

Insights Into Genome Differentiation: Pheromone-Binding Protein Variation and Population History in the European Corn Borer (*Ostrinia nubilalis*)

Christopher S. Willett and Richard G. Harrison

Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York 14853

Manuscript received April 23, 1999

Accepted for publication August 10, 1999

ABSTRACT

Examination of sequence variation at nuclear loci can give insights into population history and gene flow that cannot be derived from other commonly used molecular markers, such as allozymes. Here, we report on sequence variation at a single nuclear locus, the pheromone-binding protein (PBP) locus, in the European corn borer (*Ostrinia nubilalis*). The European corn borer has been divided into three races in New York State on the basis of differences in pheromone communication and life history. Previous allozyme data have suggested that there is a small but significant amount of genetic differentiation between these races. The PBP does not appear to be involved in the pheromone differences between these races. Examination of variation at the PBP locus in the three races reveals no fixed differences between races despite high levels of polymorphism. There also appears to have been considerable recombination in the history of the pheromone-binding protein alleles. Observation of both recombination between alleles and lack of significant nucleotide or insertion/deletion divergence between races leads us to suggest that these populations are either recently diverged or have continued to exchange genetic material subsequent to divergence in pheromone communication and life history.

ALLOZYMES and, more recently, microsatellites have been used extensively for examining population structure, and they represent genome-wide samples of variation (Lewontin and Hubby 1966; Ward *et al.* 1992; McDonald and Potts 1997). In contrast, the examination of nucleotide sequence variation at a single locus yields insights on the basis of a restricted set of closely linked (in some cases absolutely linked) polymorphisms. Nuclear genes can undergo intragenic recombination, which makes their gene genealogies complicated because different parts of a gene from the same individual will have different histories. However, this recombination and the multiple histories that each allele embodies makes calculation of effective population size and analyses of differentiation more robust (Hudson 1990; Pluzhnikov and Donnelly 1996).

Variation and differentiation at nuclear loci can be influenced by a number of different factors. A change in population size influences variation over the entire genome. Selection acts upon individual loci but influences variation at linked loci as well. The extent of this influence (and consequently the amount of polymorphism) depends upon regional levels of recombination (Kaplan *et al.* 1989; Begun and Aquadro 1992; Charlesworth *et al.* 1993). The different patterns of variation shown by nuclear genes from the same genome

can give insights into the complex history of genomes and populations.

Here, we examine patterns of variation for a single nuclear gene [a pheromone-binding protein (PBP)] in the European corn borer (ECB; *Ostrinia nubilalis*) to gain insights into the population structure of this important insect. The ECB exhibits variation in both life history and pheromone communication system, characteristics that could contribute to reproductive isolation between populations. Observed divergence between populations might then represent an early stage in the process of speciation.

The ECB is native to Europe, Northern Africa, and Western Asia (Mutuura and Monroe 1970) and was introduced to North America in the early part of the 20th century. The ECB is believed to have been introduced in shipments of broom corn from Austria, Hungary, and Italy (Smith 1920), and its spread across North America has been well documented (Caffrey and Worthley 1927; Roelofs *et al.* 1985; Hudson and LeRoux 1986). Based on the geography of early infestations and on the characteristics of the introduced moths, the ECB appears to have been introduced multiple times. A bivoltine (two generations per year) population was described first in Massachusetts, and a univoltine (single-generation) population in New York near Lake Erie. An additional area of early infestation was centered around the cities of Amsterdam and Albany, NY. The bivoltine and univoltine populations differed not only in life cycle, but also in host range. The ECB has now spread to all corn-producing areas of North America

Corresponding author: Christopher S. Willett, Marine Biology Research Division 0202, Scripps Institute of Oceanography, University of California, San Diego, CA 92093. E-mail: cwillett@ucsd.edu

east of the Rocky Mountains, resulting in considerable economic losses for farmers.

In both Europe and North America, the ECB has been shown to be variable in its pheromone communication system (Klun *et al.* 1975; Anglade *et al.* 1984). E and Z populations of the ECB have been distinguished in which the females produce and the males respond to differing ratios of E and Z isomers of 11-tetradecenyl acetate (11-14:OAc). In Z strain borers, the female moths produce and males respond to a 97:3 ratio of Z11-14:OAc to E11-14:OAc, whereas in the E strain borers, the opposite blend of isomers is used (1:99 ratio of Z11-14:OAc to E11-14:OAc; Roelofs *et al.* 1985). The Z strains are the more widespread borers in both Europe and North America and include both univoltine and bivoltine populations (Klun *et al.* 1975; Anglade *et al.* 1984). In Europe, E borers appear to be the predominant strain in parts of Italy and Switzerland; E borers are rare or absent elsewhere.

In New York State, three different races of corn borer have been described: bivoltine Z (BZ), univoltine Z (UZ), and bivoltine E (BE; Roelofs *et al.* 1985). In this article, we use the term "race" to refer to these New York populations and "strain" to refer to a population having either the E or Z pheromone communication system. In certain areas of New York, all three races can be collected, and hybridization between these races occurs. Hybridization has been documented between pheromone strains by the observation of females in the field producing a characteristic hybrid ratio pheromone (Glover *et al.* 1991). Hybridization is enabled by both a small percentage of males that will fly to females of the other pheromone strain (Linn *et al.* 1997) and some overlap in dates when adults with different life histories are flying (Eckenrode *et al.* 1983).

Several allozyme studies of genetic variation within and between the E and Z strains of the ECB have shown that there is a very small but significant amount of differentiation across the genome (Harrison and Vawter 1977; Cianchi *et al.* 1980; Glover *et al.* 1990). Harrison and Vawter (1977) found that Nei's genetic identity (I) for the two pheromone strains was 0.977. Despite the apparently low level of differentiation over most of the genome inferred from allozyme studies, differences in the pheromone system and voltinism appear to have a genetic basis. Patterns of inheritance for female pheromone production, male pheromone reception, and male pheromone response are consistent with each of these traits being determined by a single-gene locus (Klun and Maini 1979; Roelofs *et al.* 1987; Löfstedt *et al.* 1989). The change in male pheromone response behavior is caused by differences in a sex-linked locus or set of linked loci. *Tpi* and *6-Pgd* are the only allozyme loci that display large differences in allele frequencies between the pheromone strains and are sex linked (Glover *et al.* 1990). The differences in voltinism between the ECB populations are caused by a combination

of environmental cues and genetic differences (Cafrey and Worthley 1927; Sparks *et al.* 1966; Glover *et al.* 1992).

In this study, we examine patterns of variation in the gene encoding PBP, a protein involved in pheromone reception that is thought to transport pheromone molecules to neuronal receptors within the antennae (Prestwich *et al.* 1995; Vogt 1995). Willett and Harrison (1999) characterized a member of the PBP gene family from the ECB and its close relative, the Asian corn borer (*O. furnacalis*), and showed that it is a single-copy gene in these species. They also argued that this protein is apparently not involved in the discrimination of pheromone differences between the ECB pheromone races because it does not show any fixed amino acid differences between these races. However, their analyses revealed substantial variation at third-codon positions and in noncoding regions. On the basis of more extensive analyses of variation in the PBP in the ECB, we now show that there are no fixed nucleotide differences at the PBP locus between the pheromone strains despite the high level of genetic variation segregating at this locus. We also infer amounts of recombination from the pattern of variation in the alleles we have sampled, and we use these data to propose that the two pheromone strains must either have a recent origin or a more distant origin followed by continuing gene flow between the strains.

MATERIALS AND METHODS

Samples and PBP sequencing: The ECB moths we used in this study were obtained from cultures maintained by Wendell Roelofs's laboratory (Geneva, NY). The cultures derived from collections made in the field at different times, and moths were also sampled from these cultures at different times (Table 1). Colonies were started with 400 or more moths and maintained with ~150 moths. Our sample included three bivoltine E moths (BE1-3), four univoltine Z moths (UZ1-4), and a single bivoltine Z moth (BZ1). Genomic DNA was prepared as described in Willett and Harrison (1999), except for the two individuals collected in 1987, from which DNA was isolated as described in Harrison *et al.* (1987) and then stored at -20°.

We characterized the genomic sequences of the PBP from these eight ECB. Willett and Harrison (1999) describe in detail the initial characterization of the PBP of the ECB. We used a PCR-based approach using degenerate primers designed to match conserved regions of other published PBPs to initially amplify the PBP. For this study, a fragment of 1.6 kb of DNA, including the PBP, was amplified by using the primers ECEP5 and ECPA (see Willett and Harrison 1999) and the following conditions: 35 cycles of 96° for 30 sec (an additional 30 sec during the first step), 50° for 30 sec, and 72° for 2 min.

Insertion/deletion (indel) heterozygotes complicated the sequencing of many of these PCR products. We were only able to sequence the individuals BE2, UZ3, and UZ4 directly from the full-length PCR products. With the primer ECB5 (3' end at position 217 in the Figure 1 sequence) we directly sequenced the 5' region of the PBP from all of the individuals. To obtain the remainder of the PBP sequence, we used two

TABLE 1
Sample collection localities and dates

Individual	Pheromone race	Collection location	Collection date	Field or culture	Date prepped
BE1,2	Bivoltine E	Geneva, NY	May 1989	Culture	1996
BE3	Bivoltine E	Geneva, NY	May 1996	Field	1996
UZ1,2,3	Univoltine Z	Bouckville, NY	April 1994	Culture	1996
UZ4	Univoltine Z	Paris, NY	1984	Culture	1987
BZ1	Bivoltine Z	Erie, PA	1987	Field	1987

Field or Culture indicates whether the moth was taken from field collections or from a culture.

different methods to separate the two haplotypes of an individual: the cloning of PCR products and an acrylamide gel separation of the haplotypes. The two alleles of an individual often differed in length, which allowed the alleles to be separated on 4% acrylamide large-format gels that were then silver stained (Silver Sequence kit by Promega, Madison, WI) to visualize bands. We were able to reamplify each PCR-amplified band separated in this manner by pipeting 1 μ l of water onto each band on the gel and using this extract as template in a PCR reamplification. It was possible to efficiently separate and reamplify bands that were <800 bp. Obtaining three overlapping fragments of this size required three additional primers: ECUT3 (5'-TCTAAGTATGATTCCAACCTCCATCG-3'), CB1 (5'-AGATGTAATGAAACAGATGACATC-3'), and CB2 (5'-AATTCTACGAGTACCAAATAGTA-3'). The primer pairs for the three fragments were CB1/ECBI, CB2/ECEP3, and ECT3/ECUT3 (positions of the 3' ends of these primers are indicated in Figure 1). All sequencing was performed using the Thermo Sequenase radiolabeled terminator cycle sequencing kit (Amersham Life Science, Cleveland, OH) with the aforementioned primers as sequencing primers.

The overlapping sequence between these fragments allowed the haplotype phase to be reconstructed if the individual was heterozygous for a polymorphic site within that region. When individuals were not heterozygous within these overlap regions, it was necessary to clone the original PCR product (made with primers ECEP5 and ECPA) and sequence a portion of each of the fragments to determine phase. Multiple clones had to be examined to obtain both alleles in an individual and to overcome the possible problems of recombination of alleles within the PCR reaction and Taq error (Bradley and Hillis 1997). Gel-purified PCR products were cloned using the pGEM-t vector system (Promega). We sequenced between four and eight clones from each individual, which allowed us to determine the correct phase of polymorphisms for each allele. For the individuals UZ1 and UZ2, the phase relationships for polymorphisms in the 5'-most 160 bp were not determined because we had only sequenced this end directly from the PCR products. Sequences of all haplotypes have been deposited in GenBank under accession nos. AF133631–AF133643.

Sequence analysis: The MegAlign program of the Lasergene package (DNASTAR, Madison, WI) was used to align the sequences. The CLUSTAL algorithm, with a gap penalty of 8 and a gap length penalty of 2, was used for the initial alignment, which was then adjusted by eye. The program DNAsp v2.52 (Rozas and Rozas 1997) was used to calculate average pairwise differences per site (π values). These values were calculated by excluding all regions of alignment gaps and missing data from the data set. The two nucleotides of a polymorphism were randomly assigned to the two alleles for each of the polymorphisms for which haplotype phase had not

been determined because phase will not affect the value of π . DNAsp v2.52 was also used to calculate the values of average genetic distances between populations (D_a in Nei 1987).

We used several different methods to look for evidence of recombination in our sample of alleles. The first method used phylogenies constructed for different regions of the locus. The data set we used included all nucleotide substitutions; each entire indel was coded as a single character (not one character per nucleotide gap). The branch and bound algorithm using parsimony as the criterion was used to construct trees with the program PAUP* (Swofford 1998). A PBP sequence of the same region from the Asian corn borer (*O. furnacalis*) was used to root each network (GenBank accession no. AF133629). We divided the locus into three adjacent regions, each of which had approximately the same information content (25–26 parsimony informative characters). We constructed trees for each of these three regions, and our inference of recombination from these trees is described in results. We also used the method of Hudson and Kaplan (1985) for detecting the minimum number of recombination events (R_m). We used the program DNAsp v2.52 to calculate the R_m for a sample of alleles, including the indel data recoded with each indel as a single-nucleotide polymorphism (DNAsp only works with base pair substitution data). DNAsp was used to calculate the recombination parameter C (Hudson 1987) for the nucleotide data set, while the program SITES (Hey and Wakeley 1997) was used on the same data set to estimate the analogous γ (both are estimates of $4Nc$).

RESULTS

PBP sequences: We sequenced just under 1.5 kb from eight different corn borers. The sequence includes 483 bp of coding region in three exons, two introns of 103 and 256 bp (size varies slightly depending upon size of indels in introns), 100 bp of 5' flanking sequence, and 515 bp of 3' flanking sequence (Figure 1). As found in PBPs from other species, the first 20 amino acids of the PBP appear to be a signal peptide that is cleaved from the mature protein. The mature ECB protein is 45% identical in amino acid sequence to a gypsy moth (*Lymantria dispar*) PBP (Merritt *et al.* 1998) and 70% identical to the budworm moth (*Heliothis virescens*) PBP (Krieger *et al.* 1993), which represent the most dissimilar and similar PBP sequences published to date. The noncoding regions sequenced in this study display no apparent sequence identity with the corresponding

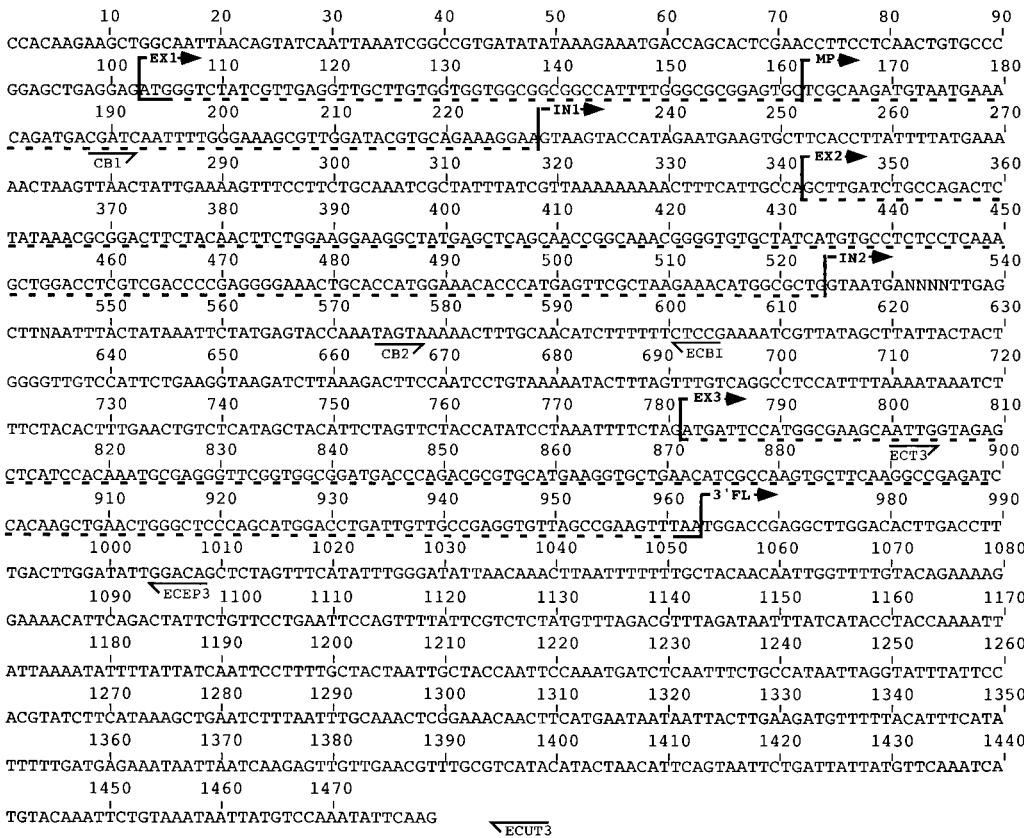


Figure 1.—Sequence of PBP locus. This is the BE1a haplotype for the region that we have sequenced in all individuals. Intron (IN) and exon (EX) boundaries are indicated on the figure. The predicted boundary between the signal peptide and the mature protein is marked with an MP. The start of the 3' flanking region is indicated with 3' FL. The positions of the 3' end of the primers are indicated with an arrow. These were the primers used to generate the three overlapping fragments for the acrylamide gel separation of alleles. The primer ECUT3 lies an additional 48 bp from the 3' end of this sequence.

noncoding regions in published sequences (data not shown).

Figure 2 depicts the variation in the PBP and surrounding region for 16 alleles from eight borers from the three different pheromone strains, BZ, BE, and UZ [complete sequence alignment in Willett (1999)]. We have also shown the sequence of a single Asian corn borer PBP sequence at the variable ECB sites. Both BE2 and UZ3 have a single heterozygous site each, while UZ4 is apparently a homozygote. The variation in the complete set of ECB alleles is composed of 94 nucleotide polymorphisms and 32 indels. All the indels occur in the noncoding regions, but within the coding regions, there are 11 synonymous and 3 nonsynonymous polymorphisms. The first two replacement sites are in the signal peptide region of the protein. Three nucleotides segregating at one position occur four times in this sample (once leading to two silent polymorphisms). At five points in the sequence, there are 2 differently sized indels. Indels ranged in size from 1 to 47 bp, with an average size of 5.7 bp (16 of the 32 were only 1 or 2 bp in size). When we use the Asian corn borer sequence as an outgroup to determine whether each indel is an insertion or a deletion, there have been an equal number of insertions and deletions.

Nucleotide diversity and population structure: The average number of pairwise differences per site (π) is 0.017 when calculated for the entire region (Table 2). We used 16 alleles for the calculation of levels of poly-

morphism, 2 for each individual. For the calculation of π , we included only 82 of the 94 segregating sites because the other 12 occur in regions with missing data or indels. When the sample is divided up into the three New York races, there is still a substantial amount of variation within each race. The BZ alleles are the most different from each other and the other alleles. The two BZ alleles contain 32 singletons (both single-nucleotide and single-indel polymorphisms unique to one of the BZ alleles). The BE race appears to be the least variable of the three races over the entire locus, and this reduced level of variation is especially pronounced at synonymous sites.

There are no fixed nucleotide or indel differences between the E and Z pheromone strains. Most of the variation is apportioned between alleles rather than between strains or races. One measure of the average pairwise distance between populations is D_a (Nei 1987), which is calculated by subtracting the average amount of pairwise distance between alleles within each population $[(D_x + D_y)/2]$ from the average pairwise distance between the populations (D_{xy}). When we compare the UZ haplotypes to the BE haplotypes, $D_a = 0.0032$.

Haplotype analysis and recombination: From the above results, it is obvious that there is a great deal of variation in the sample of ECB alleles, but little of this variation is partitioned among pheromone strains. Absence of race-specific variation at the PBP locus could be caused by past hybridization, either in Europe or

TABLE 2
Nucleotide diversity in the European corn borer PBP

Haplotypes	Total			Synonymous			Nonsynonymous			Noncoding + syn.		
	π_{total}	<i>N</i>	<i>S</i>	π_{syn}	<i>N</i>	<i>S</i>	$\pi_{non-syn}$	<i>N</i>	<i>S</i>	π_{nc+syn}	<i>N</i>	<i>S</i>
All 16	0.0173	1356	82	0.0355	104.5	11	0.00353	375.5	3	0.0227	976.5	79
6 BE	0.0134	1438	50	0.0177	105.6	5	0.00263	380.4	3	0.0173	1058	47
8 UZ	0.0161	1427	51	0.0356	104.5	8	0.0038	375.5	3	0.0206	1048	48
2 BZ	0.0259	1389	36	0.0474	105.4	5	0.00526	380.6	2	0.0337	1008	34
All 10 Z	0.0178	1359	72	0.0372	104.5	11	0.00391	375.6	3	0.0231	979.5	69

N, the total number of nucleotides from which π and *S* values are calculated; *S*, the number of segregating sites in the sample; π , the average number of pairwise differences per nucleotide between haplotypes.

subsequent to the introduction of the ECB to North America. If the hybridization between the races is a fairly recent phenomenon (if they are genetically distinct at the PBP locus in the ancestral European populations), then we may expect to find divergent haplotypes within each of the populations given the limited time that recombination and gene conversion have had to break up these haplotypes. An examination of the alleles in Figure 2 reveals regions that are distinct between haplotypes, but these regions are limited in extent. We have

examined haplotype phylogenies for the entire gene region and for three adjacent regions along the gene (Figure 3). In the "regional" haplotype trees, we see that three groups of haplotypes found in the "All Data" tree consistently appear as groups in Figure 3 (labeled A–C). Otherwise, the topologies of the regional trees are very different, presumably reflecting a history of recombination and gene conversion between alleles in the population. The allele BE3b in Figure 3 (marked with an asterisk) could have been generated by a single

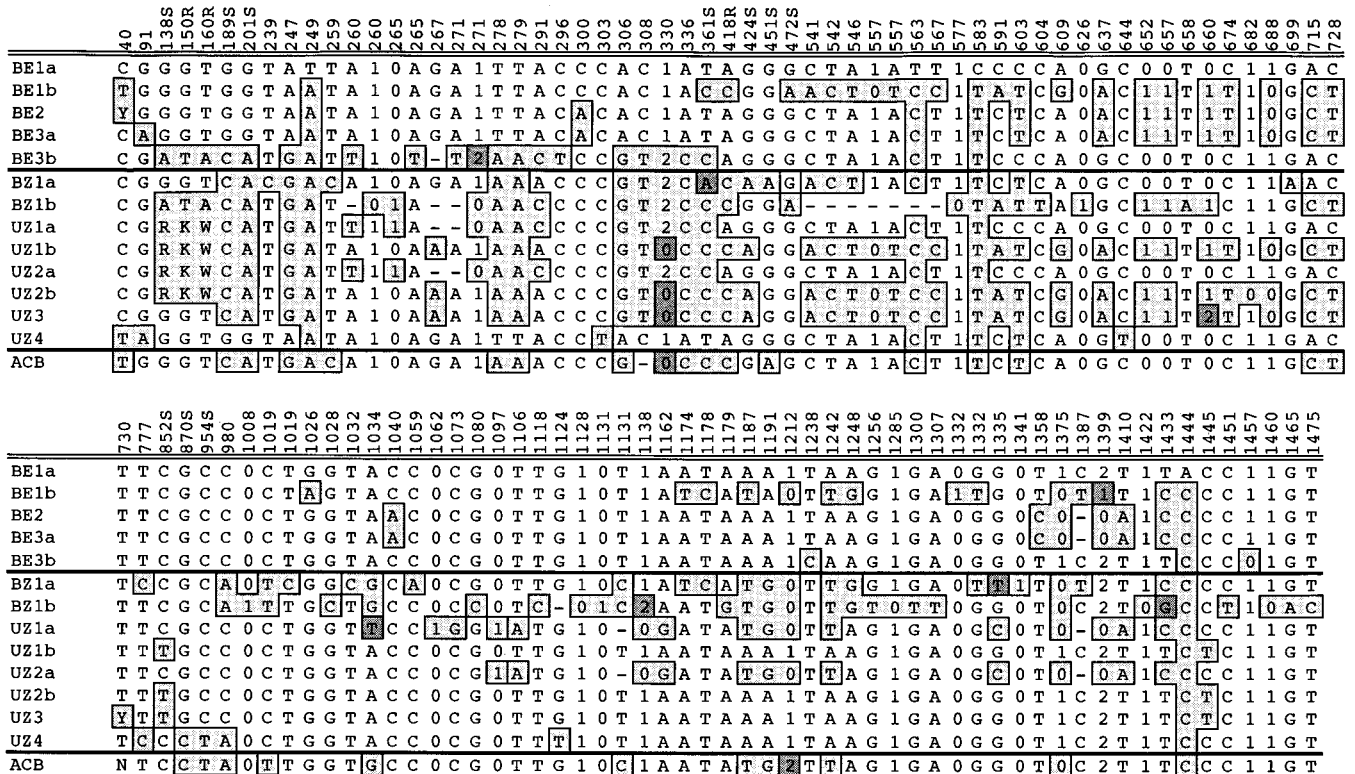


Figure 2.—Alignment of the variable sites in the ECB sample. Numbers along the top correspond to the position of the variable site in the BE1a haplotype presented in Figure 1. *S* indicates a silent polymorphism, and *R* indicates a replacement polymorphism. 138S, 150R, and 160R all fall into the predicted signal peptide region. Indels are coded in the figure with the numbers 0–2: 0, absence of DNA; 1, the presence of DNA; 2, another overlapping indel of smaller size. Ambiguous sites are denoted by the following: R-A/G, W-A/T, K-G/T, and Y-C/T. The line divides the E and Z strain haplotypes. ACB is a PBP from the Asian corn borer (*O. furnacalis*), and we have only included the sites that were variable in the ECB and not those that differed only between the ACB and the ECB sequences.

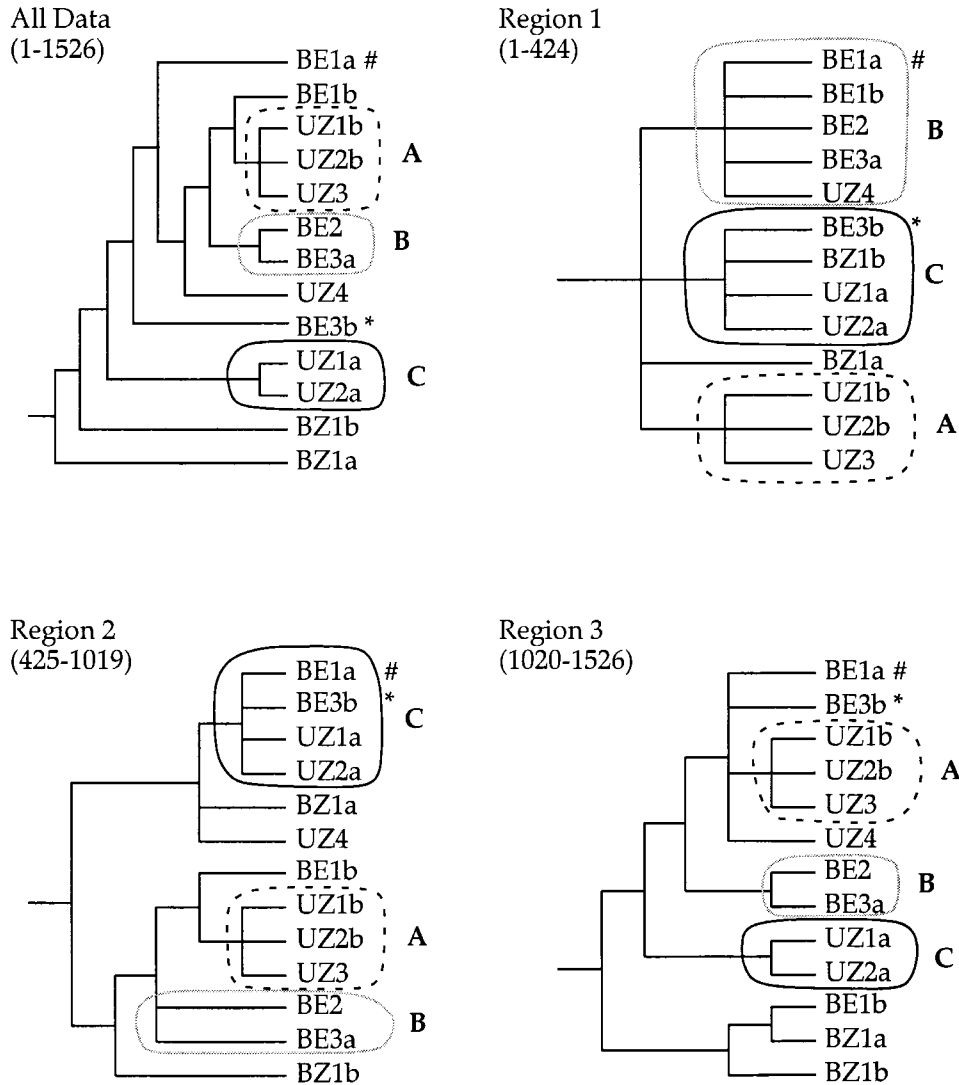


Figure 3.—Consensus haplotype trees for the PBP locus divided into three regions. The divisions were made to include 25 parsimony informative characters in each region. The three regions are bounded by the BE1a haplotype positions (Figure 1) in parentheses. Using all the data gives rise to a single most parsimonious tree of 182 steps. The region 1 tree is a consensus of 87 trees of 42 steps. Region 2 is a consensus of 4 trees of 47 steps. Region 3 gives a single tree of 60 steps. Three groups of haplotypes that remain together in each of the regions are denoted by A–C. Although group B does not appear to be monophyletic in region 2, the two alleles are actually identical in this region.

recombination event between a group C-like region 1 and 2 with a group A-related region 3. In contrast, allele BE1a (denoted with a pound sign) appears to have had a more complex history because in each region, it appears to be related to a different group of haplotypes.

From the haplotype phylogenies, it appears that there has been some recombination in the history of the ECB PBP alleles. To quantify the amount of recombination, we estimated the minimum number of recombination events (R_m) as proposed by Hudson and Kaplan (1985). The four-gamete test is the basis of this approach, and it postulates a recombination event when all four possible gamete types from a recombination event are represented in the sample. For the sample of ECB alleles, this method gives an R_m of 8. This R_m is considered a minimum estimate because the method will miss recombination between alleles having the same markers, and because not all four gametic types from recombination events are likely to be recovered in a finite sample. However, recurrent mutation can produce all four ga-

mete types, and these can be counted incorrectly as recombination events.

Other measures of recombination can be obtained from estimates of the recombination parameter $4Nc$ with either C (Hudson 1987) or γ (Hey and Wakeley 1997). These estimates are $C = 0.031$ and $\gamma = 0.017$, both expressed per base pair. γ is likely to be the more accurate estimate with smaller sample sizes (Hey and Wakeley 1997). By dividing the recombination parameter by an estimate of $4N\mu$, we can estimate the number of recombination events per mutation event. The value of π provides an estimate of $4N\mu$, and if we divide our estimate of γ by π for the total gene, we get $0.017/0.017 = 1$ or about 1 recombination event per mutation event. The minimum number of mutation events in the sample is equal to the number of segregating sites (82); if we multiply the number of mutation events by 1 recombination event per mutation event, we have an estimate of 82 recombination events in the history of this sample.

DISCUSSION

The gene loci that are instrumental in conferring specificity in pheromone communication systems should show fixed amino acid differences between strains or species. As discussed here and in Willett and Harrison (1999), there are no fixed amino acid differences between the E and Z strains of the ECB, which implies that PBP is not directly responsible for differences between the strains in pheromone response. Fixed differences in noncoding sequence could be caused by limited gene exchange or could be changes in regulatory regions. However, we show that there are no fixed nucleotide or indel differences between the pheromone strains, even when we examine introns and flanking sequences.

Allozyme studies of the ECB have suggested that replacement differences between the pheromone strains are likely only at loci involved in racial differences or closely linked to such loci (Harrison and Vawter 1977; Cianchi *et al.* 1980; Glover *et al.* 1990). Only the enzymes TPI and 6-PGD display clear differences in allele frequency between the E and Z strains. The loci encoding these enzymes are sex linked and may be closely linked to the male behavioral response trait and genetic differences in voltinism (Glover *et al.* 1990, 1992). Our PBP data, together with published allozyme studies, suggest that fixed nucleotide differences between pheromone strains may only be present in the genes responsible for pheromone strain phenotypic differences or in genes closely linked to them.

The amount of variation at synonymous and noncoding sites in the PBP locus is higher than the average values from *Drosophila melanogaster* ($\pi_{\text{syn}} = 0.014$, $\pi_{\text{nc}} = 0.011$) and about the same level as those from *D. simulans* ($\pi_{\text{syn}} = 0.030$, $\pi_{\text{nc}} = 0.019$; Moriyama and Powell 1996). In *Drosophila*, it has been shown that the level of recombination correlates with the level of polymorphism (Begun and Aquadro 1992). The π values from this study are on the order of π values from regions of high recombination in *D. melanogaster* (Moriyama and Powell 1996). Assuming that mutation rates are similar in *Drosophila* and the ECB, the level of recombination in this gene would be roughly equivalent to the level of recombination in some *Drosophila* genes: in several *Drosophila* genes, estimates of the number of recombination events per mutation event range between 2.2 and 4.9 (Hey and Wakeley 1997), and we have estimated one recombination event per mutation event for this sample of alleles.

What influence does the variation in pheromone communication and life history have on levels of polymorphism? If we divide the sample into either the two pheromone strains or the three races, the π values are still nearly as high within each (Table 2). The haplotype trees presented in Figure 3 show that PBP variation provides little evidence of genetic structure in ECB pop-

ulations in New York. For trees based on all the data or on partitions of the data, the E and Z strain borers are scattered around the trees. The three groups of alleles that are composed of either E or Z strain borers in the All Data tree are often related to borers from the opposite strain in the regional trees. Could these groups of nearly identical alleles reflect a recent common ancestry because they have been sampled from relatively small laboratory populations? In fact, only the two alleles from UZ4 are identical to one another, while other groups of closely related alleles differ from each other at one to three sites. It is unlikely that mutations would have occurred in identical alleles in the period of time that these moths have been kept in captivity (<20 generations). Most of the alleles must therefore represent independent samples from the natural populations from which the lab cultures were derived.

The haplotype trees, estimates of R_m , and estimates of recombination from the γ value suggest that there has been considerable recombination and gene conversion in the history of the alleles in this study. The level of recombination estimated by both C and γ are on the order of those found in population samples of *Drosophila* species that are thought to have few barriers to free gene exchange. These recombination events do not appear to be limited to alleles from the same strain, and they reinforce the notion that these strains can and do hybridize in the field. The amount of recombination that we have observed may be too great to have occurred over the short period of time that the ECB has been present in North America. This period is ~ 180 generations for the bivoltine races and 90 generations for the univoltine race, if an introduction about 1910 is assumed. Given the relatively short period of time for recombination to have occurred and the small size of this region as a target for recombination [recombination rates in *Drosophila* have been estimated to be ~ 3 cM/Mb, which translates to $\sim 10^{-8}$ for adjacent nucleotides (Nachman and Churchill 1996)], it would seem that the estimated number of 82 recombination events for the ECB sample of alleles is too large for all the recombination to have occurred since introduction. The E and Z strains most likely did not arrive in North America with fixed haplotype differences at the PBP locus.

How can we synthesize the observations for PBPs of high levels of variability, little differentiation between strains or races, as well as extensive recombination between E and Z alleles? The high levels of polymorphism imply that the ECB has maintained a fairly large effective population size for some time, despite a likely bottleneck in population size when introduced to North America. The introduction is thought to have included individuals from a number of different source locations to several different locations in northeastern North America. Multiple introductions would have helped to maintain variation, especially if these populations ex-

panded rapidly to the large present population size. Despite the variation between populations in pheromone systems and life history, it would appear that there is little evidence from our sample of PBP alleles for anything other than a randomly mating population. This lack of structure could be explained in two ways. One explanation is a recent origin of the pheromone differences. The origin would have been recent enough that there has not been sufficient time for loci to drift to fixation for neutral variants. The extreme of this scenario is that there is complete isolation between strains, and that all segregating polymorphism is ancestral polymorphism that has not yet sorted between the strains. The alternative explanation would be an older origin of the pheromone strains with continuing hybridization between the strains. This hybridization would homogenize the genome of the ECB across strains for loci not closely linked to the fixed pheromone strain differences. There is independent evidence for current hybridization between the pheromone strains, and this favors the second explanation for the absence of substantial differentiation; however, it is possible to have some combination of the two explanations, *e.g.*, a fairly recent origin and continuing hybridization between the strains.

The origin of the E and Z pheromone strains and their continued maintenance despite hybridization remains a mystery. If the origin is not very recent, then this pheromone difference would have been maintained despite continued gene flow between the strains. The hybrids resulting from interstrain crosses are viable, but perhaps they may have lower fitness in relatively pure populations of either strain because of the difficulty in finding mates. This pressure would cause selection to maintain the dominant pheromone type in a population. Another possibility for the maintenance or origin of the pheromone strains is that some of these pheromonal loci are linked to other selected loci (*e.g.*, voltinism and pheromone behavioral response appear to be linked) and are maintained because of hitchhiking. In either case, selection would be maintaining differences at these pheromonal loci and linked loci between strains, but other parts of the genome would be homogenized given gene flow between the strains. The pheromone-binding protein of the ECB appears to fall into the category of genes homogenized by gene flow. Direct analyses of loci involved in pheromonal differences or of loci closely linked to these may help determine the answers to questions surrounding pheromone strain origin and maintenance.

We thank Wendell Roelofs and Kathy Poole for providing European corn borers, and Chenhua Zhao for the Asian corn borers. Steve Bogdanowicz and Scott Stanley provided useful technical advice. We also thank Charles Aquadro, Wendell Roelofs, and Andrew Clark for helpful comments on this manuscript. This work was supported by National Science Foundation grant IBN-9700704 to R.G.H and C.S.W. C.S.W. was supported by a National Science Foundation predoctoral

fellowship and a National Institutes of Health traineeship from the Field of Genetics and Development at Cornell University.

LITERATURE CITED

- Anglade, P., J. Stockel, W. Awadallah, F. Bača, H. Berger *et al.*, 1984 Intraspecific sex-pheromone variability in the European corn borer, *Ostrinia nubilalis* Hbn. (Lepidoptera, Pyralidae). *Agro-nomie* **4**: 183–187.
- Begun, D. J., and C. F. Aquadro, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *Drosophila melanogaster*. *Nature* **356**: 519–520.
- Bradley, R. D., and D. M. Hillis, 1997 Recombinant DNA sequences generated by PCR amplification. *Mol. Biol. Evol.* **14**: 592–593.
- Caffrey, D. J., and L. H. Worthley, 1927 A progress report on the investigations of the European corn borer. *USDA Bull.* **1476**.
- Charlesworth, B., M. T. Morgan and D. Charlesworth, 1993 The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**: 1289–1303.
- Cianchi, R., S. Maini and L. Bullini, 1980 Genetic distance between pheromone strains of the European corn borer, *Ostrinia nubilalis*: different contribution of variable substrate, regulatory, and non-regulatory enzymes. *Heredity* **45**: 383–388.
- Eckenrode, C. J., P. S. Robbins and J. T. Andaloro, 1983 Variations in flight patterns of European corn borer (Lepidoptera: Pyralidae) in New York. *Environ. Entomol.* **12**: 393–396.
- Glover, T. J., M. Campbell, P. Robbins and W. L. Roelofs, 1990 Sex-linked control of sex pheromone behavioral responses in European corn borer moths *Ostrinia nubilalis* confirmed with TPI marker gene. *Arch. Insect Biochem. Physiol.* **15**: 67–78.
- Glover, T. J., J. J. Knodel, P. S. Robbins, C. J. Eckenrode and W. L. Roelofs, 1991 Gene flow among three races of European corn borers (Lepidoptera:Pyralidae) in New York State. *Environ. Entomol.* **20**: 1356–1362.
- Glover, T. J., P. S. Robbins, C. J. Eckenrode and W. L. Roelofs, 1992 Genetic control of voltinism characteristics in European corn borer races assessed with a marker gene. *Arch. Insect Biochem. Physiol.* **20**: 107–117.
- Harrison, R. G., and A. T. Vawter, 1977 Allozyme differentiation between pheromone strains of the European corn borer, *Ostrinia nubilalis*. *Ann. Entomol. Soc. Am.* **70**: 717–720.
- Harrison, R. G., D. M. Rand and W. C. Wheeler, 1987 Mitochondrial DNA variation in field crickets across a narrow hybrid zone. *Mol. Biol. Evol.* **4**: 144–158.
- Hey, J., and J. Wakeley, 1997 A coalescent estimator of the population recombination rate. *Genetics* **145**: 833–846.
- Hudon, M., and E. J. Leroux, 1986 Biology and population dynamics of the European corn borer (*Ostrinia nubilalis*) with special reference to sweet corn in Quebec, Canada. I. Systematics, morphology, geographical distribution, host range, economic importance. *Phytoprotection* **67**: 39–54.
- Hudson, R. R., 1987 Estimating the recombination parameter of a finite population model without selection. *Genet. Res.* **50**: 245–250.
- Hudson, R. R., 1990 Gene genealogies and the coalescent process, pp. 1–44 in *Oxford Surveys in Evolutionary Biology*, edited by D. Futuyma and J. Antonovics. Oxford University Press, Oxford.
- Hudson, R. R., and N. L. Kaplan, 1985 Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**: 147–164.
- Kaplan, N. L., R. R. Hudson and C. H. Langley, 1989 The hitchhiking effect revisited. *Genetics* **123**: 887–899.
- Klun, J. A., P. L. Anglade, F. Bača, O. L. Chapman, H. C. Chiang *et al.*, 1975 Insect sex pheromones: intraspecific pheromonal variability of *Ostrinia nubilalis* in North America and Europe. *Environ. Entomol.* **4**: 891–894.
- Klun, J. A., and S. Maini, 1979 Genetic basis of an insect chemical communication system: the European corn borer. *Environ. Entomol.* **8**: 423–426.
- Krieger, J., H. Ganble, K. Raming and H. Breer, 1993 Odorant binding proteins of *Heliothis virescens*. *Insect Biochem. Mol. Biol.* **23**: 449–456.
- Lewontin, R. C., and J. L. Hubby, 1966 A molecular approach

- to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics* **54**: 595–605.
- Linn, C. E., Jr., M. S. Young, M. Gendle, T. J. Glover and W. L. Roelofs, 1997 Sex pheromone blend discrimination in two races and hybrids of the European corn borer moth, *Ostrinia nubilalis*. *Phys. Entomol.* **22**: 212–223.
- Löfstedt, C., B. S. Hansson, W. L. Roelofs and B. O. Bengtsson, 1989 No linkage between genes controlling female pheromone production and male pheromone response in the European corn borer *Ostrinia nubilalis* Huebner (Lepidoptera: Pyralidae). *Genetics* **123**: 553–556.
- Merritt, T. J., S. LaForest, G. D. Prestwich, J. M. Quattro and R. G. Vogt, 1998 Patterns of gene duplication in Lepidopteran pheromone binding proteins. *J. Mol. Evol.* **46**: 272–276.
- McDonald, D. B., and W. K. Potts, 1997 DNA microsatellites as genetic markers at several scales, pp. 29–49 in *Avian Molecular Evolution and Systematics*, edited by D. P. Mindell. Academic Press, San Diego.
- Moriyama, E. N., and J. R. Powell, 1996 Intraspecific nuclear DNA variation in *Drosophila*. *Mol. Biol. Evol.* **13**: 261–277.
- Mutuura, A., and E. Munroe, 1970 Taxonomy and distribution of the European corn borer and allied species: genus *Ostrinia* (Lepidoptera:Pyralidae), pp. 1–112 in *Memoirs of the Entomological Society of Canada*, edited by D. P. Pielou. Entomological Society of Canada, Ottawa.
- Nachman, M. W., and G. A. Churchill, 1996 Heterogeneity in rates of recombination across the mouse genome. *Genetics*. **142**: 537–548.
- Nei, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Pluzhnikov, A., and P. Donnelly, 1996 Optimal sequencing strategies for surveying molecular genetic diversity. *Genetics* **144**: 1247–1262.
- Prestwich, G. D., G. Du and S. LaForest, 1995 How is pheromone specificity encoded in proteins? *Chem. Senses* **20**: 461–469.
- Roelofs, W. L., J. W. Du, X. H. Tang, P. S. Robbins and C. J. Eckenrode, 1985 Three European corn borer populations in New York based on sex pheromones and voltinism. *J. Chem. Ecol.* **11**: 829–836.
- Roelofs, W. L., T. Glover, X. H. Tang, I. Sreng, P. Robbins *et al.*, 1987 Sex pheromone production and perception in European corn borer moths is determined by both autosomal and sex-linked genes. *Proc. Natl. Acad. Sci. USA* **84**: 7585–7589.
- Rozas, J., and R. Rozas, 1997 DNAsp version 2.0: a novel software package for extensive molecular population genetic analysis. *Comp. Appl. Biosci.* **13**: 307–311.
- Smith, H. E., 1920 Broom corn, the probable host in which *Pyrausta nubilalis* Hubn. reached America. *J. Econ. Entomol.* **13**: 425–430.
- Sparks, A. N., T. A. Brindley and N. D. Penny, 1966 Laboratory and field studies of F1 progenies from reciprocal matings of biotypes of the European corn borer. *J. Econ. Entomol.* **59**: 915–921.
- Swofford, D. L., 1998 PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, MA.
- Vogt, R. G., 1995 Molecular genetics of moth olfaction: a model for cellular identity and temporal assembly of the nervous system, pp. 341–367 in *Molecular Model Systems in the Lepidoptera*, edited by M. R. Goldsmith and A. S. Wilkens. Cambridge University Press, New York.
- Ward, R. D., D. O. F. Skibinski and M. Woodwark, 1992 Protein heterozygosity, protein structure, and taxonomic differentiation. *Evol. Biol.* **26**: 73–159.
- Willett, C. S., 1999 The evolution of pheromone binding proteins: insights into protein evolution and population history. Ph.D. Dissertation, Cornell University, Ithaca, NY.
- Willett, C. S., and R. G. Harrison, 1999 Pheromone binding proteins in the European and Asian corn borers: no protein change associated with pheromone differences. *Insect Biochem. Mol. Biol.* **29**: 277–284.

Communicating editor: A. G. Clark