-associated divergence in herbivorous insects. Pp. 295–302 *in* M. D. Huettel, ed. Evolutionary genetics of invertebrate behavior. Plenum, New York.

———. 1986b. Evolutionary biology, 2d ed. Sinauer, Sunderland, Mass.

Futuyma, D. J., and G. C. Mayer. 1980. Non-allopatric speciation in animals. Systematic Zoology 29:254–271.

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

Hartl, D. L., and A. G. Clark. 1989. Principles of

population genetics, 2d ed. Sinauer, Sunderland, Mass.

Hedrick, P. W. 1986. Genetic polymorphism in heterogeneous environments: a decade later. Annual Review of Ecology and Systematics 17:535–567.

Hoaglin, D. C., F. Mosteller, and J. W. Tukey. 1983. Understanding robust and exploratory data analysis. Wiley, New York.

Hochberg, Y. 1988. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 75: 800–802.

Knapp, R., and T. M. Casey. 1986. Thermal ecology, behavior, and growth of gypsy moth and eastern tent caterpillars. Ecology 67:598–608.

Koehn, R. K., and T. J. Hilbish. 1987. The adaptive importance of genetic variation. American Scientist 75:134–141.

Lande, R. E. 1982. Rapid origin of sexual isolation and character divergence in a cline. Evolution 36: 213–223.

Lessios, H. A. 1992. Testing electrophoretic data for agreement with Hardy-Weinberg expectations. Marine Biology 112:517–523.

Mayr, E. 1970. Populations, species, and evolution. Harvard University Press, Cambridge, Mass.

McCauley, D. E., and W. F. Eanes. 1987. Hierarchical population structure analysis of the milkweed beetle, *Tetraopes tetraophthalmus* (Forster). Heredity 58:193–201.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590.

———. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nevo, E., A. Beiles, and R. Ben-Shlomo. 1984. The evolutionary significance of genetic diversity: ecological, demographic and life history correlates. Pp. 13–213 *in* G. S. Mani, ed. Evolutionary dynamics of genetic diversity. Lecture Notes in Biomathematics 53. Springer, Berlin.

Parsons, P. A. 1983. The evolutionary biology of colonizing species. Cambridge University Press, Cambridge.

Peterson, S. C. 1986. Host specificity of trial marking to foliage by eastern tent caterpillars, *Malacosoma americanum.* Entomologica Experimentalis et Applicata 42:91–96.

Powell, J. 1971. Genetic polymorphism in varied environments. Science 174:1035–1036.

Powers, D. A., L. DiMichele, and A. R. Place. 1983. The use of enzyme kinetics to predict differences in cellular metabolism, developmental rate and swimming performance between LDH-B genotypes of the fish *Fundulus heteroclitus.* Pp. 147–170 *in* M. C. Ratassi, J. G. Scandalois, and G. S. Whitt, eds. The isozymes: current topics in biological and medical research, Vol. 10. Liss, New York.

Rank, N. E. 1992. A hierarchical analysis of genetic differentiation in a montane leaf beetle *Chrysomela aeneicollis* (Coleoptera: Chrysomelidae). Evolution 46:1097–1111.

Rausher, M. D. 1982. Population differentiation in *Euphydryas editha* butterflies: larval adaptation to different hosts. Evolution 36:581–590.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223–225.

Sattler, P. W., and L. R. Hilburn. 1985. A program for calculating genetic distance, and its use in determining significant differences in genetic similarity between two groups of populations. Journal of Heredity 76:400.

Shoemaker, D. D., J. T. Costa III, and K. G. Ross. 1992. Estimates of heterozygosity in two social insects using a large number of electrophoretic markers. Heredity 69:573–582.

Slatkin, M. 1985. Rare alleles as indicators of gene flow. Evolution 39:53–65.

———. 1987. Gene flow and the geographic structure of natural populations. Science 236:787–792.

Slatkin, M., and H. E. Arter. 1991. Spatial autocorrelation methods in population genetics. American Naturalist 138:499–517.

Slatkin, M., and N. H. Barton. 1989. A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43:1349–1368.

Smith, D. C. 1988. Heritable divergence of *Rhagoletis pomonella* host races by seasonal asynchrony. Nature 336:66–67.

Snyder, T. P. 1974. Lack of enzyme variability in three bee species. Evolution 28:687–688.

Sokal, R. R., and D. E. Wartenberg. 1983. A test of spatial autocorrelation analysis using an isolation-by-distance model. Genetics 105:219–237.

Stehr, F. W., and E. F. Cook. 1968. A revision of the genus *Malacosoma* Hübner in North America (Lepidoptera: Lasiocampidae): systematics, biology, immatures, and parasites. U.S. National Museum Bulletin 276.

Taylor, C. E. 1976. Genetic variation in heterogeneous environments. Genetics 83:887–894.

Velleman, P. F., and D. C. Hoaglin. 1981. Applications, basics, and computing of exploratory data analysis. Duxbury, Boston.

Wade, M. J. 1982. Group selection: migration and the differentiation of small populations. Evolution 36:949–961.

Watt, W. B. 1977. Adaptation at specific loci. I. Natural selection on phosphoglucose isomerase of *Colias* butterflies: biochemical and population aspects. Genetics 87:177–194.

Watt, W. B., R. C. Cassin, and M. S. Swan. 1983. Adaptation at specific loci. III. Field behavior and survivorship differences among *Colias* PGI genotypes are predictable from in vitro biochemistry. Genetics 103:725–739.

Watt, W. B., P. A. Carter, and S. M. Blower. 1985. Adaptation at specific loci. IV. Differential mating success among glycolytic allozyme genotypes of *Colias* butterflies. Genetics 109:157–175.

Weir, B. S., and C. C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358–1370.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

———. 1946. Isolation by distance under diverse systems of mating. Genetics 31:39–59.

———. 1951. The genetical structure of populations. Annals of Eugenics 15:323–354.

Corresponding Editor: W. Watt