Genetica (2006) 127:143–161
DOI 10.1007/s10709-005-2673-z

Molecular genetic variation of boll weevil populations in North America estimated with microsatellites: Implications for patterns of dispersal

Kyung Seok Kim† & Thomas W. Sappington*
*USDA-ARS, Corn Insects & Crop Genetics Research Unit, Genetics Laboratory, Iowa State University, Ames, IA, 50011, USA; (*Author for correspondence Fax: +515-294-2265; E-mail: tsapping@iastate.edu)*

Received 13 May 2005; Accepted 1 September 2005

Key words: boll weevil, dispersal, eradication, gene flow, genetic variation, invasive species, microsatellites, migration, molecular markers, population assignment

Abstract

The boll weevil (*Anthonomus grandis* Boheman) is an insect pest of cotton that underwent a well-documented range expansion across the southeastern U.S. from Mexico beginning about 110 years ago. Eleven microsatellite loci were surveyed to infer the magnitude and pattern of genetic differentiation among boll weevil populations from 18 locations across eight U.S. states and northeast Mexico. Estimates of genetic diversity (allelic diversity and heterozygosity) were greater in Southern than Northern populations, and were greater in the west than the east among Northern populations. Boll weevil populations were genetically structured as a whole across the geographic range sampled, with a global F_{ST} of 0.241. South-central populations exhibit classic isolation by distance, but evidence suggests that populations within the Eastern and Western regions have not yet reached genetic equilibrium. Gene flow appears to be relatively high among populations within the Eastern region. Population assignment data and estimates of gene flow indicate that migration between locations separated by <300 km is frequent. The database of microsatellite genotypes generated in this study now makes it possible, through population assignment techniques, to identify the most likely geographic source of a boll weevil reintroduced to an eradication zone, which will help action agencies decide the most appropriate mitigation response.

Introduction

The boll weevil, *Anthonomus grandis* Boheman (Coleoptera: Curculionidae), has a very narrow host range, reproducing successfully only in a handful of genera in the Malvaceae (Fryxell & Lukefahr, 1967; Arzaluz & Jones, 2001). Although there is evidence that it has used cultivated cotton (*Gossypium hirsutum*) as a host since at least 900 CE (Warner & Smith, 1968), the boll weevil's range was restricted to Southern Mexico and Central America until sometime in the mid-nineteenth century. It subsequently expanded northward, probably in response to larger-scale cotton production, reaching the Southern tip of Texas in 1892. Over the course of about three decades, it spread throughout the Cotton Belt of the southeastern US, reaching the Atlantic Coast by the early 1920's (Hunter & Coad, 1923; Burke et al., 1986). A secondary range expansion into the High Plains of Texas and New Mexico began in the late 1950's (Bottrell, Rummel & Adkisson, 1972). Colonization of this region was delayed apparently by a combination of a geographic barrier – the Caprock Escarpment – and a colder climate (Stavinoha & Woodward, 2001). The boll weevil currently is expanding its range southward through cultivated cotton in South America as well (Lukefahr, Barbosa & Sabrinho, 1994; Scataglini, Confalonieri & Lanteri, 2000).

†Formerly: USDA-ARS, Areawide Pest Management Research Unit, 2771 F & B Rd., College Station, TX 77845, USA.

An eradication program initiated in the U.S. about 25 years ago has progressively eliminated the boll weevil from several states (Smith, 1998; Grefenstette & El-Lissy, 2003), but still is ongoing in eight states and parts of Mexico. It is a hugely expensive program, and the threat of reintroductions by dispersal from areas still infested is a constant concern (e.g., Allen et al., 2005; Kiser & Catanach, 2005). Thus, characterization of boll weevil dispersal capacity and patterns has become a pressing issue, but studying long-distance movement in this species, as in most insects, is fraught with difficulties. Range expansion records (Hunter & Coad, 1923; Culin et al., 1990; Lukefahr, Barbosa & Sabrinho, 1994), mark-release-recapture experiments (Johnson et al., 1975; Guerra, 1988; Raulston et al., 1996), and pheromone trapping (Jones et al., 1992; Spurgeon et al., 1997) indicate that boll weevils can disperse 100–300 km, but the frequency and magnitude of dispersal over such distances and beyond cannot be adequately addressed with such data.

To gain better insight into the patterns and extent of boll weevil dispersal among populations, we have employed a variety of DNA markers to estimate gene flow (Kim & Sappington, 2004a, b), the latter being an index of migration (Lowe, Harris & Ashton, 2004). There have been a few previous studies of boll weevil population structure based on allozyme (Bartlett, Randall & May, 1983; Terranova, Jones & Bartlett, 1990; Biggers et al., 1996), mtDNA RFLP (restriction fragment length polymorphism) (Roehrdanz & North, 1992; Roehrdanz, 1995, 2001), and RAPD (random amplification of polymorphic DNA) (Scataglini, Confalonieri & Lanteri, 2000) markers, but the focus of these studies was not on dispersal between populations. Using primers developed by Roehrdanz and Degrugiller (1998) for long PCR (polymerase chain reaction) of boll weevil mtDNA, Kim and Sappington (2004a) examined mtDNA variability in 20 populations across eight states and northeastern Mexico, and estimated that migrant exchange was frequent between populations separated by less than 300 km. The same individuals from 18 of the same populations were further examined for population structuring and gene flow using RAPD markers (Kim & Sappington, 2004b). Results of the RAPD analyses suggested that effective migration was high enough at distances of $\sim$300–400 km to prevent genetic differentiation, but the estimated numbers of migrants exchanged between populations were much lower than those estimated from mtDNA PCR-RFLP data. Although mtDNA PCR-RFLP and RAPD markers are widely used and are valuable tools for population geneticists, they have certain drawbacks. The maternal inheritance of mtDNA and the dominant mode of inheritance of RAPD markers can limit their utility in characterizing population structuring (Dowling, Moritz & Palmer, 1990; Black, 1993; Lowe, Harris & Ashton, 2004).

Differing estimates in migrant-exchange between boll weevil populations based on mtDNA PCR-RFLP and RAPD markers prompted us to develop and apply microsatellite markers to the same populations. Microsatellites are short tandem repeats of nuclear DNA sequences, and are increasingly complementing or replacing other markers for numerous applications in evolutionary genetics (Parker et al., 1998; Lowe, Harris & Ashton, 2004). Microsatellite loci are highly polymorphic, and can be scored relatively easily using PCR followed by genotyping with an automated sequencer. Because microsatellites are co-dominant markers, deviation from Hardy–Weinberg equilibrium can be assessed directly, making them ideal for assessing genetic structure within populations as well as gene flow among populations. They are an especially valuable tool for inferring levels of genetic variation and patterns of population structure among closely related or recently diverged populations (Roderick, 1996; Haig, 1998; Batley et al., 2004; Kim & Sappington, 2005), presumably the case for invasive species like the boll weevil. With microsatellites, one can perform assignment and exclusion tests of individuals relative to reference populations (Rannala & Mountain, 1997; Cornuet et al., 1999; Paetkau et al., 2004), a capability that could be especially useful in the context of characterizing the origin of boll weevils captured in an eradication zone. Like all DNA markers, microsatellites have some drawbacks, such as potential size homoplasy and null alleles, but their biggest disadvantage is the time and expense of developing them as markers (Lowe, Harris & Ashton, 2004).

Kim and Sappington (2004c) developed 14 microsatellite loci from boll weevil, 12 of which promised to be useful in population genetics studies. In this study, we used 11 of these microsatellites to analyze the genetic structuring of the

same populations of boll weevils from 18 locations previously characterized with mtDNA PCR-RFLP and RAPD markers (Kim & Sappington, 2004a, b). The samples included DNA from the same individuals, but the sample sizes were increased, where possible, with individuals from the same collections. By comparing the results obtained from microsatellite markers to those obtained previously from other markers, we hoped to obtain the most robust view possible of boll weevil dispersal patterns.

The objectives of this study were (1) to gain a better understanding of genetic structuring of geographic boll weevil populations in North America, (2) to evaluate boll weevil dispersal capacity and tendency, and (3) to use multilocus genotype data from each population to estimate the most likely source population of potential immigrants. Finally, we discuss the implications of population assignment techniques for boll weevil management in an eradication context.

Materials and methods

Sample collection

Boll weevils were sampled from 18 locations in 8 U.S. states and northeastern Mexico. Locations were assigned to three geographic regions as in previous studies (Kim & Sappington, 2004a, b). South-central region: Tampico, Mexico (MEX), and Weslaco (WTX), Kingsville (KTX), El Campo (ETX), College Station (CSTX), and Waxahachie (WATX), Texas; Western region: Hobart, Oklahoma (HOK), and Stamford (STX), Childress (CHTX), Plainview (PTX), and Big Spring (BTX), Texas, and Artesia, New Mexico (ANM); Eastern region: Winnsboro, Louisiana (WLA), Little Rock, Arkansas (LAR), Cleveland (CMS) and Yazoo City (YMS), Mississippi, Malden, Missouri (MMO), and Brownsville, Tennessee (BTN) (Figure 1). Details of collectors and dates are reported in Kim and Sappington (2004a). Within a region, each location was separated by $<$300 km from its nearest neighbor (Kim & Sappington, 2004a, b). An exception was the MEX site, which is located $\sim$430 km south of WTX, its nearest neighbor (Figure 1). Although there are occasional substantial gaps in cotton production in some areas, which potentially could serve as dispersal barriers, such as between WTX and KTX, or between MEX and WTX, cotton is generally common throughout each designated region. At each location, boll weevils of both sexes were collected in traps baited with aggregation pheromone, and then frozen. The individuals analyzed in previous RFLP and RAPD studies (Kim & Sappington, 2004a, b) also were examined in this study, along with additional individuals, when possible, collected at the same locations to increase sample size for the microsatellite work. Due to low population sizes in response to eradication efforts and other factors, the final sample size attained in a few of the populations, in particular BTX, STX and CSTX, were well below our target of 30. However, this is not likely a severe problem for those aspects of our study where comparisons are made between regions comprised of multiple populations. Nevertheless, caution should be exercised when interpreting analyses of the smaller datasets.

Microsatellites

Fourteen polymorphic dinucleotide-repeat microsatellite markers were developed for *A. grandis* (Kim & Sappington, 2004c). Because the presence of null alleles at microsatellite loci are relatively common in insects and can bias measures of population genetics parameters, we conducted a preliminary survey for detection of null alleles. Eleven of these microsatellites (AG-D1–AG-D7 and AG-9–AG-D12) were used in this study because they passed parentage analysis – where one looks for mismatches between parental and offspring genotypes – and Mendelian segregation tests in 6 controlled family crosses (K.S. Kim & T. W. Sappington, unpublished data). For these 11 loci, we assume that even if null alleles are present in a natural population, the frequency will be very low and have a negligible effect on parameter estimates. These loci were amplified in multiplexed PCR reactions, and individuals were genotyped using a Beckman-Coulter CEQ 8000 Genetic Analysis System, following the procedures described by Kim and Sappington (2004c).

Data analysis

Genetic diversity and population differentiation

Allele frequencies, the mean number of alleles per locus, observed heterozygosity (H_O), and expected

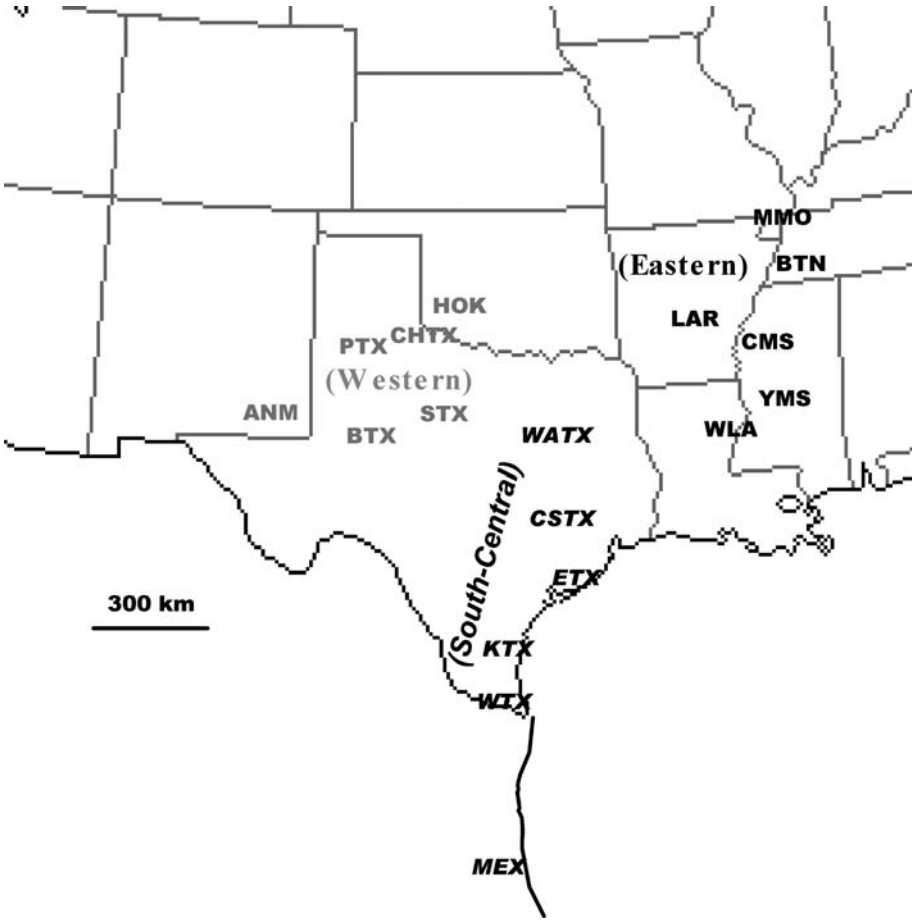

Figure 1. Geographic locations of boll weevils sampled. South-Central region, black italics font; Eastern region, black normal font; Western region, gray normal font.

heterozygosity (H_E) under Hardy–Weinberg assumptions for each locus were computed using the GENETIX software package (Belkhir et al., 2000). The two measures of heterozygosity are highly correlated, but in this study we focused on H_E because it is considered a better estimator of the genetic variability present in a population (Nei, 1987). The GENETIX program was used further to calculate $G_{ST,}$ an estimator of genetic differentiation, and the number of migrants per generation (N_m), an indirect indicator of gene flow (Wright, 1969). Adjusted allelic diversity to account for variation in sample sizes was calculated using both bootstrapping and Jackknifing techniques implemented in the program AGARST (Harley, 2001). R_{ST} (an estimate of population subdivision for stepwise mutation processes, Rousset 1996), *F*-statistics (Weir & Cockerham, 1984) and pairwise F_{ST}'s were calculated using the program FSTAT (Goudet, 1995), applying the sequential Bonferroni correction over loci in deriving significance levels (Rice, 1989).

For Hardy–Weinberg equilibrium (HWE) estimation, we followed the probability test approach (Guo & Thomson, 1992) using the program GENEPOP (Raymond & Rousset, 1995). The relationship between $F_{ST}/(1-F_{ST})$ and geographic distance between populations was calculated from 5000 replications and normalized by the Mantel statistic Z option using the MXCOMP program in NTSYSPC, version 1.70 (Rohlf, 1992). Multiple comparisons of the number of polymorphic loci, heterozygosity, and F_{ST}'s across regions were performed with the nonparametric Kruskal–Wallis test corrected for experimentwise error rate (Daniel, 1990) using Statistix software (Analytical Software, 1998).

Genetic relationship

Genetic divergence between populations based on allele frequencies was calculated as genetic distance (D_A) (Nei, Tajima & Tateno, 1983) using the DISPAN computer program (Ota, 1993). Phylogenetic trees were constructed by neighbor-joining (NJ) clustering (Saitou & Nei, 1987), and by the unweighted pair group method with the arithmetic mean (UPGMA) (Sneath & Sokal, 1973) using D_A distance. Bootstrap resampling ($n = 1000$) tested the robustness of dendrogram topologies. A principal component analysis (PCA) was applied to a covariance matrix of allele frequencies across all variable loci using the program XLSTAT (Agresti, 1990). The geometric relationship among boll weevil populations was visualized with a scattergram of the factor score data along the two PC axes that accounted for the most variation.

Bottleneck tests

Evidence of recent population bottlenecks was assessed using three approaches. Mutation-drift equilibrium and mode-shift were assessed using the program BOTTLENECK 1.2 (Cornuet & Luikart, 1996). We employed both a strict stepwise mutation model (SMM) (Kimura & Ohta, 1978), and a two-phase model (TPM) (Di Rienzo et al., 1994) in which 90% of the microsatellite mutations followed the strict SMM and 10% produced multistep changes. Wilcoxon sign-rank tests were used to determine whether deviations of observed heterozygosity relative to that expected at drift-mutation equilibrium were significant ($\alpha = 0.05$). Two-tailed tests were used because the population histories were not known (Luikart & Cornuet, 1998). We looked for a mode-shift in allele frequency distribution, which can be used as a qualitative indicator of population bottlenecks (Luikart et al., 1998). Finally, significant differences in allelic diversity between the MEX location, representing the ancestral population, and other locations were identified by a Kruskal–Wallis test ($\alpha = 0.05$) across all loci.

Assignment/exclusion test, and detection of first generation migrants

The program GeneClass2 (Piry et al., 2004) was used for both assignment/exclusion tests and the detection of first generation migrants. As an assignment criterion for both analyses, the Bayesian statistical approach of Rannala and Mountain (1997) was chosen, which has proven to be more accurate than frequency and distance based methods (Cornuet et al., 1999). We used a 'simulation' approach to evaluate the proportion of individuals in a population with genotypes compatible with having arisen in a reference population. The method computes the likelihood of a genotype occurring in the population by simulating multilocus genotypes based on allele frequencies of each reference population, and then compares the likelihood of the genotype in an individual to a distribution of likelihoods of simulated genotypes for each reference population. If the individual genotype likelihood is below a given threshold (e.g. $\alpha = 0.01$), the population is excluded as the possible origin of the individual (Cornuet et al., 1999). Unlike the 'direct' assignment method, this exclusion method does not assume that the true population of origin has been sampled, because each population is treated separately. Frequency probabilities of multilocus genotypes in each reference population were determined using Monte Carlo simulations of 10,000 independent individuals for the population.

Determination of the most likely source population of each potential immigrant in the current generation (see Paetkau et al., 2004) was conducted using the 'Detection of first generation migrants' criterion. The L_{home}/L_{max} ratio was used to compute the likelihood of migrant detection (L), where L_{home} is the likelihood computed for the population from which the individual was sampled, and L_{max} is the highest likelihood value among all population samples including the home population. This analysis was conducted with a simulation of 1000 independent individuals at thresholds of $\alpha = 0.05$ and $\alpha = 0.01$ using the assignment criterion of Rannala and Mountain (1997), and a new Monte Carlo simulation algorithm according to Paetkau et al. (2004).

Results

Allele frequencies and genetic variability

A total of 65 alleles were detected across the 11 boll weevil microsatellite loci analyzed (Table 1). The number of alleles per locus averaged 5.9 and ranged from 3 to 10. Although some unique alleles were found in South-central populations,

Table 1. Characterization of 11 microsatellite loci analyzed across all boll weevil populations

Locus	Allele number	F_{IS}	F_{IT}	F_{ST}	G_{ST}	R_{ST}
AG-D1	4	0.073	0.463***	0.420***	0.411	0.469
AG-D2	3	−0.029	0.273***	0.294***	0.279	0.286
AG-D3	6	−0.017	0.264***	0.277***	0.269	0.388
AG-D4	7	−0.006	0.170***	0.175***	0.172	0.269
AG-D5	10	0.059	0.313***	0.270***	0.265	0.043
AG-D6	6	0.063	0.299***	0.251***	0.237	0.177
AG-D7	3	0.054	0.563***	0.538***	0.525	0.543
AG-D9	5	0.359***	0.602***	0.379***	0.379	0.288
AG-D10	9	0.126***	0.222***	0.109***	0.124	0.217
AG-D11	8	0.112**	0.163***	0.057***	0.067	0.047
AG-D12	4	0.103*	0.284***	0.202***	0.203	0.222
All loci	65	0.082***	0.303***	0.241***	0.237	0.268

Probability that value is different than zero, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

they are unlikely to be useful as population specific markers because of their low frequency. Locus AG-D9 was fixed for a single allele (140 bp) in all Western and Eastern populations, and in CSTX. Locus AG-D1 showed only one allele (116 bp) in three Western populations (HOK, BTX, ANM) and in all Eastern populations. Locus AG-D2 was fixed for a single allele (127 bp) in Eastern populations, and locus AG-D3 was fixed for a single allele (158 bp) in all Eastern populations and in CSTX. All of these loci were polymorphic in all South-central populations except CSTX.

Across all boll weevil populations, all three *F*-statistics were significantly different from zero for 4 of 11 loci, and the overall means were significant as well (Table 1). F_{IS} estimates across all populations ranged from −0.029 (AG-D2) to 0.359 (AG-D9), averaging 0.082. Thus, despite variable locus-specific deviations, overall departure from HWE was in the direction of heterozygote deficiency. On average, populations exhibited a 8.2% deficit of heterozygotes, whereas the total metapopulation was 30.3% heterozygote deficient.

Three different measures (F_{ST}, G_{ST} and R_{ST}) of genetic differentiation are very similar, and levels of apparent population subdivision are considerable (Table 1). The multi-locus F_{ST} estimate indicates that about 24% of the total genetic variation is accounted for by population differences, with the remaining 76% being attributable to differences among individuals.

Allelic diversity, observed (H_O) and expected (H_E) heterozygosity, the inbreeding coefficient (F_{IS}), and *p*-values for deviations from HWE were calculated over all loci for each boll weevil population (Table 2). Allelic diversity varied from 1.4 in the sample from Yazoo city, MS to 5.0 in the sample from Tampico, Mexico, with a median of 2.5. H_E ranged from 0.078 in the sample from Cleveland, MS to 0.589 in Mexico, with a median of 0.187. The three related genetic diversity variables (allelic diversity, H_E, and H_O) showed similar tendencies within and among regions, revealing generally greater genetic diversity in Southern than Northern populations, and greater in the Western than the Eastern among Northern populations.

Estimates of genetic diversity based on allelic diversity and heterozygosity (H_O and H_E) differed significantly among the three regions (allelic diversity: KW statistic = 13.3, $p = 0.001$; H_O: KW statistic = 12.0, $p = 0.003$; H_E: KW statistic = 11.9, $p = 0.003$). All three measures of diversity differed significantly ($p < 0.05$; critical rank value = 7.38) between South-central (medians: allelic diversity = 3.6; $H_O = 0.288$; $H_E = 0.324$) and Eastern (medians: allelic diversity = 1.5; $H_O = 0.078$; $H_E = 0.089$) regions, but not between South-central and Western (medians: allelic diversity = 2.2; $H_O = 0.185$; $H_E = 0.197$) regions. Western and Eastern regions differed significantly in allelic diversity, but not in heterozygosity.

From 198 instances over all samples and over all loci (18 populations and 11 loci), 13 significant

Table 2. Number (N) of boll weevils sampled from indicated locations and characteristics of populations calculated for 11 microsatellite loci (see Table 1)

Location	N	Allelic diversity[a]		H_E	H_O	F_{IS}	p[b]
		Observed	Boot (JK)				
MEX	30	5.0 a	4.2 (4.5)	0.589	0.561	0.049	0.512 (1)[d]
WTX	30	4.4 ab	3.8 (3.9)	0.564	0.509	0.099**[c]	0.130 (1)
KTX	30	4.0 ab	3.2 (3.4)	0.441	0.400	0.094*	<0.01 (3)
ETX	27	3.2 abc	2.3 (2.5)	0.206	0.175	0.150**	0.156 (1)
CSTX	16	2.3 abcd	2.0 (2.2)	0.188	0.171	0.095	0.977 (0)
WATX	30	3.0 abcd	2.2 (2.4)	0.169	0.158	0.068	0.996 (0)
HOK	32	2.2 abcd	1.9 (2.0)	0.186	0.156	0.161**	0.062 (1)
STX	18	2.6 abcd	2.2 (2.4)	0.241	0.222	0.080	0.800 (1)
CHTX	31	2.4 abcd	2.0 (2.0)	0.202	0.199	0.013	0.432 (2)
PTX	35	2.2 abcd	1.9 (1.9)	0.192	0.171	0.111*	0.269 (1)
BTX	15	1.6 cd	1.6 (1.6)	0.214	0.279	−0.320	0.767 (0)
ANM	30	1.9 abcd	1.7 (1.8)	0.160	0.158	0.015	0.999 (0)
WLA	30	1.5 cd	1.4 (1.4)	0.093	0.073	0.223**	0.132 (1)
LAR	27	1.5 cd	1.4 (1.4)	0.084	0.088	−0.046	0.612 (0)
CMS	33	1.6 cd	1.4 (1.4)	0.078	0.069	0.119	0.844 (0)
YMS	32	1.4 d	1.4 (1.4)	0.103	0.068	0.343***	<0.01 (1)
MMO	33	1.9 bcd	1.7 (1.7)	0.154	0.132	0.142*	0.573 (0)
BTN	31	1.5 cd	1.3 (1.3)	0.082	0.082	−0.007	1.000 (0)
Mean	28.3	2.4	2.1 (2.2)	0.219	0.204	–	<0.01 (13)

Allelic diversity (mean number of alleles per locus) according to direct observation and bootstrap (Boot) or jackknife (JK) resampling, expected heterozygosity (H_E) at Hardy–Weinberg equilibrium, observed heterozygosity (H_O), inbreeding coefficient (F_{IS}), and the probability (p) of being in Hardy–Weinberg equilibrium.

[a] Means followed by the same letter are not significantly different, Kruskal–Wallis test ($\alpha = 0.05$).

[b] Probability values using Fisher's method implemented by GENEPOP program.

[c] Probability from multi-locus test that there is no heterozygote deficiency (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

[d] Number in parentheses indicates the number of loci showing a significant departure ($p < 0.05$) from Hardy–Weinberg equilibrium.

deviations from HWE were detected (Table 2). After corrections for multiple significance tests, deviations over all loci were significant in two populations. All of these significant deviations were associated with positive F_{IS} values, revealing deviation in the direction of heterozygote deficiency.

Gene differentiation among populations

Genetic differentiation between each pair of populations (pairwise F_{ST}), and the effective number of migrants exchanged per generation (N_m) were calculated (Table 3). F_{ST} estimates calculated by Weir and Cockerham's method (1984) ranged from −0.014 (WLA versus BTN) to 0.573 (MEX versus CMS). After corrections for multiple comparisons, far Southern populations (MEX, WTX, KTX) showed significant genetic differentiation from other populations whereas other South-central populations (ETX, CSTX, WATX) revealed little genetic differentiation from most populations. A low level of genetic divergence was observed in most paired comparisons among samples within the Western region (except BTX) (median $F_{ST} = 0.020$), and within the Eastern region (median $F_{ST} = 0.024$). However, most pairwise comparisons between Eastern and Western populations differed significantly from 0 (median $F_{ST} = 0.134$), indicating genetic differentiation between those regions. Differences in distribution of fixed alleles at AG-D2 and AG-D3 in Eastern and Western populations are consistent with little ongoing gene flow between these regions.

Region significantly affected within-region paired F_{ST} values (KW statistic = 13.92, $p = 0.001$). Very small pairwise F_{ST} values were

Table 3. Pairwise F_{ST} estimates (Weir & Cockerham 1984) and number of boll weevil migrants per generation (N_m) (in parentheses) (above diagonal), and Nei's D_A genetic distance (below diagonal) between locations

	MEX	WTX	KTX	ETX	CSTX	WATX	HOK	STX	CHTX
MEX	–	0.139** (1.6)	0.269** (0.7)	0.458** (0.3)	0.445** (0.3)	0.486** (0.3)	0.482** (0.3)	0.426** (0.3)	0.477** (0.3)
WTX	0.113	–	0.028^{NS} (8.7)	0.174** (1.2)	0.188** (1.1)	0.208** (1.0)	0.223** (0.9)	0.187** (1.1)	0.213** (0.9)
KTX	0.212	0.041	–	0.076 ^{NS} (3.0)	0.094** (2.4)	0.105** (2.1)	0.129** (1.7)	0.112** (2.0)	0.120** (1.8)
ETX	0.377	0.126	0.062	–	0.015^{NS} (16.9)	0.001^{NS} (225.3)	0.044^{NS} (5.4)	0.066^{NS} (3.6)	0.046^{NS} (5.2)
CSTX	0.430	0.166	0.100	0.030	–	0.016^{NS} (15.6)	0.012^{NS} (20.4)	0.017^{NS} (14.7)	0.001^{NS} (667.1)
WATX	0.393	0.150	0.080	0.029	0.029	–	0.015^{NS} (16.3)	0.055^{NS} (4.3)	0.031^{NS} (7.8)
HOK	0.421	0.188	0.125	0.054	0.040	0.021	–	0.003^{NS} (72.6)	0.008^{NS} (30.6)
STX	0.411	0.177	0.114	0.054	0.039	0.029	0.015	–	0.006^{NS} (38.9)
CHTX	0.422	0.172	0.110	0.044	0.025	0.026	0.014	0.019	–
PTX	0.422	0.172	0.111	0.042	0.033	0.024	0.014	0.019	0.012
BTX	0.428	0.183	0.138	0.062	0.063	0.041	0.040	0.052	0.051
ANM	0.444	0.194	0.131	0.042	0.042	0.025	0.015	0.023	0.027
WLA	0.518	0.234	0.138	0.055	0.028	0.036	0.041	0.051	0.032
LAR	0.522	0.242	0.145	0.059	0.031	0.036	0.040	0.050	0.035
CMS	0.511	0.235	0.140	0.061	0.031	0.036	0.043	0.054	0.034
YMS	0.530	0.249	0.153	0.069	0.036	0.046	0.046	0.053	0.036
MMO	0.467	0.201	0.128	0.047	0.017	0.034	0.028	0.036	0.017
BTN	0.520	0.239	0.142	0.058	0.033	0.041	0.043	0.058	0.034

Probability of being different than zero after corrections for multiple comparisons; *$p < 0.05$; **$p < 0.01$; NS: not significant; pan: panmictic.

observed within Western and Eastern regions, and were significantly lower than those among South-central populations (median $F_{ST} = 0.139$) (critical rank value = 11.48, $p < 0.05$).

Genetic distance inferred from $F_{ST}/(1-F_{ST})$ was positively correlated with geographic distance between populations (Figure 2), indicating that individuals from adjacent locations exchange genes more frequently than those from more distant locations. However, when the three major regions were considered separately, there was a significant correlation between genetic and geographic distance only within the South-central region (Figure 2). Correlations were not significant within the Eastern or Western regions, suggesting recent range expansion and/or frequent gene flow among populations within these regions.

Indirect estimates of gene flow (N_m values) varied from 0.2 (MEX versus Eastern populations) to panmixus. Moderate to high gene flow was indicated for most population pairs, even across different regions, but little was observed between extreme Southern (MEX and WTX) and Eastern populations. The values indicate that at least one effective migrant per generation was exchanged between populations separated by ~500 km, and three migrants exchanged between populations separated by ~300 km or less. There was relatively frequent gene flow among Eastern populations, implying panmixus in many population pairs.

Genetic relationships among populations

Nei's D_A genetic distances (Table 3) were calculated to infer genetic divergence among populations, and ranged from 0.002 (WLA versus LAR) to 0.530 (MEX versus YMS). D_A values differed significantly by region (KW statistic = 32.8, $p < 0.0001$). There was a higher degree of genetic divergence among South-central populations (median $D_A = 0.113$) than among Western populations (median $D_A = 0.023$), and genetic divergence of each was significantly higher than that among Eastern populations (median $D_A = 0.005$) (critical rank value = 11.48 at $\alpha = 0.05$). Genetic divergence between Western and Eastern populations (median $D_A = 0.047$) was significantly higher than within those regions (median $D_A = 0.015$) (KW statistic = 36.0, $p < 0.0001$). The population from Mexico showed a consistently high level of pairwise genetic divergence from other populations (median $D_A = 0.428$).

To resolve phylogenetic relationships among populations, NJ and UPGMA trees were reconstructed based on Nei's D_A genetic distances

PTX	BTX	ANM	WLA	LAR	CMS	YMS	MMO	BTN
0.484** (0.3)	0.409** (0.4)	0.496** (0.3)	0.551** (0.2)	0.548** (0.2)	0.573** (0.2)	0.558** (0.2)	0.517** (0.2)	0.562** (0.2)
0.216** (0.9)	0.153** (1.4)	0.225** (0.9)	0.310** (0.6)	0.310** (0.6)	0.341** (0.5)	0.328** (0.5)	0.255** (0.7)	0.324** (0.5)
0.129** (1.7)	0.103** (2.2)	0.139** (1.6)	0.204** (1.0)	0.207** (1.0)	0.238** (0.8)	0.230** (0.8)	0.153** (1.4)	0.219** (0.9)
0.038^{NS} (6.4)	0.068^{NS} (3.4)	0.028^{NS} (8.6)	0.155* (1.4)	0.155* (1.4)	0.198* (1.0)	0.208** (1.0)	0.072^{NS} (3.2)	0.179* (1.2)
0.024^{NS} (10.1)	0.120^{NS} (1.8)	0.043^{NS} (5.6)	0.061^{NS} (3.8)	0.061^{NS} (3.9)	0.089^{NS} (2.6)	0.106^{NS} (2.1)	0.006^{NS} (38.5)	0.083^{NS} (2.8)
0.026^{NS} (9.4)	0.075^{NS} (3.1)	0.016^{NS} (15.4)	0.130^{NS} (1.7)	0.132^{NS} (1.7)	0.181* (1.1)	0.188 * (1.1)	0.050^{NS} (4.8)	0.155 * (1.4)
0.014^{NS} (17.7)	0.082* (2.8)	0.020^{NS} (12.2)	0.090* (2.5)	0.086^{NS} (2.6)	0.122^{NS} (1.8)	0.136** (1.6)	0.038^{NS} (6.3)	0.109** (2.1)
0.011^{NS} (23.4)	0.084** (2.7)	0.035^{NS} (6.9)	0.109* (2.0)	0.108^{NS} (2.1)	0.135** (1.6)	0.125** (1.8)	0.044 * (5.4)	0.133** (1.6)
0.018^{NS} (13.3)	0.106** (2.1)	0.032* (7.5)	0.063^{NS} (3.7)	0.070 * (3.3)	0.094** (2.4)	0.091** (2.5)	0.002^{NS} (152.6)	0.078** (3.0)
–	0.070** (3.3)	0.012^{NS} (20.8)	0.144** (1.5)	0.142** (1.5)	0.176** (1.2)	0.176** (1.2)	0.049** (4.8)	0.167** (1.3)
0.040	–	0.052^{NS} (4.6)	0.303** (0.6)	0.310** (0.6)	0.359** (0.5)	0.345** (0.5)	0.182** (1.1)	0.335** (0.5)
0.023	0.029	–	0.193** (1.1)	0.191** (1.1)	0.237** (0.8)	0.244** (0.8)	0.079** (2.9)	0.218** (0.9)
0.044	0.081	0.054	–	-0.012^{NS} (pan)	-0.001^{NS} (pan)	0.010^{NS} (25.3)	0.037^{NS} (6.5)	-0.014^{NS}(pan)
0.044	0.085	0.051	0.002	–	-0.006^{NS} (pan)	0.028^{NS} (8.8)	0.051^{NS} (4.7)	-0.004^{NS} (pan)
0.046	0.091	0.058	0.004	0.004	–	0.029^{NS} (8.5)	0.073^{NS} (3.2)	-0.001^{NS} (pan)
0.048	0.097	0.064	0.004	0.005	0.007	–	0.059^{NS} (4.0)	0.024^{NS} (10.1)
0.030	0.071	0.036	0.015	0.017	0.016	0.017	–	0.053^{NS} (4.5)
0.048	0.087	0.055	0.003	0.004	0.004	0.008	0.017	–

(Figure 3). We rooted both trees with the MEX population, based on the known history of the spread of the boll weevil from Mexico into the U.S. in the late 19th century (Burke et al., 1986). The UPGMA tree (Figure 3(b)) is not well-resolved, but both trees show the far Southern populations

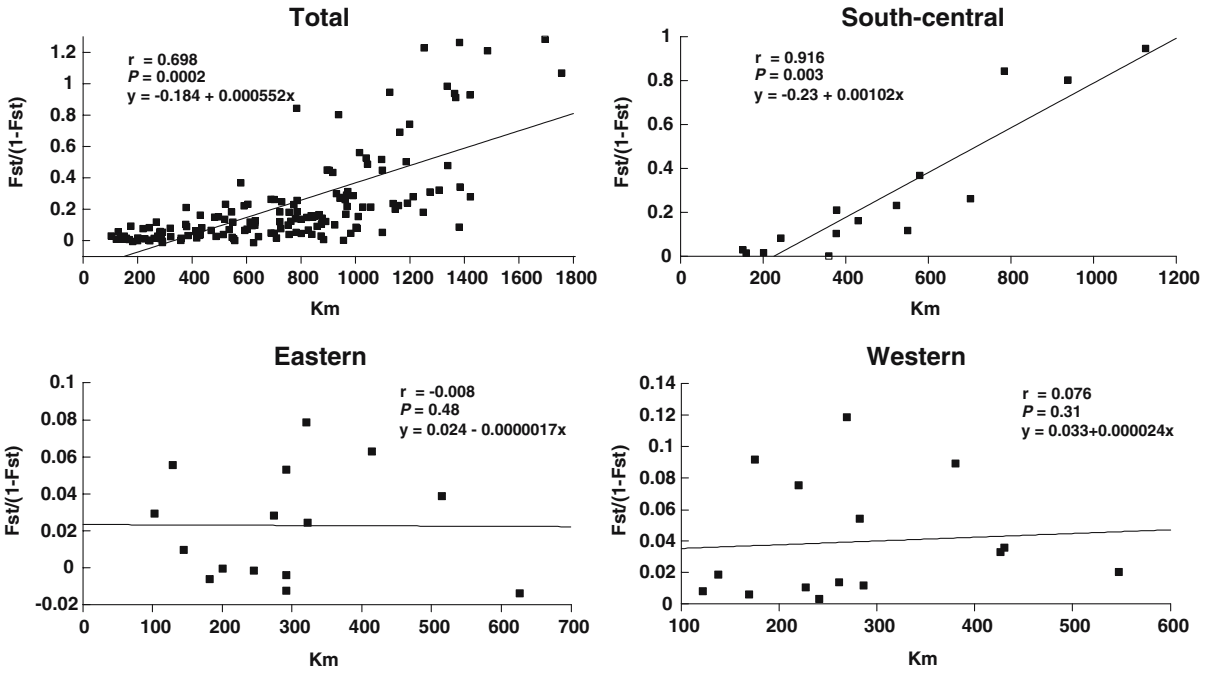

Figure 2. Isolation by distance. Relationship of genetic distance $[F_{ST}/(1-F_{ST})]$ to geographic distance among all boll weevil populations sampled (Total), or among populations within regions.

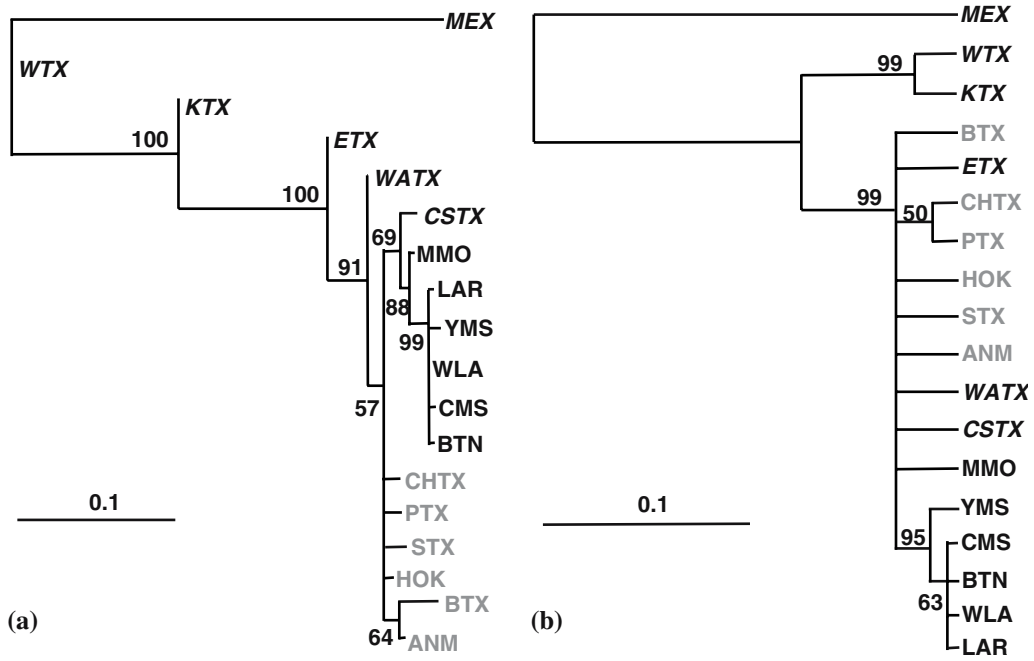

Figure 3. Dendrograms of genetic relationships among 18 boll weevil populations based on D_A genetic distance (Nei, Tajima & Tateno, 1983). (a) Neighbor-joining analysis; (b) UPGMA analysis. South-Central region, black italics font; Eastern region, black normal font; Western region, gray normal font. The numbers at nodes are percentage bootstrap values from 1000 replications of resampled loci. Nodes with bootstrap support less than 50% are collapsed into unresolved polytomies.

(MEX, WTX and KTX) as clearly diverged from all others, and basal to the other populations. Eastern populations, except MMO, were grouped into a single clade with strong bootstrap support in both trees (99% in NJ tree and 95% in UPGMA tree), but the Western populations do not cluster as

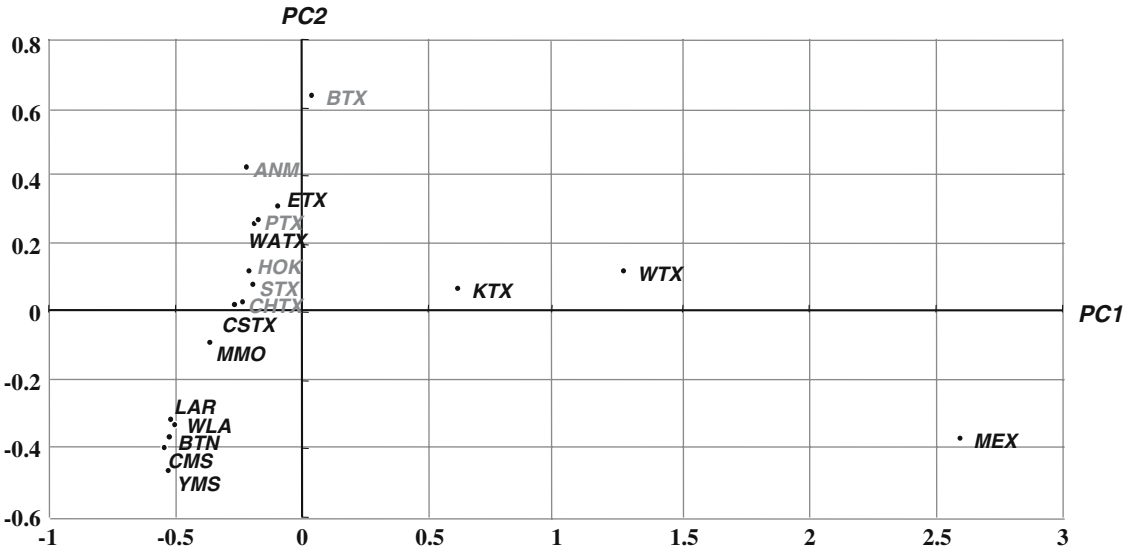

Figure 4. Scatter diagram of factor scores for 18 boll weevil populations derived from PC analysis of a covariance matrix of allele frequencies for 11 microsatellites. PC1 (*X*-axis) and PC2 (*Y*-axis) accounted for 77 and 13% of the variance, respectively. South-Central region, black italics font; Eastern region, black normal font; Western region, gray normal font.

a clade. Although Nei's DA distance is one of the distance methods that provides an accurate tree topology when using microsatellites as genetic markers (Takezaki & Nei, 1996), caution in interpreting the trees is required, because the accuracy is low when fewer than 30 markers are used (Nei, Tajima & Tateno, 1983).

To obtain further insights into the relationships among boll weevil populations, principal component (PC) analysis was conducted using allele frequency data. Mean factor scores for the 18 populations were plotted along the first two PC axes, which together accounted for 90% of the total variance in the covariance matrix (Figure 4). The largest portion of the variance (77%) was accounted for by PC1. The far Southern populations (MEX, WTX and KTX) are distinguished from the other populations along the PC1 axis, while Eastern and Western populations diverge along the PC2 axis. The more Northern South-central populations (ETX, CSTX, and WATX) cluster with the Western populations along both axes.

Population bottlenecks

Under both the SMM and TPM models of microsatellite mutation, Wilcoxon sign-rank tests indicated a significant excess of heterozygosity relative to drift-mutation equilibrium in the BTX population, while significant heterozygote deficiencies were detected in the KTX, ETX, CSTX, WATX, and STX populations. Mode-shift evidence for a bottleneck was observed in the BTX and YMS populations. In an independent and more sensitive assessment, a Kruskal–Wallis test indicated that the number of alleles per locus, or allelic diversity (AD), was significantly dependent on population (KW statistic = 98.2; $p < 0.0001$). KW separation of distributions (Table 2) showed that the presumed ancestral and demographically most-stable population, MEX, differs significantly from BTX and all Eastern populations, indicating that the latter have undergone bottlenecks.

Assignment test and detection of first generation migrants

The percentage of boll weevil individuals sampled from each population that were excluded ($\alpha = 0.05$) as potential immigrants from each possible donating (reference) population was calculated (Table 4). Populations from most locations, except from the far south (MEX, WTX and KTX), contained members whose potential origins from other populations could not be excluded with ≥95% certainty. In other words, most populations contained members whose genotypes indicated that they could have originated from another population with a probability of at least 5%. South-central populations (except MEX) were especially prone to be identified as potential sources (i.e. prone not to be excluded as potential sources) for individuals in other populations. There was less likelihood that individuals sampled in Western populations originated in the Eastern region than vice versa. Indeed few individuals sampled in the Eastern region could be excluded with ≥95% certainty from originating elsewhere, except that the farthest west populations (BTX and ANM) were unlikely sources for many Eastern individuals.

Percentage of probable first generation migrants in each population, and the number of these assigned to the most likely source population, are shown in Table 5. In total, 41 individuals were identified as probable first generation migrants at a threshold of $\alpha = 0.05$. All populations harbored at least one probable immigrant. Most migrant exchange (61%) occurred among the six populations within each region.

Discussion

Genotype data at 11 microsatellite loci from boll weevils sampled from 18 geographical locations indicate significant genetic structuring of populations throughout the central Cotton Belt of the United States. However, little genetic structuring was detected among populations located relatively near one another. Apart from some discrepancies in estimates of genetic differentiation, and therefore in magnitude of migrant exchange between populations, the conclusions of this study are generally consistent with those drawn from analyses of mtDNA PCR-RFLP and RAPD data (Kim & Sappington, 2004a, b).

Genetic diversity of the boll weevil in the central Cotton Belt

Indices of genetic diversity differed significantly among the three regions. However, we found that

Table 4. Percentage of boll weevil individuals in sampled population excluded from (i.e. determined not to be a potential immigrant from) each reference population ($p < 0.05$) based on the Bayesian approach

Sampled Pop.	Reference population																	
	MEX	WTX	KTX	ETX	CSTX	WATX	HOK	STX	CHTX	PTX	BTX	ANM	WLA	LAR	CMS	YMS	MMO	BTN
MEX	0	77	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
WTX	50	3	57	97	100	100	100	100	100	100	100	100	100	100	100	100	100	100
KTX	77	3	7	67	87	90	100	87	97	100	100	100	100	100	100	100	100	100
ETX	100	0	0	4	15	22	41	15	37	41	48	41	81	85	78	93	52	85
CSTX	100	0	0	6	0	25	31	19	25	31	50	44	56	63	63	69	31	63
WATX	100	0	0	3	13	0	20	3	13	17	40	30	73	77	63	83	27	80
HOK	97	0	3	16	25	6	6	3	6	13	50	22	81	81	75	84	31	81
STX	100	17	6	33	33	28	28	0	17	33	61	39	89	89	89	100	61	89
CHTX	100	0	0	10	6	13	13	3	6	16	61	29	71	71	71	74	26	71
PTX	100	6	6	14	14	11	11	0	11	9	40	26	71	74	71	74	34	71
BTX	100	0	0	33	33	20	33	7	33	33	0	33	87	87	87	93	60	87
ANM	100	0	0	10	3	3	0	0	3	3	23	0	73	73	73	77	27	73
WLA	100	0	0	0	0	0	0	0	0	0	57	20	3	7	7	10	3	13
LAR	100	0	0	0	0	0	0	0	0	0	56	11	4	0	0	7	0	11
CMS	100	0	0	0	0	0	3	0	0	3	48	18	9	12	6	9	0	12
YMS	100	3	3	3	0	0	6	0	3	0	72	38	6	13	16	16	3	28
MMO	100	0	0	0	0	6	9	0	6	12	67	36	39	42	39	48	3	42
BTN	100	0	0	0	0	0	0	0	0	0	61	13	3	6	3	6	0	3

Exclusion test was carried out using criterion of Rannala and Mountain (1997) and the simulation method of Paetkau et al. (2004).

there was a severe reduction of heterozygosity in the more Northern locations (ETX, CSTX, WATX) of the South-central region (mean $H_E = 0.188 \pm 0.019$SD), such that it is comparable with that of Western populations (mean $H_E = 0.199 \pm 0.027$SD). There also is evidence of bottlenecks having occurred in these populations. Most of the fixed alleles at several loci were observed in Northern populations. Thus, most of the populations north of Kingsville, TX had generally low levels of genetic diversity, suggesting that they have undergone recent population declines resulting in population bottlenecks. However, larger sample sizes than those available for the BTX, STX, and CSTX populations are needed for robust bottleneck tests (Luikart et al., 1998). These results, along with the migrant exclusion and population assignment tests, phylogenetic reconstructions, and PC analyses suggest that the South-central region is genetically heterogeneous, with the Northern three and Southern three populations comprising two different genetically defined regions. We chose to retain the South-central designation to facilitate comparisons with previous studies, but the genetic differences within the region should be kept in mind when interpreting the results.

The patterns of genetic diversity revealed by microsatellites are in keeping with those derived from analyses of mtDNA PCR-RFLP and RAPD markers applied to the same populations in earlier studies (Kim & Sappington, 2004a, b). Although the different markers have different modes of inheritance, Southern populations show consistently high levels of genetic diversity while Northern populations reveal generally low genetic diversity (Table 6). A south to north decline in boll weevil genetic diversity suggested by studies on allozymes and mtDNA PCR-RFLPs (Terranova, Jones & Bartlett, 1990; Roehrdanz, 2001) are consistent with our findings. For all three types of DNA marker, the highest genetic diversity was observed in the Tampico, Mexico sample, which is the population closest to the geographic origin of this species. In contrast, low levels of diversity characterize populations in the more recently colonized areas of the central U.S. Cotton Belt. This pattern is consistent with the known history of boll weevil range expansion into the United States from Mexico.

Table 5. Number of probable, $\alpha = 0.05$ ($\alpha = 0.01$ in parentheses), first generation migrants (FGM) identified in each population of boll weevils, and the number of those migrants assigned to the most likely population of origin

Pop.	No. FGM	Most likely population of origin																	
		MEX	WTX	KTX	ETX	CSTX	WATX	HOK	STX	CHTX	PTX	BTX	ANM	WLA	LAR	CMS	YMS	MMO	BTN
MEX	1(1)	–	1(1)																
WTX	3(2)	1(1)	–	1	1(1)														
KTX	4(0)		3									1							
ETX	3(2)			1(1)	–	1							1(1)						
CSTX	1(1)					–			1(1)										
WATX	3(0)						–				1		1				1		
HOK	4(2)							–	1	1(1)			1(1)					1	
STX	2(0)								–	1			1						
CHTX	2(1)					1(1)				–			1						
PTX	5(1)				1			1(1)			–	1							2
BTX	1(0)											–			1				
ANM	2(0)											2	–						
WLA	2(1)				1(1)									–			1		
LAR	1(0)														–		1		
CMS	2(0)										2					–			
YMS	1(0)																–	1	
MMO	3(0)					1											2	–	
BTN	1(1)																1(1)		–

Using assignment criterion of Rannala and Mountain (1997) and the Monte Carlo resampling method of Paetkau et al. (2004).

Genetic structuring within and between regions

Analyses of the 11 microsatellite loci revealed a multi-locus F_{ST} of 0.241 (Table 1), and thus significant genetic differentiation among populations of boll weevils in the central U.S. Cotton Belt. Mean within- and between-region genetic differentiation estimates were summarized and compared with those derived from mtDNA PCR-RFLP and RAPD data (Table 6). Results from all

Table 6. Mean ($\pm$SD) genetic diversity estimates [expected heterozygosity (H_E) or haplotype diversity] within regions, and genetic differentiation estimates (F_{ST} and ϕ_{ST}) within and among regions, as determined by three different types of genetic markers from the same collections of boll weevils

Region	Microsatellites		mtDNA PCR-RFLP[a]		RAPD[b]	
	H_E[c]	F_{ST}[d]	Haplotype diversity	ϕ_{ST}	H_E	F_{ST}[d]
South-central	0.360 (0.195)	0.180 (0.166)	0.703 (0.077)	0.096 (0.085)	0.184 (0.068)	0.164 (0.061)
Western	0.199 (0.027)	0.037 (0.034)	0.377 (0.216)	0.230 (0.339)	0.140 (0.042)	0.192 (0.087)
Eastern	0.099 (0.028)	0.022 (0.028)	0.655 (0.096)	0.055 (0.059)	0.088 (0.022)	0.110 (0.053)
South-central vs. Western	–	0.151 (0.156)	–	0.212 (0.189)	–	0.206 (0.081)
South-central vs. Eastern	–	0.240 (0.164)	–	0.165 (0.093)	–	0.203 (0.087)
Western vs. Eastern	–	0.152 (0.092)	–	0.375 (0.141)	–	0.257 (0.085)
All populations	0.219 (0.154)	0.151 (0.146)	0.578 (0.200)	0.214 (0.192)	0.137 (0.060)	0.202 (0.089)

[a] Data from Kim and Sappington (2004a).
[b] Data from Kim and Sappington (2004b).
[c] Estimated as expected heterozygosity in microsatellites and RAPD data and haplotype diversity in PCR-RFLP data.
[d] Genetic differentiation calculated using Weir and Cockerham's equation (1984).

three DNA markers indicate that boll weevil populations are genetically structured as a whole across the geographic range sampled. Genetic differentiation estimates from all pairwise regional comparisons revealed distinct genetic structuring among regions, despite some marker-dependent differences in trends. This conclusion also is supported by the isolation-by-distance patterns revealed by all three types of markers.

However, there are some discrepancies in the estimates of genetic differentiation, depending on the DNA marker. For example, the microsatellite data reveal a higher level of genetic differentiation between the South-central and Eastern regions than in other pairwise comparisons. In contrast, both mtDNA PCR-RFLP and RAPD data indicated profound genetic differentiation between Western and Eastern regions. Furthermore, in comparisons of mean pairwise F_{ST}'s within regions, microsatellite data indicate the highest level of genetic differentiation is among populations from the South-central region, whereas there is little genetic differentiation among populations within the Western and Eastern regions. However, data from the other two DNA markers indicate that the highest level of genetic differentiation is among populations of the Western region.

These differences in patterns of genetic differentiation are most likely due to the inheritance modes of these markers. Because mtDNA is maternally inherited without recombination, the entire mtDNA genome behaves as the equivalent of a single locus with multiple alleles (Dowling, Moritz & Palmer, 1990). The mean time of fixation or loss of new mutations is approximately twice as fast for mitochondrial genes than for nuclear genes, so mtDNA is highly sensitive to phenomena such as genetic drift, bottleneck events, and founder effects (Birky, Maruyama & Fuerst, 1983). RAPD markers have a dominant mode of inheritance, in that dominant homozygotes and heterozygotes are indistinguishable. Therefore, allele frequencies cannot be calculated directly, and indicators of population genetic structure such as *F*-statistics can be described only under the assumption of Hardy–Weinberg equilibrium. Thus, both mtDNA-RFLP and RAPD markers may provide insufficient information in evaluating the genetic structure of populations. On the other hand, microsatellites are inherited in a Mendelian fashion, provided no null alleles are present. All microsatellites used in the present study were prescreened via family analyses to ensure they exhibit Mendelian inheritance (Kim & Sappington, 2004c).

Phylogenetic reconstructions based on the various markers also are consistent with the range expansion from south to north, suggesting that migration and genetic drift have not yet reached equilibrium (Felsenstein, 1982; Neigel, 1997; Avise, 2000). The migrant exclusion and assignment tests based on microsatellite data suggest that while the Tampico, Weslaco, and Kingsville populations frequently exchange migrants, they are relatively isolated from the other populations further north. With prevailing winds in deep south Texas being from the south and southeast, there may be a continuing south to north bias in boll weevil dispersal in this region, slowing the approach to genetic equilibrium in the U.S. The Eastern locations in our studies were first colonized by the boll weevil about 50 years before the Western locations were successfully colonized (Hunter & Coad, 1923; Bottrell, Rummel. & Adkisson, 1972), most likely from Southern populations (Burke et al., 1986). The more recent the colonization of an area, the more likely it will genetically resemble the source population, perhaps accounting for the lower genetic differentiation between the South-central and Western populations than between the South-central and Eastern populations revealed by the microsatellite data.

Gene differentiation and genetic relationships among weevil populations

Examination of genetic differentiation among all population pairs as revealed by microsatellites provides further insight into the genetic structuring of boll weevil populations. Results indicate that the southernmost populations (MEX, WTX, KTX) were substantially differentiated genetically from most of the other populations sampled. On the other hand, populations (ETX, CSTX, WATX) from more Northern areas of the South-central region showed little genetic differentiation relative to most other populations in the Western and Eastern regions. Within these latter two regions, very little genetic differentiation was observed between most population pairs. Conversely, significant genetic structuring was indicated in

most pairwise comparisons between Western and Eastern populations. This finding is consistent with the results from mtDNA PCR-RFLP and RAPD data (Kim & Sappington, 2004a, b), which suggest very little genetic exchange between the Western and Eastern regions. Results of the exclusion test based on microsatellite data indicate that what little migrant exchange occurs between these two regions is in the direction of west to east, the direction of prevailing winds in this part of the continent.

These patterns are strikingly illustrated by the phylogenetic and PC analyses (Figures 3 and 4). In both NJ and UPGMA trees, all populations north of Kingsville clustered together with strong bootstrap support (NJ: 100%; UPGMA: 99%) and short branch lengths, while the southern-most populations appear basally in the NJ tree when rooted with the MEX population. In addition, separation of Eastern populations into an independent clade is conspicuous, and these Eastern populations share a common ancestor with CSTX in the NJ tree. PC analysis also reveals a clear separation of the southernmost populations (MEX, WTX and KTX) from all other populations in multivariate space, as well as distinct, clustered localizations of the Eastern and Western populations. These findings strongly suggest that the recent range expansions into the southeastern U.S. and later into the High Plains both derived mainly from populations in central Texas, rather than from independent long-range migration events from farther south.

Dispersal tendencies and the potential of gene flow of boll weevil populations

One can obtain insight into dispersal tendencies of natural populations as well as their genetic equilibrium status by examining isolation by distance relationships. Slatkin (1993) suggested that a pattern of isolation by distance should be detectable when a population is at or near equilibrium under its current patterns of dispersal. The absence of an isolation by distance pattern suggests that the population is far from equilibrium, and that genetic structuring may reflect a recent range expansion rather than current levels of gene flow.

Isolation by distance patterns derived from boll weevil microsatellite (Figure 2), mtDNA PCR-RFLP (Kim & Sappington, 2004a), and RAPD (Kim & Sappington, 2004b) markers are in good agreement. Although there is a significant positive correlation between genetic distance and geographic distance across all populations sampled, this is entirely due, in the case of microsatellite and RAPD markers, to a strong isolation by distance pattern among populations in the South-central region. Thus, the South-central populations may be close to equilibrium under current patterns of dispersal. On the other hand, the lack of correlation between genetic and geographic distance in both the Eastern and Western regions suggests that boll weevil populations there have not yet reached equilibrium after the colonizing events of a century and of a half-century ago, respectively. Therefore, considering the underlying theoretical assumptions for indirect estimates of gene flow from F_{ST} (Slatkin & Barton, 1989; Whitlock & McCauley, 1999), it is likely that the best estimates of gene flow among boll weevil populations are to be derived from the South-central region in which the populations show a strong isolation by distance relationship, as well as a relatively longer history of residence compared with the Western and Eastern populations.

Theoretically, gene flow is related to genetic differentiation by the equation $F_{ST} = 1/(4N_m + 1)$, where N_m is the number of effective migrants per generation (Wright, 1969). This relationship is imprecise but qualitatively correct (Crochet et al., 2003). Pairwise estimates of N_m for boll weevil populations based on microsatellite data were generally intermediate to those derived from mtDNA-RFLP and RAPD data. In the case of mtDNA-RFLP data, estimates of N_m's above 100 were common (Kim & Sappington, 2004a), whereas the highest N_m estimated from the RAPD data was only 6.7 for CMS-WLA, locations separated by 200 km (Kim & Sappington, 2004b). Results of microsatellite analysis indicate high levels of gene flow or panmixus between most pairs of Western populations, as well as most pairs of Eastern populations, but moderate or limited gene flow between populations across the three designated regions.

To obtain insight into boll weevil dispersal distances, the relationship between N_m and geographic distance was assessed for the South-central populations, the region with the strongest isolation by distance pattern. N_m, as calculated from both microsatellite and RAPD data, was significantly

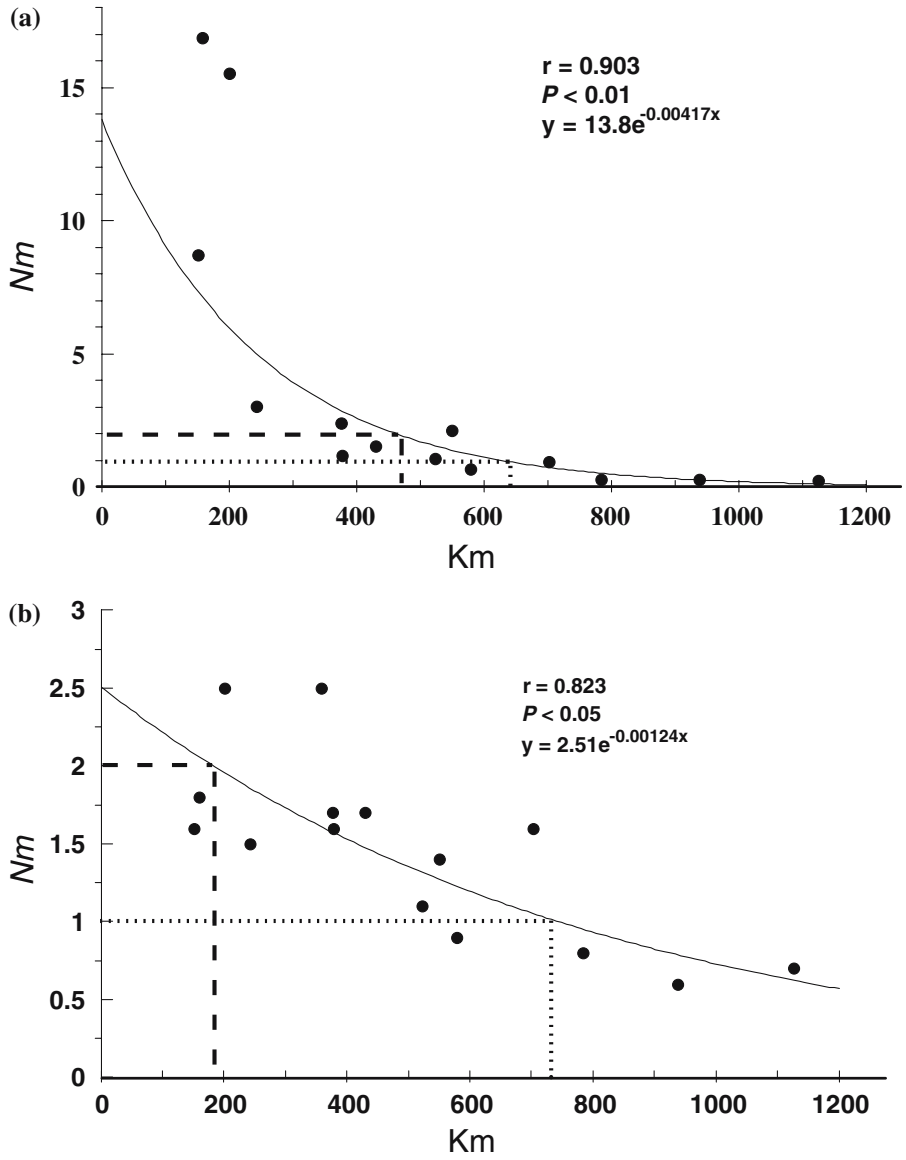

Figure 5. Relationship of gene flow estimates (*Nm*; number of effective migrants per generation) to geographic distance among boll weevil populations from the South-central region obtained from two different DNA markers. (a) Microsatellites. (b) RAPD. N_m values were calculated from Weir and Cockerham's F_{ST} (1984). RAPD data from Kim and Sappington (2004b).

correlated to geographic distance (Figure 5). Although *Nm* calculated from the mtDNA-RFLP data also was significantly correlated with geographic distance ($y = 2.35e^{-0.00157x}$; $r = 0.559$; $p < 0.05$), it is not presented because it describes female migrants only. Geographic distance between populations accounted for about 80 and 70% of the variation in N_m calculated from microsatellite and RAPD data, respectively. It has been suggested that when $N_m \geq 1$, a population is not likely to diverge genetically from the source population (Wright, 1931; Maruyama, 1970). In other words, roughly one individual exchanged between populations per generation is sufficient to prevent genetic differentiation. The regression equations derived from microsatellite (Figure 5(a)) and RAPD (Figure 5(b)) data indicate that one boll weevil can be expected to move between

locations separated by 640–740 km, respectively, and reproduce per generation. However, the curves are fairly flat, so two effective boll weevil migrants are to be expected only at distances of ~180 km (RAPD) – 470 km (microsatellites). Thus, the data from our various studies suggest that although boll weevils are capable of long-range movement over several 100's of km, most individuals disperse much shorter distances. More precise estimation of dispersal over short distances awaits ongoing analyses of data from populations separated by distances less than 100 km.

Management implications

The database of microsatellite genotypes generated by this study now makes it possible, through population assignment techniques, to identify the most likely geographic source of a boll weevil reintroduced to an eradicated zone. Population assignment of a boll weevil can help action agencies decide the most appropriate mitigation response to a reintroduction. For example, five weevils from the Plainview, Texas sample in this study were identified as probable first generation immigrants (Table 5). Two of these likely originated from the Big Spring, Texas and Hobart, Oklahoma areas, both regions near enough that natural flight could easily account for their presence. Three other probable immigrants were assigned to source populations near Brownsville, Tennessee and El Campo, Texas, areas so far distant that human-mediated transport seems the most likely mode of introduction. Boll weevils captured in eradicated zones could be screened against the microsatellite database to determine the most likely area of origin. This database of microsatellite genotype profiles can be extended and updated relatively easily, now that the microsatellite markers have been developed (Kim & Sappington, 2004c). Indeed, we are now genotyping populations of boll weevils collected in several areas of Mexico to aid in identifying the likely source of weevils recently captured in a Mexican eradication zone near Matamoros de la Laguna, Coahuila.

The rate of migrant exchange revealed in this study between Weslaco, Texas, an area just now entering an eradication program, and Kingsville, Texas, indicates that much of the difficulty in achieving final eradication after 9 years of effort in the latter area (Allen et al., 2005) is due to natural immigration of boll weevils from the Lower Rio Grande Valley. Our data also suggest that eventual eradication in the Lower Rio Grande Valley will be expedited if suppression of boll weevil populations in the cotton-growing region north of Tampico, Mexico can be achieved, because Weslaco is apparently receiving migrants from that area.

Acknowledgements

We thank Veronica Cardoza, Justin Gibson, Valentina Greenberg, and Orlando Zamora for excellent technical assistance. We are grateful to all those who collected boll weevils for this study. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

References

Agresti, A., 1990. Categorical Data Analysis. Wiley Interscience, New York.

Allen, C.T., L.W. Patton, L.E. Smith & R.O. Newman, 2005. Texas boll weevil eradication report. Proceedings of the Beltwide Cotton Conferences 2005, 1196–1205 .

Analytical Software, 1998. Statistix for Windows: User's Manual. Analytical Software, Tallahassee, FL.

Arzaluz, I.O. & R.W. Jones, 2001. Ecology and phenology of the boll weevil (Coleoptera: Curculionidae) on an unusual wild host, *Hibiscus pernambucensis*, in southeastern Mexico. J. Econ. Entomol. 94: 1405–1412.

Avise, J.C., 2000. Stability, equilibrium and molecular aspects of conservation in marine species, pp. 11–12 in Marine Genetics edited by A.M. Solé-Cava, C.A.M. Russo & J.P. Thorpe. Kluwer Academic Publishers, Norwell, MA.

Bartlett, A.C., W.C. Randall & J.E. May, 1983. Allozyme variation among populations of boll weevils in Arizona and Mexico. Southwest. Entomol. 8: 118–130.

Batley, J., K.J. Edwards, J.H.A. Barker, K.J. Dawson, C.W. Wiltshire, D.M. Glen & A. Karp, 2004. Population structure of the beetle pests *Phyllodecta vulgatissima* and *P. vitellinae* on UK willow plantations. Insect Mol. Biol. 13: 413–421.

Belkhir, K., P. Borsa, J. Goudet, L. Chikhi & F. Bonhomme, 2000. GENETIX 4.02, logiciel sous WindowsTM pour la genetique des populations. CNRS UPR 9060, Universite de Montpellier II, Montpellier (France), http://www.univ-montp2.fr/~genetix/genetix/genetix.htm.

Biggers, C.J., B.R. Jones, H.R. Bancroft & L.A. Dickey, 1996. Genetics of esterases in laboratory and feral cotton boll

weevils, *Anthonomus grandis* Boheman. Ecol. Agric. Pests 53: 469–478.

Birky, C.W., T. Maruyama & P. Fuerst, 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. Genetics 103: 513–527.

Black IV, W.C., 1993. PCR with arbitrary primers: approach with care. Insect Mol. Biol. 2: 1–6.

Bottrell, D.G., D.R. Rummel & P.L. Adkisson, 1972. Spread of the boll weevil into the High Plains of Texas. Environ. Entomol. 1: 136–140.

Burke, H.R., W.E. Clark, J.R. Cate & P.A. Fryxell, 1986. Origin and dispersal of the boll weevil. Bull. Entomol. Soc. Am. 32: 228–238.

Cornuet, J.M. & G. Luikart, 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001–2014.

Cornuet, J.M., S. Piry, G. Luikart, A. Estoup & M. Solignac, 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153: 1989–2000.

Crochet, P.A., J.Z. Chen, J.M. Pons, J.-D. Lebreton, P.D.N. Hebert & F. Bonhomme, 2003. Genetic differentiation at nuclear and mitochondrial loci among large white-headed gulls: sex-biased interspecific gene flow? Evolution 57: 2865–2878.

Culin, J., S. Brown, J. Rogers, D. Scarborough, A. Swift, B. Cotterill & J. Kovach, 1990. A simulation model examining boll weevil dispersal: historical and current situations. Environ. Entomol. 19: 195–208.

Daniel, W.W., 1990. Applied Nonparametric Statistics. PWS-KENT Publishing Company, Boston.

Di Rienzo, A., A.C. Peterson, J.C. Garza, A.M. Valdes, M. Slatkin & N.B. Freimer, 1994. Mutational processes of simple-sequence repeat loci in human populations. Proc. Nat. Acad. Sci. USA 91: 3166–3170.

Dowling, T.E., C. Moritz & J.D. Palmer, 1990. Nucleic acids II: Restriction site analysis, pp. 250–317 in Molecular Systematics edited by D.M. Hillis & C. Moritz. Sinauer Associates, Sunderland, MA.

Felsenstein, J., 1982. How can we infer geography and history from gene frequencies? J. Theor. Biol. 96: 9–20.

Fryxell, P.A. & M.J. Lukefahr, 1967. *Hampea* Schlecht: possible primary host of the boll weevil. Science 155: 1568–1569.

Goudet, J., 1995. FSTAT, Version 1.2, a computer program to calculate *F*-statistics. J. Hered. 86: 485–486.

Grefenstette B. & O. El-Lissy, 2003. Boll weevil eradication update. Proceedings of the Beltwide Cotton Conferences 2003: 131–141.

Guerra, A.A., 1988. Seasonal boll weevil movement between northeastern Mexico and the Rio Grande Valley of Texas, USA. Southwest. Entomol. 13: 261–271.

Guo, S.W. & E.A. Thompson, 1992. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48: 361–372.

Haig, S.M., 1998. Molecular contributions to conservation. Ecology 79: 413–425.

Harley, E.H., 2001. AGARst. A Programme for Calculating Allele Frequencies, G_{ST} and R_{ST} from Microsatellite Data, Version 2. University of Cape Town, Cape Town.

Hunter, W.D. & B.R. Coad, 1923. The boll-weevil problem. USDA, Farmers' Bulletin 1329. (A1.9, 1329).

Johnson, W.L., W.H. Cross, J.E. Leggett, W.L. McGovern, H.C. Mitchell & E.B. Mitchell, 1975. Dispersal of marked boll weevil: 1970-1973 studies. Ann. Entomol. Soc. Am. 68: 1018–1022.

Jones, R.W., J.R. Cate, E.M. Hernandez & R.T. Navarro, 1992. Hosts and seasonal activity of the boll weevil (Coleoptera: Curculionidae) in tropical and subtropical habitats of northeastern Mexico. J. Econ. Entomol. 85: 74–82.

Kim, K.S. & T.W. Sappington, 2004a. Boll weevil (*Anthonomus grandis* Boheman) dispersal in the southern United States: evidence from mitochondrial DNA variation. Environ. Entomol. 33: 457–470.

Kim, K.S. & T.W. Sappington, 2004b. Genetic structuring of boll weevil populations in the U.S. based on RAPD markers. Insect Mol. Biol. 13: 292–303.

Kim, K.S. & T.W. Sappington, 2004c. Isolation and characterization of polymorphic microsatellite loci in the boll weevil, *Anthonomus grandis* Boheman (Coleoptera: Curculionidae). Mol. Ecol. Notes 4: 701–703.

Kim, K.S. & T.W. Sappington, 2005. Genetic structuring of Western corn rootworm (Coleoptera: Chrysomelidae) populations in the U.S. based on microsatellite loci analysis. Environ. Entomol. 34: 494–503.

Kimura, M. & T. Ohta, 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. Proc. Nat. Acad. Sci. USA 75: 2868–2872.

Kiser, D. & M. Catanach, 2005. Boll weevil eradication update – Arkansas, 2004. Proceedings of the Beltwide Cotton Conferences, 2005, pp. 1074–1090.

Lowe, A., S. Harris & P. Ashton, 2004. Ecological Genetics: Design, Analysis, and Application. Blackwell Science Ltd, Malden, MA.

Luikart, G. & J.M. Cornuet, 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conserv. Biol. 12: 228–237.

Luikart, G., F.W. Allendorf, J.M. Cornuet & B. Sherwin, 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. J. Hered. 89: 238–247.

Lukefahr, M.J., S. Barbosa & R.B. Sabrinho, 1994. The introduction and spread of the boll weevil, *Anthonomus grandis* Boh. (Coleoptera) in Brazil. Southwest. Entomol. 19: 414–417.

Maruyama, T., 1970. Analysis of population structure. I. One-dimensional stepping-stone models of finite length. Ann. Human Genet. 34: 201–219.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Nei, M., F. Tajima & Y. Tateno, 1983. Accuracy of estimated phylogenetic trees from molecular data. J. Mol. Evol. 19: 153–170.

Neigel, J.E., 1997. A comparison of alternative strategies for estimating gene flow from genetic markers. Ann. Rev. Ecol. Syst. 28: 105–128.

Ota, T., 1993. DISPAN, Genetic Distance and Phylogenetic Analysis. Institute of Molecular Evolutionary Genetics, The Pennsylvania State University.

Paetkau, D., R. Slade, M. Burden & A. Estoup, 2004. Genetic assignment methods for the direct, real-time estimation of

migration rate: a simulation-based exploration of accuracy and power. Mol. Ecol. 13: 55–65.

Parker, P., A. Snow, M. Schug, G. Booton & P. Fuerst, 1998. What molecules can tell us about populations: choosing and using a molecular marker. Ecology 79: 361–382.

Piry, S., A. Alapetite, J.M. Cornuet, D. Paetkau, L. Baudouin & A. Estoup, 2004. GeneClass2: a Software for Genetic Assignment and First-Generation Migrant Detection. J. Hered. 95: 536–539.

Rannala, B. & J.L. Mountain, 1997. Detecting immigration by using multilocus genotypes. Proc. Nat. Acad. Sci. USA 94: 9197–9201.

Raulston, J.R., T.J. Henneberry, J.E. Leggett, D.N. Byrne, E. Grafton-Cardwell & T.F. Leigh, 1996. Short- and long-range movement of insects and mites, pp. 143–162 in Cotton Insects and Mites: Characterization and Management. Cotton Foundation Reference Book Series, No. 3 edited by E.G. King, J.R. Phillips &R.J. Coleman. The Cotton Foundation Publisher, Memphis, TN.

Raymond, M. & F. Rousset, 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J.Hered. 86: 248–249.

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

Roderick, G.K., 1996. Geographic structure of insect populations: gene flow, phylogeography, and their uses. Ann. Rev. Entomol. 41: 325–352.

Roehrdanz, R.L., 1995. Amplification of complete insect mitochondrial genome in two easy pieces. Insect Mol. Biol. 4: 169–172.

Roehrdanz, R.L., 2001. Genetic differentiation of Southeastern boll weevil and Thurberia weevil populations of *Anthonomus grandis* (Coleoptera: Curculionidae) using mitochondrial DNA. Ann. Entomol. Soc. Am. 94: 928–935.

Roehrdanz, R.L. & M.E. Degrugillier, 1998. Long sections of mitochondrial DNA amplified from fourteen orders of insects using conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 91: 771–778.

Roehrdanz, R.L. & D.T. North, 1992. Mitochondrial DNA restriction fragment variation and biosystematics of the boll weevil, *Anthonomus grandis* . Southwest. Entomol. 17: 101–108.

Rohlf, F.J., 1992. NTSYS-PC Numerical Taxonomy and Multivariate Analysis System, Version 1.70. Exeter Publishers, Setauket, NY.

Rousset, F., 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. Genetics 142: 1357–1362.

Saitou, N. & M. Nei, 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406–425.

Scataglini, M.A., V.A. Confalonieri & A.A. Lanteri, 2000. Dispersal of the cotton boll weevil (Coleoptera: Curculionidae) in South America: evidence of RAPD analysis. Genetica 108: 127–136.

Slatkin, M., 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47: 264–279.

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

Smith, J.W., 1998. Boll weevil eradication: area-wide pest management. Ann. Entomol. Soc. Am. 91: 239–247.

Sneath, P.H.A. & R.R. Sokal, 1973. Numerical Taxonomy. W.H. Freedman and Co, San Francisco.

Spurgeon, D.W., J.R. Raulston, O. Zamora & J. Loera, 1997. Spatial and temporal patterns of boll weevil trap captures in northeastern Mexico. Proceedings of the Beltwide Cotton Conferences 1997, pp. 984–986.

Stavinoha, K.D. & L.A. Woodward, 2001. Texas boll weevil history, pp. 451–502 in Boll Weevil Eradication in the United States through 1999 edited by W.A. Dickerson, A.L. Brashear, J.T. Brumley, F.L. Carter, W.J. Grefenstette & F.A. Harris. The Cotton Foundation Publisher, Memphis, TN.

Takezaki, N. & M. Nei, 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics 144: 389–399.

Terranova, A.C., R.G. Jones & A.C. Bartlett, 1990. The southeastern boll weevil: an allozyme characterization of its population structure. Southwest. Entomol. 15: 481–496.

Warner, R.E. & C.E. Smith Jr, 1968. Boll weevil found in Pre-Columbian cotton from Mexico. Science 162: 911–912.

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

Whitlock, M.C. & D.E. McCauley, 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4N_m+1)$. Heredity 82: 117–125.

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

Wright, S., 1969. Evolution and the Genetics of Populations, Vol. 2. University of Chicago Press, Chicago.