2003 BELTWIDE COTTON CONFERENCES, NASHVILLE, TN – JANUARY 6-10

DNA FINGERPRINTING BOLL WEEVIL POPULATIONS FROM NON-ERADICATED STATES AND NORTHEAST MEXICO

Kyung Seok Kim and Thomas W. Sappington
USDA-ARS, IFNRRU
Weslaco, TX

Abstract

An understanding of boll weevil (*Anthonomus grandis*, Boheman) dispersal behavior is essential to characterizing and responding to the threat of migration into eradicated zones. Variation in boll weevil mitochondrial DNA (mtDNA) was sampled and analyzed to make inferences on the magnitude and geographic pattern of genetic differentiation among weevil populations. PCR-RFLP analysis was conducted on a large fragment of mtDNA from each of 419 individuals from 20 locations across northeast Mexico and eight US states. A total of 28 distinct mtDNA haplotypes, 17 of which were unique to single locations, were identified from restriction reactions of ten informative endonucleases. The value of within-location haplotype diversity varied from 0 to 0.81, and nucleotide diversity ranged from 0 to 0.36%. Nucleotide sequence divergence among weevil populations ranged from -0.01 to 0.68% with a mean value of 0.13%. Haplotype and nucleotide diversity was generally greater in eastern than western populations, and haplotype frequencies differed greatly in these two regions. Phylogenetic reconstruction of populations revealed two major clades corresponding to eastern and western regions, and is consistent with historical boll weevil range expansion into the southeastern U.S. from Mexico, and a secondary colonization of the High Plains. Evidence suggests that gene flow between eastern and western populations is limited. However, it appears that migration between populations separated by < 200 km is frequent.

Introduction

The boll weevil (*Anthonomus grandis*) first entered the U.S. from Mexico in 1892 through the southern tip of Texas (Burke et al. 1986). Its ability to disperse is evident from the history of its subsequent range expansion in the U.S. of 64-193 km per year (Hunter and Coad 1923). Trapping data indicated that the spread of the boll weevil through previously uninfested areas of southern Brazil could occur at the rate of 97 km in a 3-day period, and 160 km in a 9-day period (Lukefahr et al. 1994). Weevil flight activity increases late in the growing season when cotton begins to senesce and the insects seek overwintering sites or more favorable conditions for feeding and ovipositing (Guerra 1986). Marked individuals have been recaptured 105-272 km from the point of release (Guerra 1988, Raulston et al. 1996). However, the frequency, impact, and geographic patterns of such long-range movement are still unknown. Potential reinfestation of eradicated areas by long-range movement (>100 km) of boll weevils is of great concern to eradication programs. Although circumstantial evidence often implicates migrants as the source of unexpected increases in pheromone trap captures in eradication zones, such evidence is seldom unequivocal, and the source of migrants is even less certain.

The boll weevil has been eradicated from several states in the southeastern and far western U.S., but many regions are still infested (El-Lissy and Grefenstette 2002). Movement of weevils from infested areas into eradicated or nearly eradicated zones can occur naturally through flight (Allen et al. 2001), or inadvertantly through human-mediated transport (Jones and Wilson 2002). The economic consequences of a reinfestation are great, so detection and prevention of such events is a high priority. When a reintroduction does occur, a question of paramount importance is the source of the weevils, but deduction by circumstantial evidence is the only method currently available for guessing the origin of a reintroduction. We have begun research to determine the usefulness of DNA markers in providing better information on the distance and magnitude of boll weevil dispersal, and in identifying potential sources of migrants captured in eradicated areas. A better understanding of boll weevil dispersal patterns and population interchange will permit scientists and action agencies to develop realistic and more effective strategies for monitoring and responding to boll weevil reintroductions to eradication and post-eradication zones.

Mitochondrial DNA (mtDNA) analysis has proven particularly useful in studying genetic variation within species as well as differences among closely related species. The properties of mtDNA have been reviewed in detail by Avise et al. (1987). We surveyed genetic variation among widely separated populations of boll weevils by restriction fragment length polymorphism (RFLP) analyses of a long polymerase chain reaction (PCR) product of mtDNA. Our goal was to characterize the magnitude and geographic patterns of genetic differentiation in boll weevils from populations in the south central Cotton Belt of the U.S. and northeast Mexico. The results of this study shed light on boll weevil gene flow between populations, and form the foundation necessary to proceed with more thorough DNA fingerprinting of weevil populations.

Materials and Methods

Sample Collection

Boll weevils were collected by pheromone trapping from 20 locations in NM, TX, OK, MO, AR, LA, TN, and MS in the U.S., and from northeastern Mexico (MX) in the cotton growing region just north of Tampico (Table 1). Weevils were collected in multiple traps at each location. For the phylogeographic study, locations were grouped into three main regions designated as south-central, western, and eastern. Within these three regions, each location was separated by no more than 300 km from its nearest neighbor, except in the case of MEX. The MEX was located about 430 km south of WTX, but for convenience was included in the south-central region.

MtDNA Analysis

Total genomic DNA was extracted from individual boll weevils using Promega's Wizard isolation kit (Promega, WI), according to the manufacturer's protocol. A long fragment of mtDNA was amplified by PCR using the 12S and C1 primers described by Roehrdanz and Degrugillier (1998). Sequences of the primers are 12S, 5'-AAACTAGGATTAG ATACCCTATTAT-3' and C1, 5'-TTGATTTTTTGGTCATCCAGAAGT-3'. This fragment comprises most of the weevil mtDNA except for the AT-rich and ND2 regions. The PCR reaction was performed in a total volume of 50 µl using 15 to 50 ng of genomic DNA, 2.5 mM $MgCl_2$, 400 µM of each dNTP, 0.4 µM of each primer, and 1.5 units of LA Taq polymerase (Panvera, Madison, WI). Amplification was carried out in a GeneAmp PCR system 9700 thermocycler (Perkin Elmer, Norwalk, CT). PCR cycling conditions were as described by Roehrdanz (1995) with slight modifications. The reaction began with a 'cool start' followed by cycling parameters: 94°C 1 min; 15 cycles of 94°C 1 min, 60°C 12 min; 20 cycles of 94°C 1 min, 60°C 12 min with 15 s auto extend; 72°C 10 min; 4°C hold.

Aliquots of 3-8 µl (depending on DNA quantity) of the amplified mtDNA fragment were digested with each restriction enzyme in 96-well microtiter plates using 1.5U enzyme in a total volume of 20 µl. PCR products and resulting restriction fragments were separated electrophoretically in 1.0-3.5% agarose gels (depending on the enzyme) in 1 X TBE buffer (90mM Tris-borate, 2mM EDTA), followed by staining with ethidium bromide (0.2µg/ml). Stained gels were documented with a Chemi Doc imaging system, and restriction fragments were scored with Quantity One Software™ (Bio-Rad Laboratories, Hercules, CA). A Kb DNA ladder (Stratagene, La Jolla, CA) and a 100-bp DNA ladder (Fermentas Inc., Hanover, MD) were used as molecular weight size standards. In a preliminary test, amplified fragments from 60 weevils, which included 3 individuals from each location, were screened for polymorphisms using 28 restriction enzymes (Table 2). Of these, 10 enzymes (*Eco*RI, *Msp*I, *Rsa*I, *Hinf*I, *Taq*I, *Vsp*I, *Bsi*YI, *Dde*I, *Nde*II, and *Hae*III) revealed mtDNA variation between at least two populations, and were therefore selected for further study.

Data Analysis

Differences in banding patterns arose from presence or absence of restriction fragments of certain sizes, and were used to analyze genetic structuring among boll weevil populations. A single letter was used to designate each restriction fragment length profile. A multi-letter code, based on the restriction patterns across all enzymes, was assigned to the composite mtDNA haplotype observed for each weevil. The minimum path network interconnecting the composite haplotypes was constructed by the parsimony approach of Avise et al. (1979). Some hypothetical fragments were assumed to explain all conjectured mutational steps. The mean number of nucleotide substitutions per site (d) between all pairs of haplotypes was calculated based on restriction fragment information (Nei and Li 1979). Haplotype frequency distributions for each population and the associated d values were used to estimate haplotype and nucleotide diversity within populations (Nei 1987). Nucleotide divergence among populations was estimated according to the equation described by Nei and Tajima (1981). Geographic heterogeneity between all pairs of populations was estimated from population frequency distributions using the Monte Carlo simulation (Roff and Bentzen 1989).

All calculations above were computed using programs included in the REAP package (McElroy et al. 1991). The genetic relationships among populations were reconstructed using the FITCH and NEIGHBOR programs from the PHYLIP 3.5c computer package (Felsenstein 1993). The phylogenetic trees were based on nucleotide divergence (estimates of net nucleotide substitutions) between two populations (Nei and Tajima, 1981). The correlation of genetic distance (nucleotide divergence) with geographic distance between populations was calculated from 5000 replications and normalized by the Mantel statistic Z option of the NTSYS-PC program, version 1.70 (Rohlf 1992).

Results

Haplotype Distribution and mtDNA Diversity in Weevils

Of 28 restriction enzymes used in a preliminary screening, only ten enzymes revealed polymorphic banding patterns in weevils (Table 2). The sum of the inferred fragment sizes for the patterns ranged from 12.4 to 12.7 kilobases (kb). The slight variation in the sum of the fragments is probably due to undetectable small fragments (< 100 base pairs) generated by some enzymes, and/or to imprecise measurements of the larger fragments. The number of detectable restriction fragments for a

single enzyme varied from six with *Eco*RI to 33 with *Dde*I, and the number of variable patterns produced by an enzyme ranged from two for *Eco*RI and *Hae*III to eight for *Dde*I. In the survey of all weevil populations, the ten selected restriction enzymes produced 48 unique digestion patterns. These generated 28 distinct haplotypes among the 419 weevils analyzed (Table 3). Eleven haplotypes were found in more than two locations and 17 haplotypes were unique to single locations. The two most common haplotypes (1 and 2) were found in 61.3% of the weevils sampled, and they were widely distributed across weevil populations (Table 3).

Within-location haplotype diversity ranged from 0 to 0.81 and nucleotide diversity ranged from 0% to 0.36%. The number of observed haplotypes within locations varied from one in BTX to eleven in KTX. The highest levels of haplotype diversity were found in populations from Mexico (MEX), southern Texas (WTX, KTX), and some eastern locations (MMO, SMS, YMS, BTN), whereas low levels of diversity characterized populations from western Texas (BTX, STX) and adjacent states (HOK, ANM). In general, the highest levels of haplotype and nucleotide diversity were observed in south-central populations (mean h = 0.70, π = 0.29%), and were greater in eastern populations (mean h = 0.64, π = 0.24%) than in western populations (mean h = 0.38, π = 0.09%).

A parsimony network showing the relationship of the composite haplotypes indicates that most haplotypes are removed from the two most common haplotypes (H1 and H2) by only one or two mutational steps (Fig. 1). However, these two central haplotypes are separated from one another by at least 7 mutations.

There were significant differences in haplotype frequencies among boll weevil populations as a whole (χ^2 = 1103.1, P < 0.001) and in most pairwise comparisons (Table 4 and 5). However, apart from the WLA and BTX populations, which were characterized by unusual haplotype frequencies, there were no significant differences in haplotype frequency distributions among geographically adjacent populations (<200 km) (above diagonal in Table 4). When populations from each region were pooled, the three regions differed significantly. There were also significant differences among populations within each region, but when the monomorphic BTX population was excluded, there was no significant difference among western populations (χ^2 = 37.3, n.s.) (Table 5).

Phylogeographic Relationships Among Populations

Nucleotide divergence among weevil populations ranged from –0.01% (BTN vs. MMO) to 0.68% (BTX vs. WLA) (below diagonal in Table 4). Estimates of nucleotide divergence among the three main geographical regions were much higher than those within regions (Table 5). The highest nucleotide divergence was observed between the western and eastern regions.

A qualitative method of data analysis can provide additional information not revealed by quantitative methods, especially regarding phylogeographic inferences among populations. Geographic distributions of mtDNA genotypes revealed by *Eco*RI, *Msp*I, and *Taq*I are shown in Fig. 2. The geographic patterns produced by each of the other enzymes are similar to one of the above distributions, and can be inferred from Table 3. Genotype variation revealed by digestion with *Hae*III was found in only two individuals from the MEX population. The "B" patterns of *Eco*RI and of *Nde*II were observed in 24 weevils collected from only extreme-south locations (MEX, WTX, KTX). Geographic distribution of genotypes produced by some enzymes (*Msp*I, *Hinf*I, *Bsi*YI, *Dde*I and *Vsp*I) exposed a deep genetic disparity between weevils from western and eastern regions, except in the case of the BTN and MMO populations, which showed intermediate genotypes. However, *Taq* I and *Rsa* I did not reveal any obvious geographical differences among populations. Overall, the genotpyes of south-central populations were more variable than those of other populations, and were intermediate in composition between those of western and eastern populations.

To resolve genetic relationships among populations, two phylogenetic trees (Fitch and Neighbor-joining tree) were reconstructed based on values of nucleotide divergence between populations (Fig. 3). The trees were rooted with the MEX population, based on historical information on the spread of the boll weevil from Mexico into the U.S. in the late 19th century (Burke et al. 1986). The trees have similar topologies, with slight differences in branch length, and reveal two major clades corresponding to the eastern and western regions. The unrooted NJ tree indicated that there are two distinct and linearly connected patterns diverging from the MEX and WTX populations (Fig. 3C).

The relationship between genetic distance (nucleotide divergence) and geographic distance can provide insight on possible long-range dispersal in weevils. Genetic distance is positively correlated with geographic distance between populations, though with a rather scattered topology (r = 0.392, P < 0.001) (Fig. 4).

Discussion

Hypotheses on the origin and subsequent range expansion of the boll weevil through Mexico and the U.S. have been discussed by several authors (Fryxell and Lukefahr, 1967; Burke and Cate 1979; Burke et al. 1986). Most authors have proposed a Meso-American (southern Mexico and northern Central America) origin. This hypothesis is based mainly on evi-

dence from geographic variation of morphological characters, and on host plant associations (Burke et al. 1986). The explosive range expansion of the boll weevil from its native host began in the late 19th century and was made possible through the availability of cultivated cotton as a host (Burke et al. 1986). Burke et al. (1986) hypothesized two possible routes of northward dispersal of the boll weevil. The first involved dispersal up the east coast of Mexico and into the southeastern United States. The second required a crossing of the central highlands to northwestern Mexico and into southern Arizona from the Pacific coast. The initial dispersal of the boll weevil across the Cotton Belt of the southeastern United States was rapid, with range expansion averaging 95 km per year (Culin et al. 1990). A secondary range expansion into the High Plains of Texas and New Mexico occurred much later, beginning in the late 1950s (Bottrell et al. 1972).

Our data revealed high levels of haplotype and nucleotide diversities in southern populations of boll weevils, with lower levels in the more northerly populations (Table 3). This is the pattern one would expect if there are lingering genetic founder effects from the recent colonization. Phylogenetic trees revealed two major clades corresponding to the eastern and western regions, which represent the two historical range expansions into the southeastern Cotton Belt and into the High Plains, respectively (Hunter and Coad 1923, Bottrell et al. 1972). Populations from both the eastern and western regions are apparently derived from Mexico and Weslaco ancestral populations (Fig. 3). Evidence from the geographic distributions of mtDNA genotypes, as well as from nucleotide divergence values in pair-wise comparisons of the three regions, indicate that southern populations are genetically intermediate between western and eastern populations (Fig. 2, Table 5). Overall, these findings are consistent with the initial pattern of range expansion observed when the boll weevil first entered the U.S. from Mexico (Burke et al. 1986). Eastern populations (MO, LA, TN, MS, AR) showed generally higher mtDNA diversity than western populations (northwestern TX, OK, NM). Successful colonization in the latter areas has occurred only in more recent decades (Bottrell et al. 1972).

Of the 28 distinct mtDNA haplotypes observed in the weevils sampled, 17 were unique to single locations, and three were found in only southern populations (haplotypes 5, 7, and 8). In addition, the "B" patterns of *Eco*RI and *Nde*II also were found only in southern locations (MEX, WTX, KTX). Thus, these haplotypes and region-specific patterns of *Eco*R I and *Nde*II can be considered diagnostic genotypes, or genetic fingerprints, for southern weevils.

Our data reveal that haplotypes 1 and 2 are the most common among boll weevils in the U.S. and northeastern Mexico, accounting for 61.3% of the total haplotypes determined (257 out of 417 samples) (Table 3). In a parsimony network (Fig. 1), they occupy a central position in the two major clusters and connect separately to all the other haplotypes. Haplotype 1 is the most widespread geographically but haplotype 2 is found almost exclusively in southern and eastern populations. Therefore, although haplotype 2 cannot be excluded as an ancestral type, haplotype 1 is the most likely ancestral mtDNA genotype, or is very close to the original colonizing boll weevil haplotype. However, data from larger samples will be needed to more rigorously address this question.

Geographic patterns of mtDNA variation observed in a number of animals were described and categorized by Avise et al. (1987). Our data permit us to consider assignment of these categories in the special case of boll weevils, an animal with a known history of recent dispersal into the southeastern United States. The presumed ancestral haplotype (1) occurs over a broad area, whereas most of the haplotypes we identified were found within single locations or a cluster of adjacent locations. The patterns exhibited by some haplotypes (2, 3, 4, 6) suggest limited gene flow between western and eastern population, which do not appear to be isolated by long-term zoogeographic barriers to dispersal. The results of a Monte Carlo simulation indicate a significantly different geographic distribution of the haplotypes (Table 4, 5). Furthermore, we found a positive relationship between genetic distance and geographic distance (Fig. 4). Therefore, the relationship between boll weevil mtDNA haplotypes and geography corresponds to examples of phylogeographic categories III and V as described by Avise et al. (1987). Both categories are characterized by a continuous genetic divergence pattern, but Category III patterns are typified by limited gene flow, and Category V patterns by intermediate gene flow. In practice, it is not easy to distinguish clearly between categories III and V. This is the case with our data, probably because boll weevils have a recent history of colonization, with likely repeated genetic bottlenecks occurring at regional scales for several reasons, including human activity. Therefore, it is not surprising that we observe rather complex population structures.

The populations from Big Spring, TX (BTX) and Winsboro, LA (WLA) displayed unusual haplotype frequencies within their respective regions. MtDNA variation was less in these populations than in adjacent locations (Table 3), especially in the case of BTX which was monomorphic for haplotype 3 (Table 3). Except for BTX, western populations did not differ significantly in their haplotype frequency distributions (Table 5). Mitochondrial DNA is highly sensitive to phenomena such as genetic drift, bottleneck events and founder effects, because the loss of mtDNA variation occurs roughly four times faster than that of nuclear DNA (one-fourth of the effective population number of nuclear genes) (Birky et al. 1983). Therefore, it is likely that founder or bottleneck events have contributed to the loss of mtDNA variation in these populations. Overall, the geographic pattern of haplotype frequencies, and the presence of numerous unique haplotypes suggest that gene flow between eastern and western populations is limited. However, recent gene flow analyses that we have conducted (data not shown) indicates that migration between populations separated by < 200 km is frequent. These analyses will be reported in a forthcoming paper.

In conclusion, we found that mtDNA PCR-RFLP analysis is a fast and technically unambiguous tool, providing a very powerful approach for the study of dispersal and population structure of boll weevil. In particular, the use of a long fragment of the mtDNA for RFLP analysis made it possible to identify enough polymorphisms to analyze population genetic structuring in an animal like the boll weevil with a recent colonization history.

Acknowledgements

We thank Lisa Saenz, Veronica Cardoza, Valentina Greenberg, and Orlando Zamora for their technical assistance. We are grateful to all the cooperators who sent us weevils from across the Cotton Belt. 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

Allen, C. T., L. W. Patton, L. E. Smith, and R. E. Newman. 2001. Texas boll weevil eradication update. Proc. Beltwide Cotton Conf. Pp. 934-937.

Avise, J. C., R. A. Lansman, and R. O. Shade. 1979. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I. Population structure and evolution in the genus Peromyscus. Genetics 92: 279-295.

Avise, J. C., J. Amond, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18: 489-522.

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

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

Burke, H. R. and J. R. Cate. 1979. A new species of Mexican *Anthonomus* related to the boll weevil (Coleoptera: Curculionidae). Ann. Entomol. Soc. Am. 72: 189-192.

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

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

El-Lissy, O. and B. Grefenstette. 2002. Boll weevil eradication in the U.S., 2001. Proc. Beltwide Cotton Conf. (CD), 10 pp.

Felsenstein J. 1993. PHYLIP-phylogenetic inference package, version 3.57c. University of Washington, Seattle, WA.

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

Guerra, A. A. 1986. Boll weevil movement: dispersal during and after the cotton season in the Lower Rio Grande Valley of Texas. Southwest. Entomol. 11: 10-16.

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

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

Jones, R. G. and J. A. Wilson. 2002. Boll weevil: post eradication outbreaks in cotton in the southeastern United States. Beltwide Cotton Conferences (CD) 3 pp.

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

McElroy, D., P. Moran, E. Bermingham, and J. Kornfield. 1991. REAP. The Restriction Enzyme Analysis Package. Version 4.0. University of Maine, Orono, ME, U.S.A.

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

Nei, M., and W. H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Nat. Acad. Sci. USA. 76: 5269-5273.

Nei, M. and F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97: 145-163.

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

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

Roehrdanz, R. L. and 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.

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

Roff, D. A. and P. Bentzen. 1989. The statistical analysis of mitochondrial DNA polymorphisms: χ^2 and the problem of small samples. Mol. Biol. Evol. 6: 539-545.

Table 1. Locations of boll weevil collections, corresponding list of abbreviations for populations, sample sizes, collection date and collector.

Region	Location	Abbreviation	Sample size	Collection date	Collector
South Central	Tampico, Mexico	MEX	27	07-Apr-99	Greenburg, ARS-IFNRRU
	Weslaco, TX	WTX	24	05-Jun-00	Sappington, ARS-IFNRRU
	Kingsville, TX	KTX	27	07-Oct-02	Montgomery, TBWEF
	El Campo, TX	ETX	25	01-Aug-02	Mote, TBWEF
	College Station, TX	CSTX	16	30-May-00	Spurgeon, ARS-APMRU
	Waxahachie, TX	WATX	26	16-Sep-02	Knutson, Texas A&M
Western	Hobart, OK	HOK	18	11-Dec-01	Massey, OBWEO
	Stamford, TX	STX	15	14-Aug-01	Cleveland, TBWEF
	Childress, TX	CHTX	25	01-Aug-01	Isbell, TBWEF
	Plainview, TX	PTX	19	11-Sep-01	Jones, TBWEF
	Big Spring, TX	BTX	15	14-Aug-01	Melendez, TBWP
	Artesia, NM	ANM	21	16-Oct-01	Norman, PVCBWCC
Eastern	Gilliam, LA	GLA	19	18-Jun-01	Courtright, LDAF
	Winnsboro, LA	WLA	21	06-Jul-01	Pylant, LADF
	Little Rock, AR	LAR	21	17-Jul-01	Kiser, ABWEF
	Cleveland, MS	CMS	20	24-Sep-01	Sprouse, SEBWEP
	Yazoo City, MS	YMS	20	11-Oct-01	Keene, SEBWEF
	Smithville, MS	SMS	18	08-Jul-02	Boyd, SEBWEF
	Malden, MO	MMO	21	30-Jan-02	Smith, SEBWEF
	Brownsville, TN	BTN	21	21-Jun-01	Seward, SEBWEF

Table 2. Restriction endonucleases, and total numbers of fragments and fragment patterns revealed in boll weevil mtDNA by each enzyme.

Type of enzyme	Enzyme	Total fragments	Fragment Patterns
Tetrameric	*Bsi*YI	14	6
	*Dde*I	33	8
	*Hae*III	10	2
	*Hha*I	4	1
	*Hinf*I	32	7
	*Mse*I	Unresolvable [b]	?
	*Msp*I	9	3
	*Nde*II	21	5
	*Rsa*I	23	7
	*Sau*96I	6	1
	*Scr*FI	4	1
	*Taq*I	21	4
Hexameric	*Bam*HI	0 [a]	1
	*Bgl*II	3	1
	*Bsp*106I	3	1
	*Dra*I	>19	1
	*Eco*RI	6	2
	*Eco*RV	3	1
	*Hha*I	0[a]	1
	*Hind*III	6	1
	*Hpa*I	3	1
	*Kpn*I	3	1
	*Pst*I	0[a]	1
	*Sac*I	2	1
	*Sfu*I	3	1
	*Vsp*I	24	4
	*Xba*I	5	1
	*Xho*I	3	1

[a] denotes no cut by restriction endonuclease.

b There were too many fragments to be resolved by electrophoresis.

Table 3. Boll weevil mtDNA haplotype frequency distributions, haplotype diversity (*h*) with standard error (S.E.), and percentage nucleotide diversity (π).

	Composite genotypes	MEX	WTX	KTX	ETX	CSTX	WATX	HOK	STX	CHTX	PTX	BTX	ANM	GLA	WLA	LAR	CMS	YMS	SMS	MMO	BTN
1	AAAAAAAAAA	0.074	0.333	0.444	0.577	0.563	0.4	0.833	0.8	0.68	0.579	0	0.667	0.158	0.048	0.19	0.25	0.15	0.278	0.143	0.333
2	ABABABBBAA	0	0.083	0.111	0.077	0.25	0.44	0	0.067	0	0	0	0.048	0.737	0.429	0.619	0.6	0.45	0.222	0.381	0.381
3	AABABAAAAA	0.037	0.083	0	0.038	0.063	0.04	0.111	0.133	0.08	0.263	1.0	0.286	0	0	0	0	0	0	0.143	0
4	ABAFABBBAA	0	0	0.074	0.077	0	0	0.056	0	0	0	0	0	0.105	0.524	0.048	0.1	0.1	0.222	0.143	0.19
5	BAAEAAAABA	0.37	0.417	0.111	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	AAAAAAACAA	0	0.042	0	0.115	0.125	0.04	0	0	0	0	0	0	0	0	0.095	0	0	0.278	0.19	0.095
7	AADAAAAAAA	0.222	0	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	ABCEAACBAA	0.074	0.042	0.037	0.077	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	ABABCBBBAA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.25	0	0	0
10	AAAAAAAFAA	0.037	0	0	0	0	0	0	0	0.04	0.105	0	0	0	0	0	0	0	0	0	0
11	AABABDAAAA	0	0	0	0	0	0.04	0	0	0.04	0	0	0	0	0	0	0	0	0	0	0
12	ABABABBEAA	0	0	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0.05	0	0	0	0
13	AABABADAAA	0	0	0	0	0	0	0	0	0.08	0	0	0	0	0	0	0	0	0	0	0
14	AAAAAAAAAB	0.074	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	AABAAAAAAA	0	0	0	0	0	0	0	0	0	0.053	0	0	0	0	0	0	0	0	0	0
16	AAAAACAAAA	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	AAECAAEACA	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	AAAADAAADA	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	ABADABBBAA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.05	0	0	0
20	ACABABBBAA	0	0	0	0	0	0	0	0	0.04	0	0	0	0	0	0	0	0	0	0	0
21	AAFAAAAAAA	0	0	0	0	0	0	0	0	0.04	0	0	0	0	0	0	0	0	0	0	0
22	ABABABBDAA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.048	0	0	0	0	0
23	AAAAAAAGAA	0	0	0	0	0	0.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	AAAEAAAAAA	0	0	0	0.038	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	AAAAAAFAAA	0	0	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	BAGGAAAHEA	0	0	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	AAAAAACAAA	0	0	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	AAAAAAAABA	0	0	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total samples	27	24	27	26	16	25	18	18	25	19	15	21	19	21	21	20	20	18	21	21
	Total haplotypes	10	6	11	7	4	6	3	3	7	4	1	3	3	3	5	4	5	4	5	4
	h	0.81	0.71	0.78	0.65	0.62	0.65	0.30	0.35	0.53	0.60	0.00	0.48	0.43	0.55	0.58	0.58	0.72	0.77	0.78	0.72
	S.E.	0.04	0.04	0.05	0.07	0.07	0.04	0.09	0.10	0.08	0.07	0.00	0.06	0.09	0.04	0.08	0.07	0.05	0.02	0.04	0.03
	π (%)	0.29	0.26	0.33	0.24	0.26	0.33	0.10	0.11	0.13	0.09	0.00	0.12	0.16	0.07	0.26	0.23	0.19	0.32	0.36	0.30

Table 4. χ^2 values and their significance from pairwise tests for heterogeneity of haplotype frequencies (above diagonal), and percentage nucleotide divergence among weevil populations (below diagonal).

	MEX	**WTX**	**KTX**	**ETX**	**CSTX**	**WATX**	**HOK**	**STX**	**CHTX**	**PTX**	**BTX**	**ANM**	**GLA**	**WLA**	**LAR**	**CMS**	**YMS**	**SMS**	**MMO**	**BTN**
MEX	0	19.2*	31.8***	39.9***	33.9***	43.3***	34.9***	31.6***	40.2***	32.8***	37.9***	37.4***	41.1***	45.3***	42.6***	41.2***	42.1***	39.1***	40.1***	41.7***
WTX	-0.003	0	15.7^{NS}	16.7**	11.2*	19.8**	16.6**	11.7*	23.2***	18.4**	31.5***	15.8**	27.1***	34.9***	24.6***	24.7***	28.6***	20.6***	21.8***	20.9***
KTX	0.023	0.013	0	14.9^{NS}	13.7^{NS}	20.7*	14.4^{NS}	12.6^{NS}	23.8**	22.3**	42.0***	18.7**	21.8***	28.2***	23.2**	16.6*	23.9**	17.6*	24.5**	15.8^{NS}
ETX	0.036	0.031	-0.001	0	5.3^{NS}	15.2*	7.5^{NS}	6.5^{NS}	16.5*	15.6*	37.0***	11.5*	23.5***	29.7***	19.7***	19.7***	25.1***	9.7^{NS}	15.6**	11.0^{NS}
CSTX	0.043	0.034	-0.004	-0.010	0	3.9^{NS}	8.8*	4.5^{NS}	13.5*	11.7*	27.3***	7.9*	13.4**	22.1***	9.2*	10.8*	15.7**	7.3^{NS}	8.5^{NS}	6.0^{NS}
WATX	0.094	0.072	0.012	0.020	0.005	0	15.6*	10.0*	20.2***	19.3***	36.0***	15.3**	9.5^{NS}	22.4***	7.8^{NS}	8.3^{NS}	15.6**	13.8**	11.8*	8.1^{NS}
HOK	0.035	0.041	0.028	0.011	0.013	0.074	0	2.1^{NS}	6.2^{NS}	5.9^{NS}	25.9***	3.8^{NS}	24.3***	31.5***	24.3***	20.3***	25.3***	17.8***	21.1***	16.6***
STX	0.034	0.039	0.026	0.009	0.010	0.068	-0.006	0	5.7^{NS}	4.9^{NS}	22.9***	1.2^{NS}	20.5***	28.5***	19.8***	16.8***	21.5***	15.5***	17.5***	14.2**
CHTX	0.041	0.048	0.039	0.020	0.022	0.086	-0.003	-0.004	0	8.2^{NS}	32.5***	9.0^{NS}	33.6***	42.2***	33.0***	29.4***	34.7***	27.1***	30.9***	26.0***
PTX	0.049	0.059	0.057	0.034	0.037	0.109	0.003	0.002	-0.002	0	18.8***	4.4^{NS}	28.6***	36.3***	28.2***	25.2***	29.6***	23.2***	23.0***	22.9***
BTX	0.183	0.192	0.206	0.176	0.174	0.244	0.127	0.120	0.101	0.080	0	18.4***	34.0***	36.0***	36.0***	35.0***	35.0***	33.0***	25.7***	36.0***
ANM	0.043	0.049	0.041	0.022	0.023	0.085	-0.001	-0.003	-0.003	-0.003	0.081	0	26.4***	34.7***	25.8***	22.6***	27.5***	21.0***	20.6***	19.8***
GLA	0.336	0.285	0.174	0.212	0.181	0.080	0.338	0.329	0.362	0.405	0.561	0.358	0	8.2*	3.4^{NS}	1.6^{NS}	7.1^{NS}	9.7^{NS}	8.8^{NS}	5.8^{NS}
WLA	0.442	0.384	0.257	0.303	0.269	0.146	0.449	0.440	0.477	0.524	0.682	0.472	0.006	0	13.9**	10.3**	13.2**	12.7**	12.6**	9.8*
LAR	0.243	0.200	0.105	0.132	0.106	0.032	0.238	0.230	0.259	0.297	0.452	0.257	-0.001	0.029	0	4.5^{NS}	10.2^{NS}	8.8*	7.0^{NS}	4.8^{NS}
CMS	0.265	0.220	0.120	0.152	0.124	0.043	0.262	0.253	0.283	0.322	0.478	0.281	-0.005	0.020	-0.010	0	7.9^{NS}	10.6*	9.5^{NS}	4.8^{NS}
YMS	0.348	0.296	0.185	0.223	0.190	0.088	0.349	0.340	0.372	0.414	0.560	0.367	-0.004	0.009	0.005	0.000	0	14.0**	13.2*	10.3*
SMS	0.106	0.082	0.018	0.023	0.009	-0.008	0.087	0.083	0.103	0.129	0.282	0.103	0.075	0.134	0.027	0.040	0.084	0	4.9^{NS}	2.7^{NS}
MMO	0.134	0.105	0.035	0.046	0.027	-0.007	0.117	0.111	0.130	0.155	0.277	0.126	0.047	0.098	0.010	0.019	0.053	-0.010	0	5.4^{NS}
BTN	0.156	0.124	0.045	0.063	0.043	-0.001	0.144	0.138	0.162	0.193	0.347	0.161	0.029	0.073	-0.001	0.005	0.036	-0.006	-0.010	0

The significance of the χ^2 was calculated through 10,000 repeated resamplings using a Monte Carlo simulation (Roff & Benzen 1989).
*** P<0.001, ** P<0.01, * P<0.05, ^{NS} not significant.

Table 5. Heterogeneity test of haplotype frequencies and % nucleotide divergence among populations within regions, and among regions.

Region	Heterogeneity test: χ^2	Nucleotide divergence (%): Mean value (range)
South-central	163.6***	0.024 (-0.010 ~ 0.094)
Western	84.5*** (37.3^{NS}) [a]	0.033 (-0.006 ~ 0.127)
Eastern	126.2***	0.027 (-0.010 ~ 0.134)
South-central vs. Western	92.4*** [b]	0.074 (0.020 ~ 0.244)
South-central vs. Eastern	116.4*** [b]	0.149 (-0.008 ~ 0.442)
Western vs. Eastern	162.4*** [b]	0.297 (0.083 ~ 0.682)
All Populations	1103.1***	0.134 (-0.010 ~ 0.682)

[a] Value when BTX was excluded.

[b] Populations pooled from each location were compared with each other.

*** $P<0.001$, ** $P<0.01$, * $P<0.05$, ^{NS} not significant.

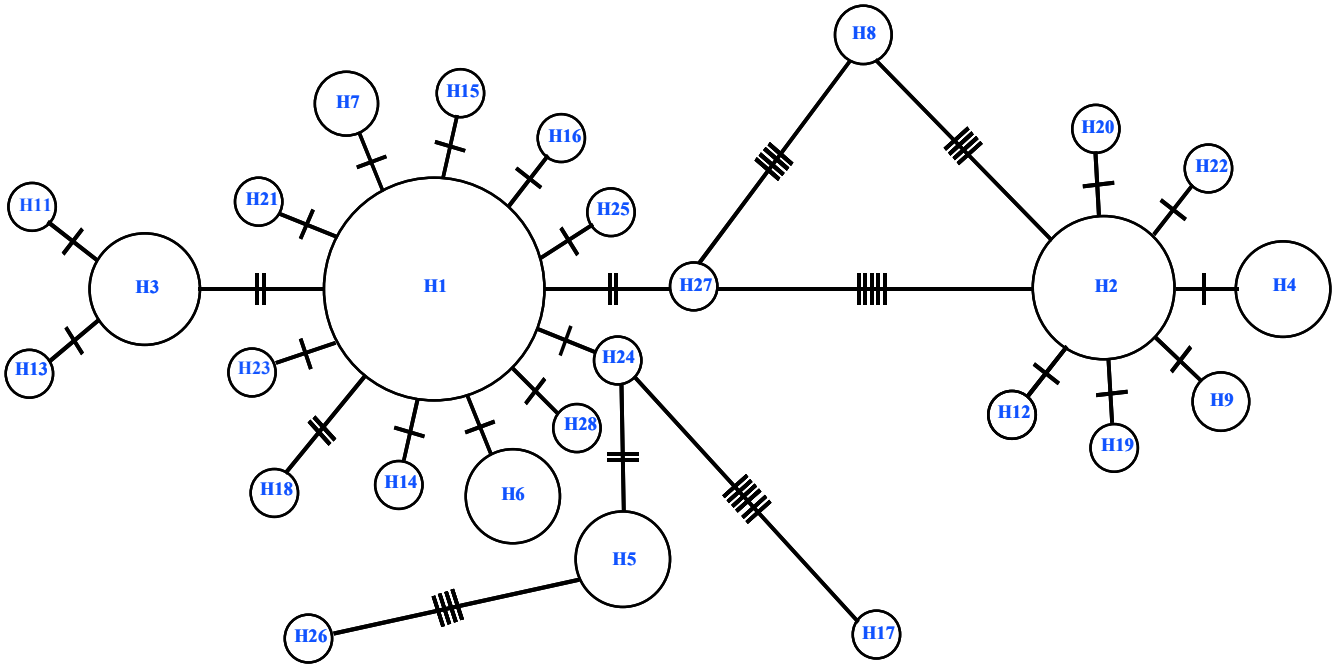

Figure 1. A parsimony network showing the relationship among the 28 mtDNA haplotypes (H1 to H28) of weevils. Short solid lines crossing branches indicate the number of hypothesized restriction site changes that occurred along a path. The area of each circle represents the frequency with which the haplotype occurs.

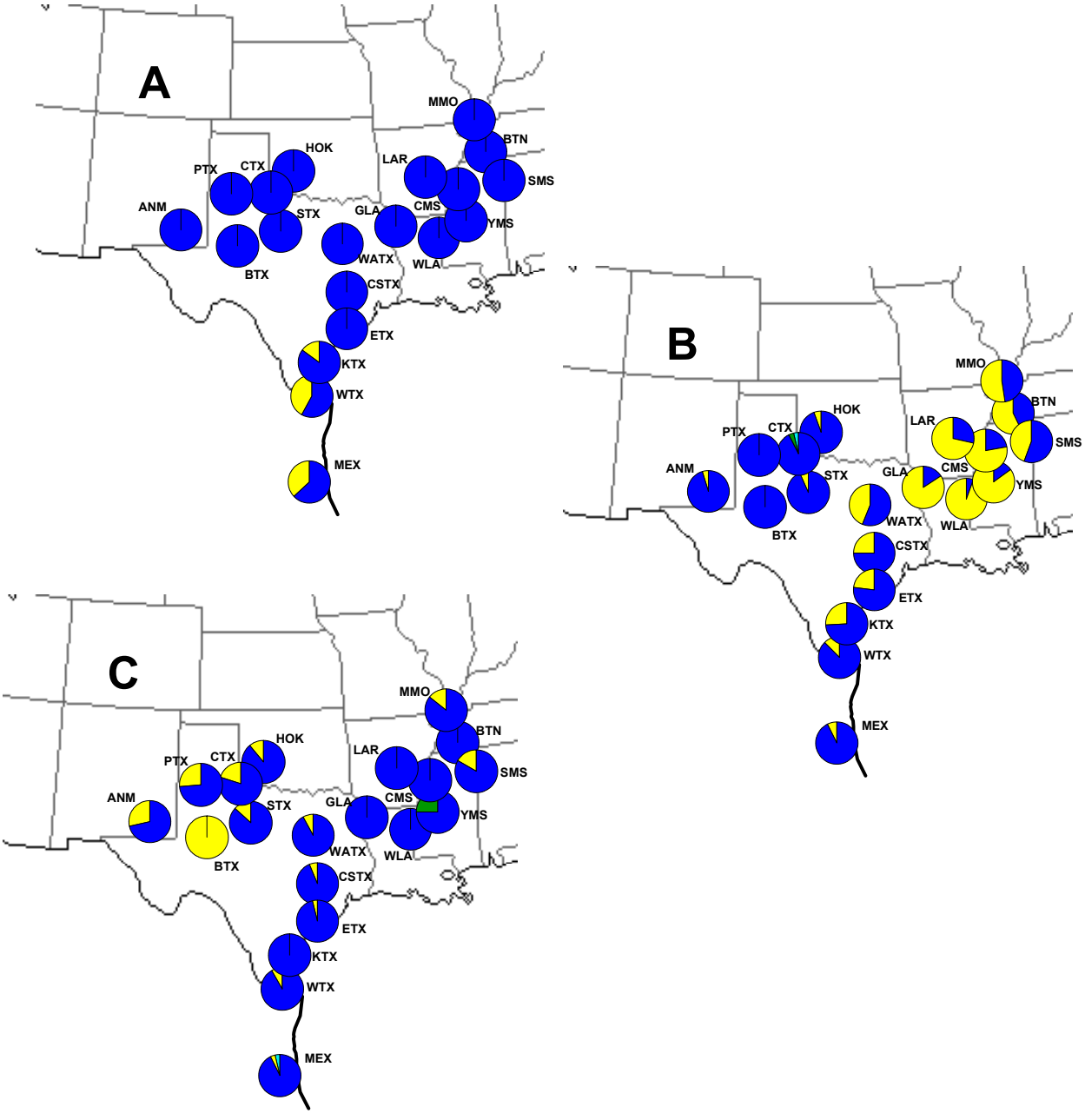

Figure 2. Geographic distributions of variable mtDNA genotypes produced by three representative endonucleases. **A**: *Eco*RI, **B**: *Msp*I, **C**: *Taq*I.

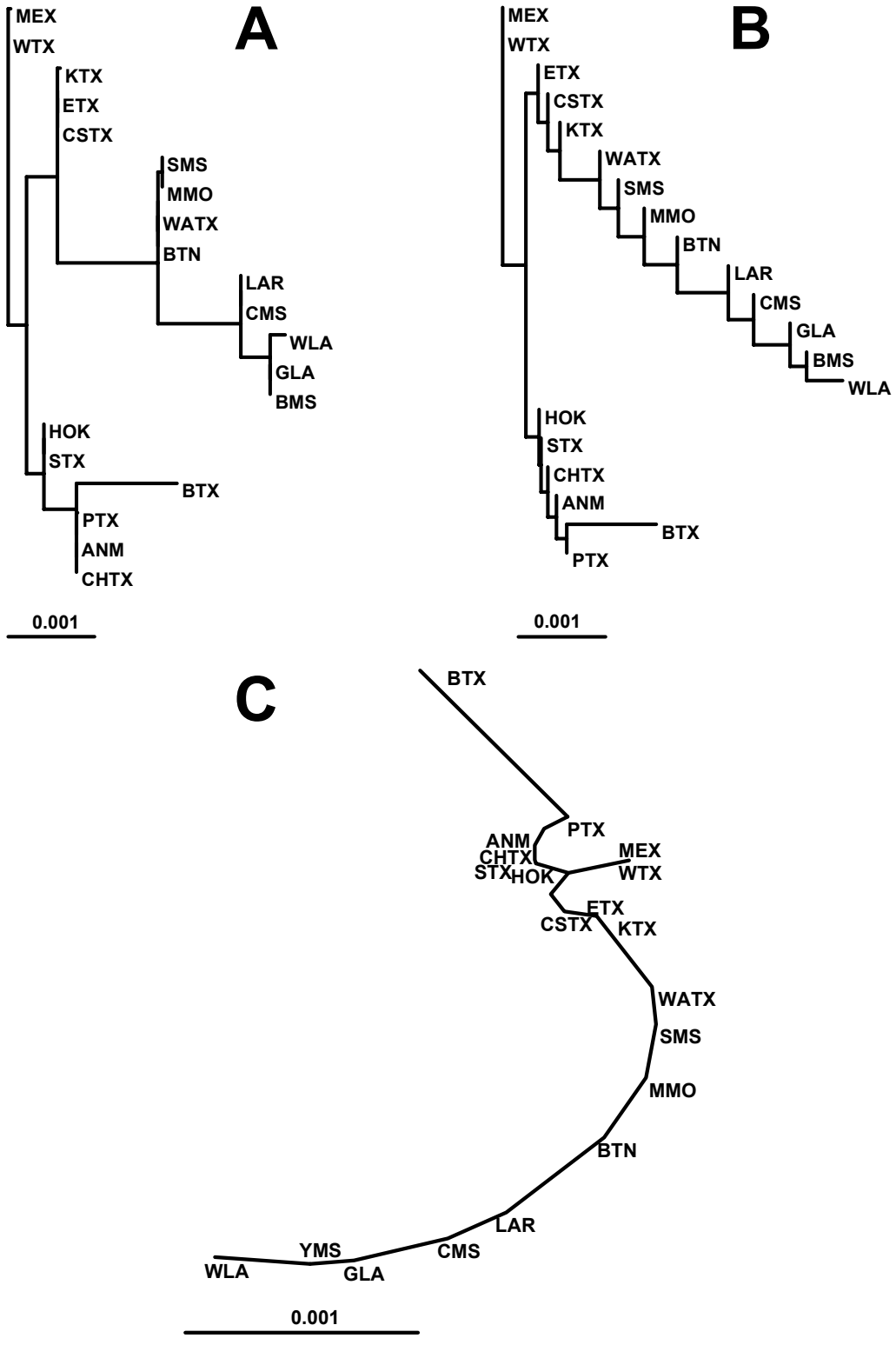

Figure 3. Phylogenetic trees showing the genetic relationships among the weevil populations sampled. A: FITCH tree, B: Neighbor-joining (NJ) tree, C: Unrooted NJ tree. Scale bar indicates nucleotide substitution rate per site.

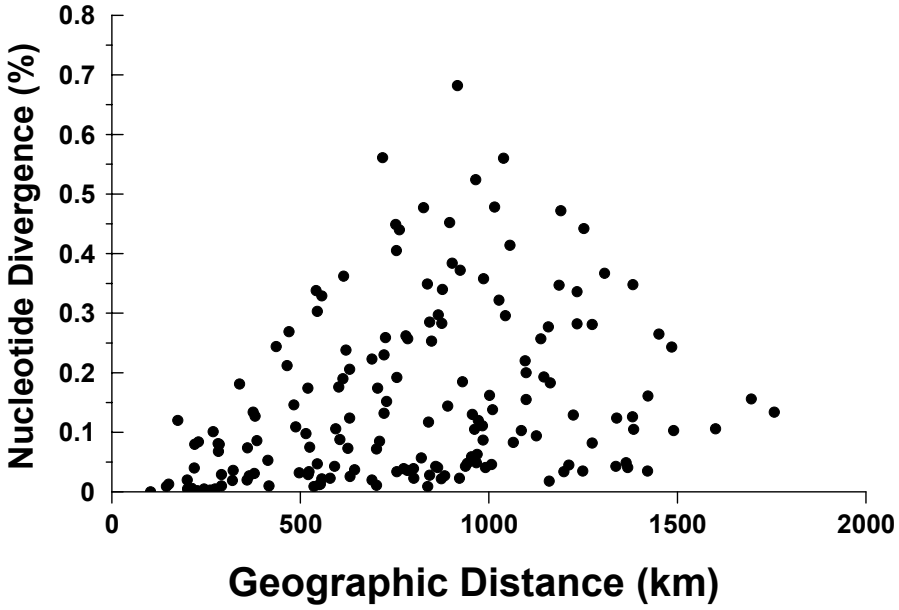

Figure 4. Relationship of genetic distance (% nucleotide divergence) to geographic distance between boll weevil populations. The correlation between genetic distance and geographic distance was calculated from 5000 replications and normalized by the Mantel statistic Z option of the NTSYS-PC program (Rohlf 1992).