AN ANALYSIS OF GENE FLOW AMONG MIDWESTERN POPULATIONS OF THE MOSQUITO OCHLEROTATUS TRISERIATUS

ERIC T. BECK,* CHRISTOPHER F. BOSIO, DAVID A. GESKE, CAROL D. BLAIR, BARRY J. BEATY, AND WILLIAM C. BLACK, IV
Arthropod-Borne and Infectious Diseases Laboratory, Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, Colorado; Division of Vector Control, La Crosse County Health Department, La Crosse, Wisconsin

Abstract. A population genetics study of the mosquito Ochlerotatus triseriatus was performed on 36 collections from adjoining regions of Iowa, Minnesota, and Wisconsin covering approximately 120 km². Single nucleotide polymorphism analysis was used to estimate variation in the mitochondrial NADH dehydrogenase subunit 4 (ND4) gene. The heated oligonucleotide ligation assay was used to identify the ND4 haplotype of each mosquito. No evidence of genetic isolation by distance was found, nor did Interstate 90 or the Mississippi River serve as barriers to gene flow. The effective migration rate varied from 18 to 45 reproductive migrants/generation, which is similar to estimates from an earlier study. The collections belong to a single, large, panmictic population. However, within this panmictic population, local genetic drift arises, possibly due to a single or a few females ovipositing in larval breeding containers. From generation to generation, there is sufficient gene flow to mix families arising from individual breeding sites and eliminate founder effects due to drift.

INTRODUCTION

The Eastern treehole mosquito, Ochlerotatus triseriatus (Sensu Reinart, 2000),1 is the primary vector of La Crosse virus (LACV) in the Midwestern United States.2 Sinsko and Craig suggested that the species may exist in “ecological islands” maintained by limited gene flow between forested tracts of land in otherwise agricultural landscapes.3 Matthews and Craig used variation at 14 allozyme loci to examine gene flow among collections from Michigan, Indiana, and Illinois.4 They found that collections shared >99% of their genetic markers suggesting, contrary to Sinsko and Craig, that Oc. triseriatus exists as a large panmictic population.5,6

In this study, we use mitochondrial DNA (mtDNA) markers to reassess the results of Matthews and Craig.4 Mitochondrial DNA is maternally inherited and does not recombine; therefore, it can be used to examine maternal lineages.3 Mitochondrial DNA has previously been used to study phylogenetic relationships in several Anopheles species, as well as in the mosquito Aedes (Steigomyia) aegypti.6–13

Ochlerotatus triseriatus were collected from 36 sites in southwestern Wisconsin, southeastern Minnesota, and northeastern Iowa covering an area of approximately 120 km². We assessed whether the Mississippi River and Interstate 90 serve as barriers to gene flow in adjoining regions of the states of Iowa, Minnesota, and Wisconsin. We also determined genetic diversity at each collection site. We examined the effective migration rates (Nm) (the number of reproductive migrants per generation) and the effective population size (Ne) (the number of reproductive individuals).

MATERIALS AND METHODS

Mosquito collection and isolation of DNA. Ochlerotatus triseriatus eggs were collected from 5 oviposition traps in each of the 36 sites listed in Table 1. Collections were made from mid-July through September of 2002 by the La Crosse County Health Department in areas where La Crosse encephalitis cases occurred or that contained clusters of people at risk (e.g., wooded areas adjacent to houses with children, schools or playgrounds) (Figure 1). Each trap consisted of a can (6.5 × 11 cm) painted black, half filled with tap water, and containing paper toweling along the inside perimeter. Each trap was placed at or slightly above ground level. Egg papers were recovered from traps after 10 days, sent to Colorado State University, hatched, and reared to adults. Mosquitoes were analyzed for LACV infection by immunofluorescence assay.14 DNA was extracted from the thorax of each mosquito using the salt extraction method.15 The DNA was dissolved in 200 μL of Tris-EDTA (10 mM Tris, 1 mM EDTA), pH 8.0, buffer and stored at −70°C.

Polymerase chain reaction (PCR) amplification, single-strand conformation polymorphism (SSCP) analysis, and DNA sequencing. The mitochondrial NADH dehydrogenase subunit 4 gene (ND4) was amplified using the thermocycling parameters and the primers described previously.12 However, Tag DNA polymerase (Promega, Madison, WI) was added at the beginning of each reaction. The SSCP analysis follows that of Black and DuTeau.13 Prior to loading the gel, 4 μL of PCR product was combined with 4 μL of denaturing loading buffer and heated at 95°C for 5 minutes and placed directly on ice for 5 minutes. Samples were loaded onto a 3% polyacrylamide gel containing 1x Tris-borate-EDTA buffer (89 mM Tris, 89 mM borate, 2 mM EDTA) and electrophoretically separated using an IPC gel apparatus (Bio-Rad, Hercules, CA) at 15 milliamps for 10–12 hours. Gels were stained with silver and examined for different banding patterns suggestive of differences in primary sequences.15 DNA sequencing was performed by Davis Sequencing, Inc. (Davis, CA). Fifteen samples were sequenced representing all four ND4 haplotypes. Samples from 27 of the 36 sites were chosen for SSCP screening. The chosen sites for SSCP sampling covered the entire region of the study.

Single nucleotide polymorphism (SNP) analysis. The oligonucleotide ligation assay (OLA) is an inexpensive SNP assay that uses ligation between a biotinylated allele-specific detector and a 3’ fluorescein labeled reporter oligonucleotide. Heated OLA (HOLA) uses a thermal stable ligase and cycles of denaturing and reannealing on a thermal cycler for ligation

* Address correspondence to Eric T. Beck, Arthropod-Borne and Infectious Diseases Laboratory, Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO 80523. E-mail: etbeck@lamar.colostate.edu

80523. E-mail: etbeck@lamar.colostate.edu

534
and SNP detection. Allele-specific detectors (bold and underlined bases indicate polymorphic sites in the ND4 gene) at SNP site 21 were ND4tris21A-dtc: 5'-Biotin-CCT AAG GCY CAT GTT GAA GCT-3', and ND4tris21G-dtc: 5'-Biotin-CCT AAG GCY CAT GTT GAA GCC-3', and the reporter was ND4tris21-rpt: 5'-PO4-CCT GTT TCA GGA TCA ATA A-Fluorescein-3'. Allele-specific detectors at SNP site 23 were ND4tris23C-dtc: 5'-Biotin-GCT TAT TCT GTT GAC CAT ATG-3', and ND4tris23T-dtc: 5'-Biotin-GCT TAT TCT GTT GAC CAT ATA-3', and the reporter was ND4tris23-rpt: 5'-PO4-GGA ATT GTA TTA AGA GGG T-Fluorescein-3'. One 96-well plate was used for an ND4 PCR, a second plate for HOLA, and a third for ligation detection. Successful ligation was detected using horseradish peroxidase (HRP)--conjugated, anti-fluorescein antibody (Roche, Indianapolis, IN). The HRP activity was detected by addition of 3,3',5,5' tetramethylbenzidine (Sigma, St. Louis, MO).

Population genetics analyses. The ND4 frequencies were estimated in each collection using ARLEQUIN version 2.000 (University of Geneva, Geneva, Switzerland). Genetic diversity (the probability that two randomly chosen haplotypes in a collection are different) was estimated in each collection (equation 8.17). This genetic diversity is equivalent to heterozygosity in diploid data. Pairwise $F_{ST}$ (a measure of non-random mating among subpopulations) and linearized $F_{ST}$ $(F_{ST} / [1 - F_{ST}])$ values were estimated among populations and analyses of molecular variance (AMOVA) were performed. The data were analyzed by grouping sites to the north and south of Interstate 90, sites to the east and west of the Mississippi River, and sites in the northwest, northeast, southwest, and southeast quadrants of the study area using the interstate highway and the river as the quadrant dividers. We constructed a dendrogram among all of the collections using the NEIGHBOR program in PHYLIP3.61 using linearized $F_{ST}$ values. This value was also plotted against the natural logarithm of pairwise geographic distances among collections to test for isolation by distance. We used the Mantel test to determine if geographic distance acts as a barrier to gene flow in Oc. triserius.21

ArcGIS (Environmental Systems Research Institute, Redlands, CA) was used to examine the spatial distribution of genetic diversity. Genetic diversity was transformed by taking the arcsine of the square root of genetic diversity to linearize the data (equation 13.5) because values near 0 or 1 (transformed values) are non-linear. Inverse distance weighted (IDW) interpolation was used to predict genetic diversity values outside of collection sites. IDW assumes that genetic diversity values will be more similar among nearby points. The interpolated diversities are therefore a weighted average of the scatter points where the weights are assigned based on the distance to each scatter point. The weight diminishes as the...
distance from the interpolation point to the scatter point increases. Similarly, population genetics assumes that genetic diversity should be spatially autocorrelated. Proximate collection sites are more likely to have similar genetic diversity than distant sites.

RESULTS

Mitochondrial haplotype analysis. We used SSCP analysis on 564 of the 1,697 mosquitoes from different collection sites encompassing the entire study area. Only four different ND4 haplotypes were observed (Figure 2). These contained eight variable sites, seven of which were in complete linkage disequilibrium. Haplotypes 1 and 3, as well as haplotypes 2 and 4, differed by only one basepair and all haplotypes could be differentiated using SSCP analysis. All eight substitutions among the four haplotypes occurred in a third codon position and seven involved transitions. Only substitution 234 encoded a nonsynonymous Met ←→ Ile mutation.

Frequencies of haplotypes 1–4 among 564 mosquitoes were 0.812, 0.023, 0.158, and 0.007, respectively (Table 2). The SSCP analysis required more work and haplotypes were more difficult to discern than with the HOLA SNP detection assay. The haplotypes determined by SSCP and HOLA were identical in all 564 mosquitoes. Haplotypes were therefore determined using HOLA. Frequencies of haplotypes 1–4 among all 1,697 mosquitoes by HOLA were 0.784, 0.031, 0.183, and 0.003, respectively (Table 2).

Nested analysis of haplotype frequencies. The ND4 haplotype frequencies were partitioned using AMOVA within collections, among collections in a quadrant, and among quadrants. Most (86.3%) of the variation arose within collections, 13.2% arose among collections within a quadrant, and very little arose among quadrants (0.6%) (Table 3). The effective migration rates were determined using the equation \( N_e \mu = (1 - F_{ST})/4F_{ST}^{20} \). The \( N_e \mu \) among mosquitoes in all four quadrants was 45 reproductive migrants/generation, an extremely high rate of gene flow.

Haplotype frequencies were next partitioned within collections, among collections, and between collections north or south of Interstate 90. Again, most (86.9%) of the variation arose within collections, whereas the remainder arose among collections either north or south of the interstate. There was no variance among collections on either side of the interstate (Table 4). Interstate 90 does not serve as a barrier to gene flow.

Results of the AMOVA comparing collections east and
west of the Mississippi River were similar (Table 5). Most (86.2%) of the variance arose within collections, 12.4% arose among collections located either east or west of the river, and only 1.4% occurred between collections on either side of the river. The $N_{m/m}$ between groups on either side of the Mississippi River was 18 reproductive migrants/generation.

Cluster analysis and genetic diversity. Cluster analysis placed the collections into two groups. One group contained collections with high genetic diversity and the other contained collections with low diversity. Clusters did not correspond to the geographic locations of collections. The variation in genetic diversity within each collection was enormous, varying between 0.0000 for collections WHVR, INNL, STEU, and WAZP to 0.5846 in the VICT collection (Figure 3). In many cases, populations that were located within a few kilometers of one another had large differences in genetic diversity.

Transformed genetic diversity values were used to predict genetic diversity throughout the study area (Figure 1). The equation of the model is $y = 0.3093x + 22.818$, $r^2 = 0.195$, where $y$ is the predicted genetic diversity and $x$ is the arcsine of the square root of genetic diversity. The index of dispersion ($ID = \text{variance} / \text{mean}$, equation 4.1$^{23}$) was estimated using the mean and variance of the transformed values to test the hypothesis that genetic variability was randomly distributed over the study area. The ID was significantly greater than 1 ($P = 1.77 \times 10^{-8}$), indicating a contagious or clustered distribution of variability. This is evident from the map in that there is no smooth pattern of increasing or decreasing genetic diversity. Also, the map shows that sites of high variability are often located very close to sites with low variability (e.g., CALB and CALG, VICT and NEAL, and PDCB and PPSP).

**Figure 2.** Sequence alignment of the four mitochondrial NADH dehydrogenase subunit 4 (ND4) gene haplotypes detected by single-strand conformation polymorphism analysis of 564 Ochlerotatus triseriatus (Oc. tris) mosquitoes. Bases 21 and 234 were used for single-strand conformation polymorphism analysis with the heated oligonucleotide ligation to determine the ND4 haplotype of all 1,697 mosquitoes.
The linearized pairwise F\text{ST} in this study area
This 10-fold disparity
\textit{Oc. triseriatus}†
Variability ranged from 0.000 to
among quadrants
\textit{Ochlerotatus triseriatus}
\textit{Ochlerotatus triseriatus}
Only a
from throughout
"Ae. aegypti" among
is low when compared with the same gene in
between north and south of I-90
\textit{Ae.} collections
P in the eastern half of North America including Mexico will
Total 1,646 290.758 0.17760
Among collections
in a collection
1,612 247.043 0.15325 86.87
Total 1,646 290.758 0.17721
* Fixation indices: F\text{ST}north and south of I-90 = -0.002 (no barrier to gene flow); F\text{ST}among collections in region = 0.133; F\text{ST}haplotype in a collection = 0.131†
† P ≤ 0.0001.

**DISCUSSION**

The amount of diversity in the mtDNA ND4 gene in \textit{Oc. triseriatus} is low when compared with the same gene in \textit{Ae. aegypti}. A similar sample of \textit{Ae. aegypti} from throughout Mexico detected 25 different ND4 haplotypes.\textsuperscript{12,13} Only a survey from throughout the geographic range of \textit{Oc. triseriatus} in the eastern half of North America including Mexico will determine if low mtDNA diversity is general to the species. Such a survey should also involve nuclear diploid markers. A continued pattern of low mtDNA diversity would suggest a historical bottleneck in \textit{Oc. triseriatus}. Alternatively, high mtDNA diversity in other collections would suggest that a local historical bottleneck occurred during the establishment of northern midwestern \textit{Oc. triseriatus} populations.

We estimated F\text{ST} among \textit{Oc. triseriatus} collections from northwestern Indiana and southwestern Michigan using the allozyme frequency data at 14 enzyme loci reported by Matthews and Craig.\textsuperscript{5} The N/m of 22 reproductive migrants/generation estimated from their study falls within the range of 18–45 reproductive migrants/generation found in the present study. Our results therefore largely agree with the results of Matthews and Craig and do not support the suggestion by Sinsko and Craig that the species exists in “ecological islands” in agricultural landscapes.\textsuperscript{3,4}

A major difference between our results and those of Matthews and Craig concerns the genetic variability encountered at individual collection sites.\textsuperscript{5} Variability ranged from 0.000 to 0.585 in our study, whereas it only ranged from 0.202 to 0.257 in the report by Matthews and Craig.\textsuperscript{5} This 10-fold disparity could arise from the types of genetic markers used. Variation in isozyme loci may be constrained by purifying selection,
whereas nucleotide variation in the mitochondrial ND4 gene was probably largely neutral. Seven of the eight substitutions were synonymous and involved transitions (Figure 2).

However, substitution rates alone fail to explain why many collections exhibited no variation, whereas others were highly variable (Figure 3) and why there were geographic clusters of collections with high genetic diversity. The quadrant abbreviation (north/south refers to location relative to Interstate 90, east/west refers to location relative to the Mississippi River), and genetic diversity values are listed next to the site abbreviation.

It is also possible that the collection methods are responsible for the clustering effect. In some of our collections, eggs only occurred in 1–2 ovitraps. However, we did not record how many ovitraps at each site contained eggs. Instead, a simple observation was made that not all ovitraps contained samples and it is possible that one female laid all of the eggs. We believe that this is highly unlikely but cannot assess this possibility because eggs from the five ovitraps at a location were combined prior to rearing them to adults. For future studies, it would be wise to record this information so that an accurate comparison can be made between genetic diversity and the number of oviposition cans containing eggs.

The present study used different field collecting techniques from those of Matthews and Craig. We collected five ovitrap papers at each site, whereas Matthews and Craig removed first and second instar larvae from a minimum of five tree-holes at each collection site. In both studies, adults in which genetic analyses were completed were reared in the laboratory. We would need to conduct a formal ecologic or genetic analysis of eggs found in individual traps to assess the frequency of egg dumping or skip oviposition in Oc. triseriatus.

**Table 6** F_{ST} values from allele frequencies in Matthews and Craig (N_{m} ≃ 22 reproductive migrants/generation)*

<table>
<thead>
<tr>
<th>Locus</th>
<th>F_{ST} among collections</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODH</td>
<td>0.026</td>
</tr>
<tr>
<td>PGM</td>
<td>0.008</td>
</tr>
<tr>
<td>HBD</td>
<td>0.009</td>
</tr>
<tr>
<td>IDH-2</td>
<td>0.008</td>
</tr>
<tr>
<td>MDH-2</td>
<td>0.008</td>
</tr>
<tr>
<td>ME</td>
<td>0.013</td>
</tr>
<tr>
<td>GPD</td>
<td>0.003</td>
</tr>
<tr>
<td>PGI</td>
<td>0.013</td>
</tr>
<tr>
<td>SOD-5</td>
<td>0.001</td>
</tr>
<tr>
<td>IDH-1</td>
<td>0.015</td>
</tr>
<tr>
<td>FUM</td>
<td>0.004</td>
</tr>
<tr>
<td>AK</td>
<td>0.004</td>
</tr>
<tr>
<td>HK-4</td>
<td>0.000</td>
</tr>
<tr>
<td>6-PGD</td>
<td>0.003</td>
</tr>
<tr>
<td>Average</td>
<td>0.011</td>
</tr>
</tbody>
</table>

* ODH = octanol dehydrogenase; PGM = phosphoglycerate mutase; HBD = hydroxybutyric dehydrogenase; IDH = isocitrate dehydrogenase; MDH = malate dehydrogenase; ME = malic enzyme; GPD = glyceral-3-phosphate dehydrogenase; PGI = phosphoglucoisomerase; SOD = superoxide dismutase; IDH = isocitrate dehydrogenase; FUM = fumarase; AK = adenylic kinase; HK = hexokinase; 6-PGD = 6-phosphogluconate dehydrogenase.
tus. If egg dumping (semelparity) is common in *Oc. triseriatus*, this would explain why a large proportion (approximately 13%) of variation arose among collections (Tables 3–5) within a quadrant or region even though mosquitoes across the study area were panmictic.

This survey of variation in mtDNA was completed rapidly because we switched from an SSCP to a SNP format for haplotype detection. The haplotypes were initially determined using SSCP, which identified 4 unique ND4 haplotypes among 564 initial mosquitoes. A subset of 27 of the 36 collection sites were screened for polymorphisms using SSCP analysis to obtain a broad overview of diversity in the study area. It is possible that SSCP analysis of all samples would have revealed additional low frequency haplotypes. However, addition of these haplotypes would not have greatly affected the diversity analysis. After sequence analysis of all four SSCP haplotypes, only eight polymorphic sites were detected. Seven of the eight SNP sites were in linkage disequilibrium, suggesting two maternal lineages of *Oc. triseriatus* in the study area. This disequilibrium also allowed us to detect the four ND4 haplotypes by simply identifying the base present at two SNP sites (bases 21 and 234). Use of HOLA allowed us to more rapidly determine genotypes in all 1,697 mosquitoes.16

These results suggest a general model of gene flow in *Oc. triseriatus* throughout our study area in the upper Midwest. The species exists as a single, panmictic adult population. However, within this panmictic population, local genetic drift arises, caused by one or a few females ovipositing in larval breeding containers (e.g., tires, cans, and treeholes). From generation to generation, there is sufficient gene flow to mix families arising from individual breeding sites and eliminate founders effects due to drift.

Received January 5, 2005. Accepted for publication April 18, 2005.

Acknowledgments: We thank William A. Tholfne and the rest of the people involved in La Crosse County Health Department’s vector control section for collecting the samples used in this study, as well as Cynthia Meredith for her work rearing these collections to adulthood. We would also like to thank Saul Lozano-Fuentes for his help using the ArcGIS program.

Financial support: This work was supported by National Institutes of Health grant AI-32543. Eric T. Beck was supported by Centers for Disease Control and Prevention Fellowship Training Program grant T01-CC82307.

Authors’ addresses: Eric T. Beck, Christopher F. Bosio, Carol D. Blair, Barry J. Beaty, and William C. Black IV, Arthropod-Borne and Infectious Diseases Laboratory, Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO 80523; E-mails: etbeck@lamar.colostate.edu, cbosio@colostate.edu, eblair@colostate.edu, bbeaty@colostate.edu, and william.black@colostate.edu. David A. Geske, Division of Vector Control, La Crosse County Health Department, La Crosse, WI 54601; E-mail: geske.dave@co.la-crosse.wi.us.

REFERENCES