ORIGINAL PAPER

# Strong genetic structure among populations of the invasive avocado pest *Pseudacysta perseae* (Heidemann) (Hemiptera: Tingidae) reveals the source of introduced populations

**Paul F. Rugman-Jones · Mark S. Hoddle · Phil A. Phillips · Gilsang Jeong · Richard Stouthamer**

**Abstract** In 2004, the avocado lace bug (ALB) *Pseudacysta perseae*, was discovered in San Diego County, CA, USA. Historically, California avocado producers have relied on biological control for suppression of injurious pests. A fundamental step in identifying biological control agents is determining the source of the invasive population, thus allowing a focused search of that area for natural enemies most closely adapted to the invasive pest genotype. In an attempt to determine the geographic area of origin for ALB to assist incipient biological control efforts, mitochondrial and microsatellite markers were used to compare the genetic profile of the California population of ALB with that of populations from the pest's known range in the US, Guatemala, Mexico and the Caribbean. Both types of marker revealed evidence for strong genetic structure among the sampled populations. Mitochondrial sequences narrowed the geographic origin of the California population, to populations in Texas and several Mexican states. This geographic range was refined further with the analysis of allele frequencies at eight microsatellite loci which identified ALB populations in the state of Nayarit, Mexico, as the most probable origin of the California population. Any effort to identify natural enemies of ALB for use in biological control in California should therefore focus on Nayarit. The strong genetic structure present among the ALB populations studied is discussed with regards to the potential influence of cytoplasmic incompatibility-inducing endosymbionts, host plant variation, and the putative native range of this pest.

**Keywords** mtDNA · Microsatellites · Exotic insect pest · COI · *Wolbachia*

P. F. Rugman-Jones (✉) · M. S. Hoddle · R. Stouthamer
Department of Entomology, University of California, Riverside, CA 92521, USA
e-mail: paulrj@ucr.edu

P. F. Rugman-Jones · M. S. Hoddle · R. Stouthamer
Center for Invasive Species Research, University of California, Riverside, CA 92521, USA

P. A. Phillips
University of California Cooperative Extension, 669 County Square Drive, Ventura, CA 93003, USA

G. Jeong
Department of Agricultural Biology, National Academy of Agricultural Science, RDA, 61, Seodundong Kweonseonku, Suwon 441-853, Republic of Korea

## Introduction

California is the leading producer of US grown avocados (*Persea americana* Miller [Lauraceae]), accounting for about 95% of domestic production. Annual crop value typically exceeds $320 million, and around 6,000 growers farm approximately 60,000 acres stretching south from San Luis Obispo, to San Diego at the US–Mexico border (http://www.avocado.org/california-avocado-history). In recent years, the California avocado industry has been impacted by the

introduction of a series of exotic foliar feeding arthropod pests that are specific to avocados. Following on from the establishment of red-banded whitefly *Tetraleurodes perseae* Nakahara (Hemiptera: Aleyrodidae) in 1982 (Rose and Wooley 1984) and persea mite *Oligonychus perseae* Tuttle, Barker & Abatiello (Acari: Tetranychidae) in 1990 (Bender 1993), California avocado growers were dealt a major economic blow with the establishment of avocado thrips *Scirtothrips perseae* Nakahara (Thysanoptera: Thripidae). After its discovery in 1996, *S. perseae* spread very rapidly (Hoddle and Morse 2003), and annual losses attributable to this pest have been estimated at $4.45–8.51 million per year (Hoddle et al. 2003). Thus, the discovery in 2004, that California had acquired yet another exotic arthropod pest of avocados, avocado lacebug (ALB) *Pseudacysta perseae* (Heideman) (Hemiptera: Tingidae), was disquieting.

*Pseudacysta perseae* was originally described as *Acysta perseae* from specimens collected from avocados in Florida, USA (Heidemann 1908) and was for a long time thought to have a limited distribution, primarily in peninsular Florida, and was regarded as no more than an occasional minor pest (Mead and Peña 1991). However, during the 1990s, this situation changed and damaging population outbreaks of ALB became commonplace in Florida and throughout the Caribbean as this pest began to invade new islands (Mead and Peña 1991; Peña 2003). At the outset of the present study, the geographic range of ALB was known to extend to the US states of Florida, Georgia, Louisiana and Texas; east to Bermuda; west to the gulf coast of Mexico; and south, through the Caribbean islands of Cuba, Puerto Rico, the Dominican Republic, and into Venezuela (Mead and Peña 1991; Almaguel et al. 1999; Sandoval Cabrera and Cermeli 2005).

In all areas where ALB is known to occur it is considered a pest of avocados that can be of economic importance (Peña 2003; Hoddle et al. 2005). Adults and nymphs of ALB form dense feeding colonies on the underside of leaves causing large necrotic patches, and in severe cases this feeding damage causes significant defoliation, which results in reduced fruit yields (Peña et al. 1998). Experimental evidence from Florida indicates that avocado varieties vary in their susceptibility to feeding damage (Peña et al. 1998), but in the Dominican Republic, severe damage to 'Hass' avocado (*P. americana* var. Hass) caused by ALB outbreaks has been observed (Hoddle unpublished).

The 'Hass' cultivar accounts for around 95% of avocados grown in California (http://www.avocado.org/the-hass-avocado-a-california-native/). Thus, in September 2004, the detection of ALB attacking backyard avocado trees in the Chula Vista and National City areas south of the city of San Diego (Hoddle et al. 2005; Humeres et al. 2009a) placed this pest in very close proximity to major 'Hass' avocado production areas. San Diego County produces around 60% of all California avocados (http://www.avocado.org/california-avocado-history) and it was feared that outbreaks similar to those witnessed in the Caribbean, could occur in California, and would cost the California avocado industry millions of dollars in pesticide applications and reduced crop yields. In addition, ALB is also known to feed on ornamental camphor *Cinnamomum camphora* (L.) and red bay *Persea borbonia* (L.) (Mead and Peña 1991) and therefore, potentially poses a significant pest threat to other native and exotic plant species in the Lauraceae.

California avocado production was once renowned for its almost exclusive reliance on biological control for suppression of injurious pests (McMurtry 1992). However, the importance of natural enemies for control of avocado pests has been slowly eroding, with the regular arrival of exotic pests driving a subsequent increase in pesticide use for their control (Hoddle 2004). While pesticides can result in immediate and rapid decline in a pest population, they also increase the likelihood of eliminating non-target species, such as beneficial natural enemies and pollinators (Mills and Daane 2005). In California avocados, the removal of natural enemies because of insecticide applications may result in secondary flare ups, particularly of mite and lepidopteran species that are normally well controlled (Hoddle 2004). Increased application frequencies of insecticides may also promote the development of pesticide resistance in already established pests. This is of particular concern in avocado thrips because Californian populations have developed varying levels of resistance to certain pesticides (Humeres and Morse 2006). An alternative strategy for the management of new invasive pests would be to revert to classical biological control; suppressing the introduced pest population's growth by the deliberate introduction and establishment of natural enemies from the pest's home range, in the invaded range. Several natural enemies of ALB have been identified from Florida (Peña et al. 1998, 2009;

Gagné et al. 2008; Henry et al. 2009) and Cuba (de la Torre et al. 1999), but their potential for use in classical biological control programs has not been tested. A fundamental step in the initial assessment of potential biological control agents is identifying the geographic area from where the target pest arrived. In theory, natural enemies from the actual donor region could be expected to be best adapted to the ALB genotype that invaded California, in comparison with candidate natural enemies from other potential donor regions. Thus, identifying the donor region(s) of California's ALB population would help focus a search for candidate biological control agents to that particular area.

To answer this question pertaining to the geographic origin of California's ALB population, systematic surveys to collect ALB were conducted, by the authors (MSH and PAP) or with cooperators, throughout avocado growing regions in the southeastern states of the US, the Caribbean, Mexico, Central America, and north eastern South America. Mitochondrial and nuclear genetic markers were used to ascertain if these collections constituted a single species (*P. perseae*), and to investigate population genetic variation and structure in an attempt to pin-point the geographic origin of the ALB population that invaded California, from within this natural distribution.

## Materials and methods

### Specimen collection

Between June 2005 and January 2008, specimens of ALB were collected from avocado trees (Hass and non-Hass) in the USA, Mexico, Guatemala, the Caribbean islands, and French Guyana (for complete collecting details see Table 1). Specimens were collected into 95% ethanol and returned to the University of California Riverside (UCR), USA. Where possible, these specimens were maintained below 4°C during transit. On arrival at UCR, specimens were stored at −20°C. These specimens were supplemented in 2010 with additional collections from the Dominican Republic (Table 1).

### DNA extraction

Initially, isolation of genomic DNA from individual ALB was achieved using a 'salting-out' technique similar to that used by ourselves for thrips specimens (Rugman-Jones et al. 2007). The method for ALB differed from that for thrips in that specimens were ground up, and that 5 μL of Proteinase K (Qiagen, Valencia, CA) was added to the extraction at the first step. However, the majority of extractions were performed using the Gentra® Puregene® Cell Kit (Qiagen). Extractions using this kit followed the manufacturer's supplementary protocol for extraction from a single *Drosophila melanogaster* (http://www.qiagen.com/literature/render.aspx?id=103451), with the exception that at the initial cell lysis step, we added 0.6 μL of Proteinase K and then incubated the samples for 3 h at 55°C prior to treating the lysates with Rnase A for 1 h. In addition, the resulting DNA was eluted in 30 μL of the kit's "DNA Hydration Solution". Extracted DNA was stored at −20°C.

### Generation and analysis of mitochondrial DNA sequences

The polymerase chain reaction (PCR) was used to amplify a section of the cytochrome C oxidase subunit I (COI) of mitochondrial DNA (mtDNA). Initial amplifications were performed using the primer pair LCO1490 (5′-GGTCAACAAATCATAAAGATATTGG-3′) and HCO2198 (5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) (Folmer et al. 1994). PCR was performed in 25 μL reactions containing 1× Thermopol Reaction Buffer (New England BioLabs, Ipswich, MA), 200 μM each dNTP, 0.2 μM of each primer, 2 mM MgCl$_2$, 1.2 μL BSA (NEB), 1 U *Taq* polymerase (NEB) and 2 μL DNA template (concentration not determined). PCR was performed on a Mastercycler® 5331 or Mastercycler® ep gradient S thermocycler (Eppendorf North America Inc., New York, NY) with an initial denaturation step of 5 min at 95°C; followed by 35 cycles of 95°C for 30 s, 43.5°C for 60 s, and 72°C for 90 s; and, a final extension step of 10 min at 72°C. Successful amplification was confirmed by visualizing PCR products after electrophoresis on 1% agarose gels stained with ethidium bromide. PCR products were purified using the Wizard® PCR Preps DNA Purification System (Promega, Madison, WI) and sequenced in both directions at the Institute for Integrative Genome Biology, University of California, Riverside. Sequences were aligned manually using BioEdit

**Table 1** Collection records for specimens of avocado lace bug, *Pseudacysta perseae* used in this study

| Collection site | Latitude, longitude | Altitude (m) | Date | Collector |
|---|---|---|---|---|
| USA | | | | |
| Chula Vista, CA[a] | 32.6947°N, 117.0700°W | 92 | September 2005 and July 2006 | E. Humeres and M. Hoddle |
| Weslaco, TX[a] | 26.1411°N, 97.9731°W | 22 | March 2006 | M. Ciomperlik |
| Dade Co., FL | 25.5078°N, 80.4997°W | 2 | June 2005 | J. E. Peña and W. Roltsch |
| Mexico | | | | |
| Oxkutcab, Yucatan | 20.2977°N, 89.3756°W | 11 | April 2006 | W. Roltsch |
| Hunucma, Yucatan | 21.0352°N, 89.8185°W | 5 | April 2006 | W. Roltsch |
| Merida, Yucatan | 21.0008°N, 89.6171°W | 5 | April 2006 | W. Roltsch |
| Jaltipan, Veracruz[a] | 19.9592°N, 94.7260°W | 53 | March 2006 | M. Hoddle and P. Phillips |
| La Tinaja, Veracruz[a] | 18.7657°N, 96.4538°W | 114 | March 2006 | M. Hoddle and P. Phillips |
| Cuamio, Michoacán | 20.0397°N, 101.1522°W | 1,798 | January 2008 | M. Hoddle and P. Phillips |
| Las Penas, Michoacán[a] | 18.0219°N, 102.5012°W | 71 | January 2008 | M. Hoddle and P. Phillips |
| Las Vivosas, Nayarit[a] | 21.5337°N, 104.9317°W | 888 | January 2008 | M. Hoddle and P. Phillips |
| San Blas, Nayarit[a] | 21.5435°N, 105.2795°W | 6 | January 2008 | M. Hoddle and P. Phillips |
| Xalisco, Nayarit[a] | 21.4340°N, 104.8998°W | 978 | January 2008 | M. Hoddle and P. Phillips |
| Puerto Vallarta, Jalisco[a] | 20.6224°N, 105.2224°W | 15 | January 2008 | M. Hoddle and P. Phillips |
| Manzanillo, Colima | 19.1093°N, 104.2617°W | 75 | January 2008 | M. Hoddle and P. Phillips |
| Puertocita de Lajas, Colima | 19.2362°N, 104.1886°W | 428 | January 2008 | M. Hoddle and P. Phillips |
| Tierra Colorada, Guerrero[a] | 17.1681°N, 99.5350°W | 286 | January 2008 | M. Hoddle and P. Phillips |
| Guanajuato, Guanajuato | 21.0194°N, 101.2683°W | 2,027 | January 2008 | M. Hoddle and P. Phillips |
| La Venta, Tabasco[a] | 18.0978°N, 94.0486°W | 24 | October 2007 | M. Hoddle and P. Phillips |
| Palenque, Chiapas[a] | 17.4882°N, 92.0410°W | 44 | October 2007 | M. Hoddle and P. Phillips |
| Guatemala | | | | |
| Santa Maria, Escuintla | 14.3992°N, 91.1059°W | 423 | March 2008 | M. Hoddle and C. Hoddle |
| Caribbean | | | | |
| Dominican Republic | – | – | April 2006 | C. Serra |
| | – | – | May 2010 | P. Phillips |
| Jamaica | – | – | March 2006 | M. Hoddle and P. Phillips |
| Puerto Rico | – | – | March 2006 | M. Hoddle and P. Phillips |
| St. Thomas and St. John | – | – | March 2006 | M. Hoddle and P. Phillips |
| St. Lucia | – | – | May 2006 | J. Leroy |
| St. Kitts | – | – | May 2006 | L. Rhodes |
| French Guyana | – | – | August 2006 | J. C. Streito |

GPS coordinates are given for individual US, Mexican and Guatemalan populations sampled. Caribbean islands were sampled at multiple locations (Dominican Republic = 7, Jamaica = 7, Puerto Rico = 11, St. Thomas = 5, St. John = 2) but data for individual sites has been omitted for conciseness and can be obtained by contacting the authors

[a] Populations used in microsatellite study

version 7.0.9.0 (Hall 1999) and a set of internal ALB-specific primers were designed using Primer3 (v.0.4.0) (Rozen and Skaletsky 2000); *ALB*-COI-f (5′-GGGCCATTTATTGGAAATGA-3′) and *ALB*-COI-r (5′-GATAGGATCCCCTCCTCCTG-3′). The internal primers provided more consistent amplification and

were combined in 25 μL reactions containing 1× Thermopol Reaction Buffer (NEB), 200 μM each dNTP, 0.2 μM of each primer, 0.5 mM $MgCl_2$, 1 U *Taq* polymerase (NEB) and 1 μL DNA template (concentration not determined). The thermocycler profile remained the same except the annealing

temperature was increased to 59°C. All PCR products were purified and sequenced. Primers were removed and each sequence was translated (http://www.ebi.ac.uk/Tools/emboss/transeq/index.html) to confirm the absence of nuclear pseudogenes (Song et al. 2008). Sequences were deposited in GenBank® (Benson et al. 2008).

Sequences were again aligned manually using BioEdit and trimmed to match those produced by the ALB-specific primers. There were no gaps in the COI fragment, making alignment straightforward. The resulting data matrix, totaling 429 sequences, each 515 bp long, was examined for evidence of gross population structuring and the existence of cryptic species (Hebert et al. 2003, 2004), across the geographic range from which ALB specimens were collected. A haplotype network was constructed using the statistical parsimony method of Templeton et al. (1992) implemented in TCS version 1.21 (Clement et al. 2000) with the default 95% connection limit. Genetic diversity was simply measured as the average number of pairwise differences ($\kappa$), within and between regions, calculated using DnaSP Ver. 5.10.00 (Librado and Rozas 2009). Regions were defined by grouping sample locations by state (for the USA and Mexico), country (Guatemala and French Guyana) or larger geo-political area (i.e., the Caribbean islands) (Table 1). One exception to this was the inclusion of a single specimen collected at high altitude (>5,800 ft) in Michoacán, Mexico, which was included with 3 specimens collected close by in Guanajuato, Mexico. For simplicity, this region is referred to in our analyses as Guanajuato.

While variation in the COI haplotypes detected in ALB was well within the limits expected for a single species (Hebert et al. 2004), the geographic distribution of mitochondrial haplotypes was highly structured (see Results). Therefore, as an additional qualitative measure of species identity, we looked for DNA sequence uniformity in the conserved D2 domain of 28S (28S-D2) nuclear ribosomal DNA (rDNA). This was done for at least one specimen displaying each COI haplotype detected at each collecting location. The forward (5′-CGTGTTGCTTGATAGTGCAG C-3′) and reverse (5′-TTGGTCCGTGTTTCAAGAC GGG-3′) primers of Campbell et al. (1993, 2000) were combined to amplify a 687 bp stretch of 28S-D2, in 25 μL reactions containing 1× Thermopol Reaction Buffer (NEB), 200 μM each dNTP, 0.2 μM each

primer, 1 U *Taq* polymerase (NEB) and 2 μL DNA template (concentration not determined). The thermocycler was programmed for an initial denaturing step of 3 min at 94°C; followed by 32 cycles of 94°C for 45 s, 55°C for 30 s, and 72°C for 90 s; and, a final extension step of 5 min at 72°C. PCR products were purified and sequenced as detailed above, and sequences were deposited in GenBank® (Benson et al. 2008).

Screening populations for infection by maternally inherited symbionts

The population genetics of mtDNA can be greatly influenced by the co-transmission of maternally inherited symbiotic bacteria that manipulate sexual reproduction of their hosts (Hurst and Jiggins 2005). We screened sampled ALB populations for infection with the three most commonly reported endosymbiotic reproductive manipulators, *Wolbachia*, *Cardinium*, and *Rickettsia* (Stouthamer et al. 1999; Hunter et al. 2003; Perlman et al. 2006). Initially, the presence of each bacterium was checked in a total of 24 specimens, drawn from populations across the sampled range, using bacterial genus-specific PCR protocols. The presence of *Wolbachia* was checked using the *wsp*-81f and *wsp*-691r primers of Braig et al. (1998) which amplify approximately 600 bp of the *wsp* gene. A *Wolbachia*-infected *Trichogramma pretiosum* was used as a positive control. The presence of *Cardinium* was assessed using the CLOf and CLOr primers of Weeks et al. (2003) which amplify approximately 450 bp of 16S rDNA. A *Cardinium*-infected unisexual form of *Aspidiotus nerii* was used as a positive control. The presence of *Rickettsia* was checked using the Rb-F and Rb-R primers of Gottlieb et al. (2006) which amplify approximately 900 bp of 16S rDNA. No positive control was available for the *Rickettsia* PCR. All PCRs included a negative control (sterile water).

Since the majority of specimens tested, across populations and COI haplotypes, were infected with *Wolbachia* (see "Results"), the particular *Wolbachia* strain involved was determined for each of the four most populous haplotypes (Hap-A, Hap-D, Hap-E and Hap-G; see "Results") by sequencing the five genes of the *Wolbachia* multi-locus sequence typing (MLST) system of Baldo et al. (2006). These haplotypes covered specimens from the vast majority of the

sampled range. The *wsp* gene (see above), an alternative method for typing *Wolbachia* strains (van Meer et al. 1999), was also sequenced for these specimens plus additional specimens drawn from other areas/haplotypes.

Isolation, amplification and sizing of microsatellite loci

ALB microsatellite loci were isolated by Genetic Identification Services (Chatsworth, CA). Four cDNA libraries were produced, each enriched for a different microsatellite repeat motif; CA-, GA-, ATG- and TAGA-. This resulted in the isolation of 24 potential loci, of which 16 were found to be uninformative (monomorphic) or difficult to amplify (data not shown). The remaining 8 loci are characterized in Table 2. Each locus was amplified individually via PCR, in 20 μL reactions containing 1× GeneAmp® PCR Buffer II (Applied Biosystems, Carlsbad, CA), 200 μM each dNTP, 0.2 μM each respective primer (Table 2), 2 mM MgCl$_2$, 0.5 U AmpliTaq Gold® DNA Polymerase (Applied Biosystems), and 1 μL DNA template (concentration not determined). PCR was performed on a Mastercycler® ep gradient S thermocycler programmed for: 10 min at 95°C; followed by 36 cycles of 94°C for 40 s, 55°C for

40 s, and 72°C for 30 s; and, a final extension step of 10 min at 72°C. The resulting amplicons were sized on an ABI PRISM 3100® Genetic Analyser using GENEMAPPER® software version 3.7 and the internal-lane size standard GeneScan®-500 LIZ™ (Applied Biosystems).

Microsatellite analyses

Individuals from 11 potential source locations were analyzed alongside the invasive California population (Table 1). For the microsatellite analyses, each potential source population contained only individuals collected from a single geographic location (i.e. no mixing of individuals from neighboring areas). Potential source populations were designated as those in which at least one resident individual shared a COI haplotype with the California population (see "Results"), and then only included in our analyses if at least 10 individuals were available for genotyping. Intrapopulation genetic diversity was assessed in terms of a range of standard parameters for each locus and across all loci. Allele frequencies and estimates of expected ($H_E$) and observed heterozygosity ($H_O$) were obtained using POPGENE version 1.32 (Yeh and Boyle 1997). $F_{IS}$ was calculated in FSTAT 2.9.3.2 (Goudet 1995, 2001) and deviations from Hardy–

**Table 2** Characterization of eight microsatellite loci isolated from *Pseudacysta perseae* including PCR primer sequences

| Locus | Primer sequences (5′–3′) | Repeat motif |
|---|---|---|
| *Pper*-B6 | F: CTCTGCAACGTGTAGTTAATTG | $(TC)_{14}$ |
|  | R: TGGTAGTAGAGGGAGAGTGATC |  |
| *Pper*-B7 | F: GTGGATGAAATAGAGGATGAAC | $(GA)_6TAGACA(GA)_7GTA(GA)_4GC(GA)_8$ |
|  | R: GCATGAGGAAGAGTATTCTAGC |  |
| *Pper*-B106 | F: TTCTCGCCTAACTTGAACTTC | $(TC)_{14}$ |
|  | R: CGGACTTAATGCAATTTATACC |  |
| *Pper*-B107 | F: GGAAGCCTCTGGACAAATG | $(AG)_2AA(AG)_{11}$ |
|  | R: TCTCTTCGTCCGAGCTGTAG |  |
| *Pper*-B120 | F: TGTGCCCATTTTTTGAGG | $(AG)_{12}(TGAG)_2TGATAG$ |
|  | R: ACTTGCGTAATACATTGTGGTG |  |
| *Pper*-C4 | F: CGATGTAGATGAAGGGAAATAG | $(TGA)_8$ |
|  | R: AATCAGTGGGAATGAAACTTAG |  |
| *Pper*-C9 | F: CCATCGCAGTCTTCATAAAC | $(GAC)_3(GAT)_4(GAC)_6(GAT)_5(GACGAT)_2$ |
|  | R: CATCGTCATCATCTTCTTCAT |  |
| *Pper*-D9 | F: CTCGATAATTCTCCTCGACTAG | $(AG)_3(AGAT)_{13}(AG)_{26}$ |
|  | R: CGCTTTCCTACTTACTTTTCAG |  |

Clone sequences for these primer pairs correspond to GenBank accessions JF834534–JF834541

Weinberg equilibrium (HWE) and linkage disequilibrium (LD) between pairs of loci was examined using a Markov chain method (Guo and Thompson 1992) provided in GENEPOP 4.0.10 (Raymond and Rousset 1995; Rousset 2008). Significance was estimated using 1000 dememorization steps, 100 batches, and 1,000 iterations per batch, and *P* values were adjusted using sequential Bonferroni corrections to control for multiple comparisons. Evidence for the presence of null alleles was sought using MICRO-CHECKER version 2.2.3 (van Oosterhout et al. 2004).

Interpopulation genetic variation was investigated using two standard measures. Multilocus estimates of $F_{ST}$ were calculated for all population pairs using FSTAT. Significance of the $F_{ST}$ estimates was evaluated with 1,000 random permutations of genotypes among samples, and *P* values were adjusted using sequential Bonferroni corrections. This method does not assume HWE within populations. Population differentiation was also examined by calculating Nei's unbiased genetic distances (Nei 1978) and constructing a dendrogram of relationships using the UPGMA method (unweighted pair group method with arithmetic mean), both in POPGENE. The resulting dendrogram was redrawn in FigTree v1.3.1 (Rambaut 2009). The distribution of microsatellite variation within and among populations was investigated using AMOVA (Excoffier et al. 1992) based on number of different alleles, implemented in ARLEQUIN v.3.11 (Excoffier et al. 2005). Significance was evaluated using 10 000 random permutations.

The origin of the invasive California population of ALB was investigated using the software program Structure version 2.3.3 (Pritchard et al. 2000). The Structure analysis examined the entire data set without considering geographic population information and used a Bayesian, Monte Carlo Markov Chain (MCMC) clustering approach to group individuals with similar multilocus genotypes into the most likely number of clusters (*K*) in the data set, estimating the log-likelihood of data belonging to a cluster, and calculating the probability of each individual belonging to each of the *K* clusters. Geographic population information was overlaid onto these clusters. If the invasive population clusters clearly with one of the potential source populations, this may reveal the origin of the invasive population. We applied an admixture ancestry model with independent loci (Pritchard et al. 2000), and explored likelihoods for models where the potential number of clusters for *K* ranged from 2 to 9. Ten independent runs were performed for each potential *K* and each run consisted of a burn-in period of 20,000 MCMC iterations and a data collection period of 50,000 MCMC iterations. At the end of these runs, the most likely value of *K* was determined by calculating $\Delta K$ (Evanno et al. 2005). The "*Full-Search*" algorithm in the program CLUMPP (Jakobsson and Rosenberg 2007) was used to control for label switching and genuine multimodality across the 10 replicate runs at the most likely value of *K*. Calculation of $\Delta K$ and creation of CLUMPP input files was done online using Structure Harvester v0.56.4 (Earl 2010).

## Results

Among the 16 regions (Table 1) and 429 individuals sequenced, we found 16 polymorphic sites and 9 distinct COI haplotypes (Fig. 1; Table 3; GenBank® accessions JF838244–838277). Synonymous substitutions accounted for 13 of the 16 polymorphic sites. A single, and unique, non-synonymous substitution was found in each of the three haplotypes represented by only a single specimen (Hap-C, Hap-F, and Hap-H). The distribution of haplotypes across the 16 regions was highly structured, with only 5 of 16 regions displaying more than one haplotype (Fig. 1, diagonal element Table 3). Hap-A (Fig. 1) was the largest with around 62% of all individuals analyzed and 100% of the California individuals. Hap-A also included all individuals from Texas (USA), Guerrero, Chiapas and Tabasco, along with the majority of individuals from Nayarit, Jalisco and Michoacán (all Mexico). Of the remaining haplotypes, 7 differed from Hap-A by no more than 2 nucleotides (0.39%). One of these, Hap-G, contained all individuals from Florida (USA), Yucatan (Mexico), French Guyana (South America), and the Caribbean islands. Indeed, Hap-G was the second most common haplotype in our sample but was not detected outside of these four regions (Fig. 1). Similarly, the third most common haplotype, Hap-E, was only found in the Vera Cruz population. The most divergent haplotype, Hap-I, differed from its nearest neighbor by 8 nucleotides (1.55%) and contained only the 4 individuals collected at high altitude (around 1,800 m) close to the border between the Mexican states of Guanajuato and Michoacán. Sequences of 28S-D2 (GenBank® accessions JF838221–838243)

**Fig. 1** Network of
*Pseudacysta perseae*
mitochondrial haplotypes
constructed from 429
sequences of the COI gene.
*Open circles* represent
hypothetical haplotypes not
detected in our sample and
*lines* connecting adjacent
haplotypes represent a
single nucleotide
substitution. Connecting
lines in *bold* represent non-
synonymous substitutions
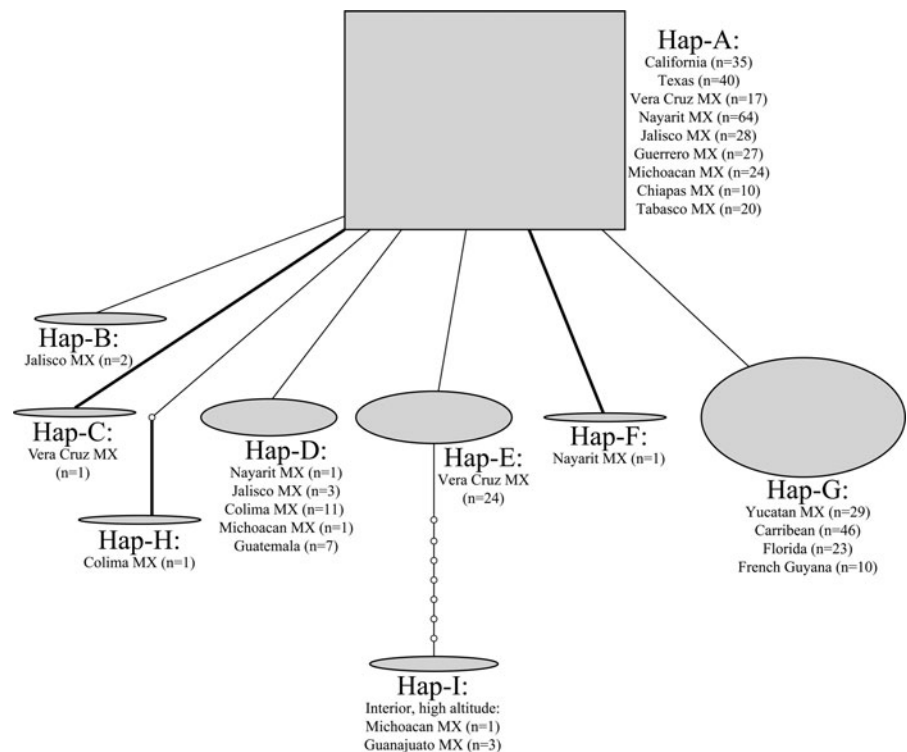


were identical across mitochondrial haplotypes suggesting the absence of cryptic species in the samples analyzed.

Of 24 specimens tested for infection with endosymbionts, 23 tested positive for *Wolbachia*, but none tested positive for either *Cardinium* or *Rickettsia*. Multi-locus sequence typing of the *Wolbachia* of a single individual from each of the four most populous mtDNA haplotypes (Hap-A, Hap-D, Hap-E and Hap-G; Fig. 1) found no nucleotide differences across five genes and 2,375 bp, indicating that the *Wolbachia* strain infecting different haplotypes (and consequently different populations) was identical (GenBank® accessions JN609257–609276). The sequence of the *wsp* gene (a further 555 bp) of these four specimens plus a further 9 drawn from other areas/haplotypes was also identical (GenBank® accessions JN859608–859620).

Across twelve populations, a total of 74 microsatellite alleles were observed in the 8 loci, with the total number of alleles at a single locus ranging from 3 (*Pper*-C4 and *Pper*-C9) to 27 (*Pper*-D9), and a mean (±SE) per locus of 9.25 ± 2.85 (Table 4). Of these alleles, 32 were unique to a single population, with the number of private alleles ranging from 0 (*Pper*-C4) to 16 (*Pper*-D9), and a mean per locus of 4.00 ± 1.79.

Significant deviation from HWE was detected in 8 of 96 locus by population comparisons (Table 5). In 6 of the 8 cases, positive $F_{IS}$ estimates indicated deviation was the result of an excess of homozygotes. In turn, results from MICRO-CHECKER revealed that the presence of a null allele may have contributed to four of these departures (Table 5). MICRO-CHECKER also revealed the potential presence of null alleles in one further locus by population combination (*Pper*-B107 by CA; Table 5). Across the twelve populations, only three locus pairs (from 199 tested) displayed significant LD, but in each case the significant pair was limited to a different population (Na-LV, *Pper*-B106 & *Pper*-D9; Na-SB, *Pper*-B107 & *Pper*-D9; Mi, *Pper*-B107 & *Pper*-B106).

All pairwise multilocus estimates of $F_{ST}$ were very high, ranging from 0.1061 to 0.8356 suggesting a high degree of population differentiation (Table 6). The high estimates of $F_{ST}$ were most likely the result of the high numbers of private alleles detected in our data; across all loci a mean (±SE) of 2.67 ± 0.76 (Table 5). Nei's unbiased genetic distance between populations was also high, ranging from 0.0959 to 1.4665 Table 6). For both measures, the two most similar populations were Na-LV and Na-X, and the

**Table 3** Variation in the DNA sequence of a 515 bp stretch of the mitochondrial COI gene of *Pseudacysta perseae*

| | CA | TX | FL | VC | Na | Ja | Co | Mi | Gu | Ta | Ch | Gn | Yu | GT | Carib. | GF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CA (n = 35) | 0.000 | | | | | | | | | | | | | | | |
| TX (n = 40) | 0.000 | 0.000 | | | | | | | | | | | | | | |
| FL (n = 23) | 1.000 | 1.000 | 0.000 | | | | | | | | | | | | | |
| VC (n = 42) | 0.595 | 0.595 | 1.595 | 0.549 | | | | | | | | | | | | |
| Na (n = 66) | 0.030 | 0.030 | 1.030 | 0.626 | 0.061 | | | | | | | | | | | |
| Ja (n = 33) | 0.152 | 0.152 | 1.152 | 0.747 | 0.179 | 0.288 | | | | | | | | | | |
| Co (n = 12) | 1.083 | 1.083 | 2.083 | 1.679 | 1.086 | 1.068 | 0.500 | | | | | | | | | |
| Mi (n = 25) | 0.040 | 0.040 | 1.040 | 0.635 | 0.069 | 0.184 | 1.050 | 0.080 | | | | | | | | |
| Gu (n = 27) | 0.000 | 0.000 | 1.000 | 0.595 | 0.030 | 0.152 | 1.083 | 0.040 | 0.000 | | | | | | | |
| Ta (n = 20) | 0.000 | 0.000 | 1.000 | 0.595 | 0.030 | 0.152 | 1.083 | 0.040 | 0.000 | 0.000 | | | | | | |
| Ch (n = 10) | 0.000 | 0.000 | 1.000 | 0.595 | 0.030 | 0.152 | 1.083 | 0.040 | 0.000 | 0.000 | 0.000 | | | | | |
| Gn (n = 4) | 9.000 | 9.000 | 10.000 | 8.452 | 9.030 | 9.152 | 10.083 | 9.040 | 9.000 | 9.000 | 9.000 | 0.000 | | | | |
| Yu (n = 29) | 1.000 | 1.000 | 0.000 | 1.595 | 1.030 | 1.152 | 2.083 | 1.040 | 1.000 | 1.000 | 1.000 | 10.000 | 0.000 | | | |
| GT (n = 7) | 1.000 | 1.000 | 2.000 | 1.595 | 1.000 | 0.970 | 0.250 | 0.960 | 1.000 | 1.000 | 1.000 | 10.000 | 2.000 | 0.000 | | |
| Carib. (n = 46) | 1.000 | 1.000 | 0.000 | 1.595 | 1.030 | 1.152 | 2.083 | 1.040 | 1.000 | 1.000 | 1.000 | 10.000 | 0.000 | 2.000 | 0.000 | |
| GF (n = 10) | 1.000 | 1.000 | 0.000 | 1.595 | 1.030 | 1.152 | 2.083 | 1.040 | 1.000 | 1.000 | 1.000 | 10.000 | 0.000 | 2.000 | 0.000 | 0.000 |

Average number of pairwise nucleotide differences within (diagonal element) and between geographic regions

*CA* California, USA; *TX* Texas, USA; *FL* Florida, USA; *VC* Vera Cruz, Mexico; *Na* Nayarit, Mexico; *Ja* Jalisco, Mexico; *Co* Colima, Mexico; *Mi* Michoacan, Mexico; *Gu* Guerrero, Mexico; *Ta* Tabasco, Mexico; *Ch* Chiapas, Mexico; *Gn* Guanajuato, Mexico; *Yu* Yucatan, Mexico; *GT* Guatemala; *Carib.* Caribbean Islands; *GF* French Guyana

**Table 4** Distribution of microsatellite alleles across populations of *Pseudacysta perseae* from the USA and Mexico Private alleles (i.e. alleles present in only one population) are in bold

| Locus | Size (bp) | CA (n = 30) | TX (n = 30) | VC-1 (n = 15) | VC-2 (n = 20) | Na-LV (n = 17) | Na-SB (n = 23) | Na-X (n = 13) | Gu (n = 19) | Ta (n = 22) | Ch (n = 10) | Ja (n = 23) | Mi (n = 17) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Pper*-B6 | **256** | – | – | – | – | – | – | – | – | – | – | – | – |
| | 260 | 0.033 | 0.017 | – | – | – | – | 0.039 | – | – | – | – | – |
| | 262 | 0.717 | 0.933 | 0.700 | 0.775 | 0.588 | 0.261 | 0.961 | 0.921 | 1.000 | 0.900 | 0.630 | – |
| | 264 | – | – | 0.167 | 0.225 | 0.265 | 0.652 | – | 0.053 | – | – | 0.348 | – |
| | 266 | – | 0.050 | 0.133 | – | – | – | – | – | – | 0.100 | – | – |
| | **270** | – | – | – | – | – | – | – | – | – | – | – | – |
| | 273 | 0.250 | – | – | – | 0.147 | 0.087 | – | 0.026 | – | – | 0.022 | 0.971 |
| | **280** | – | – | – | – | – | – | – | – | – | – | – | 0.029 |
| *Pper*-B7 | 231 | 0.117 | – | 0.300 | – | – | 0.913 | – | – | 0.636 | 0.100 | 0.174 | – |
| | 237 | – | – | – | 0.325 | – | – | – | – | 0.023 | – | – | – |
| | 238 | 0.883 | 1.000 | – | 0.300 | 0.941 | 0.087 | 1.000 | 0.816 | 0.341 | 0.900 | 0.609 | 0.912 |
| | 240 | – | – | 0.600 | 0.200 | – | – | – | 0.184 | – | – | – | 0.088 |
| | 241 | – | – | – | 0.175 | 0.029 | – | – | – | – | – | – | – |
| | 243 | – | – | 0.100 | – | – | – | – | – | – | – | 0.065 | – |
| | **245** | – | – | – | – | – | – | – | – | – | – | 0.087 | – |
| | **246** | – | – | – | – | – | – | – | – | – | – | 0.065 | – |
| | **256** | – | – | – | – | 0.029 | – | – | – | – | – | – | – |
| *Pper*-B106 | 226 | – | 0.967 | 0.633 | 0.750 | 0.294 | 0.348 | 0.269 | 0.421 | – | 0.500 | 0.522 | 0.676 |
| | 228 | – | – | – | – | 0.206 | 0.043 | 0.308 | – | – | – | 0.087 | – |
| | **230** | – | – | – | – | – | – | – | – | – | – | 0.043 | – |
| | **235** | 0.033 | – | – | – | – | – | – | – | – | – | – | – |
| | 237 | 0.967 | 0.033 | 0.367 | 0.250 | 0.500 | 0.609 | 0.423 | 0.579 | 1.000 | 0.500 | 0.348 | 0.324 |

**Table 4** continued

| Locus | Size (bp) | CA (n = 30) | TX (n = 30) | VC-1 (n = 15) | VC-2 (n = 20) | Na-LV (n = 17) | Na-SB (n = 23) | Na-X (n = 13) | Gu (n = 19) | Ta (n = 22) | Ch (n = 10) | Ja (n = 23) | Mi (n = 17) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Pper*-B107 | 206 | 0.667 | – | – | – | 0.735 | 0.522 | 0.731 | – | – | – | – | – |
|  | **212** | – | – | – | 0.350 | – | – | – | – | – | – | – | – |
|  | **214** | – | – | – | 0.025 | – | – | – | – | – | – | – | – |
|  | 216 | 0.250 | – | – | 0.075 | – | – | – | – | – | – | – | – |
|  | 218 | 0.083 | – | – | – | 0.029 | 0.239 | – | – | – | – | 0.043 | 0.353 |
|  | 222 | – | – | – | – | – | – | – | – | 0.023 | – | 0.196 | – |
|  | **224** | – | – | – | – | – | – | – | – | 0.068 | – | – | – |
|  | **225** | – | – | – | – | 0.059 | – | – | – | – | – | – | – |
|  | 227 | – | – | 0.400 | – | – | – | – | – | 0.909 | – | – | – |
|  | 229 | – | – | – | – | – | 0.065 | – | – | – | – | 0.391 | – |
|  | 231 | – | – | – | – | 0.177 | 0.174 | 0.269 | – | – | – | – | – |
|  | 234 | – | 0.500 | 0.600 | 0.550 | – | – | – | 0.868 | – | 1.000 | 0.370 | 0.647 |
|  | **239** | – | 0.367 | – | – | – | – | – | – | – | – | – | – |
|  | 241 | – | 0.133 | – | – | – | – | – | 0.132 | – | – | – | – |
| *Pper*-B120 | 281 | – | – | 0.333 | – | – | – | – | – | – | – | 0.283 | – |
|  | 283 | 1.000 | – | 0.667 | 0.425 | 0.471 | 0.935 | 0.192 | 1.000 | 1.000 | 1.000 | 0.152 | 1.000 |
|  | 285 | – | – | – | 0.575 | 0.529 | 0.065 | 0.808 | – | – | – | 0.457 | – |
|  | **287** | – | – | – | – | – | – | – | – | – | – | 0.109 | – |
|  | **308** | – | 1.000 | – | – | – | – | – | – | – | – | – | – |
| *Pper*-C4 | 127 | – | – | 0.067 | 0.150 | 0.147 | 0.261 | 0.269 | – | 0.023 | – | 0.174 | – |
|  | 130 | – | 1.000 | 0.867 | 0.725 | – | – | – | 0.974 | 0.273 | 0.750 | 0.826 | 0.206 |
|  | 137 | 1.000 | – | 0.067 | 0.125 | 0.853 | 0.739 | 0.731 | 0.026 | 0.705 | 0.250 | – | 0.794 |
| *Pper*-C9 | 259 | 1.000 | – | – | – | 0.323 | – | 0.308 | – | – | – | 0.283 | – |
|  | 262 | – | 1.000 | 1.000 | 1.000 | 0.677 | 1.000 | 0.692 | 1.000 | 1.000 | 1.000 | 0.652 | 1.000 |
|  | **264** | – | – | – | – | – | – | – | – | – | – | 0.065 | – |

**Table 4** continued

| Locus | Size (bp) | CA (n = 30) | TX (n = 30) | VC-1 (n = 15) | VC-2 (n = 20) | Na-LV (n = 17) | Na-SB (n = 23) | Na-X (n = 13) | Gu (n = 19) | Ta (n = 22) | Ch (n = 10) | Ja (n = 23) | Mi (n = 17) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Pper-D9 | **222** | – | – | – | – | – | – | – | – | – | – | 0.087 | – |
| | **227** | – | – | – | – | 0.059 | – | – | – | – | – | – | – |
| | 229 | 0.117 | – | – | – | 0.294 | 0.435 | – | – | – | – | – | – |
| | 231 | – | 1.000 | – | 0.350 | – | – | – | 0.947 | – | 0.050 | 0.435 | 0.294 |
| | 233 | – | – | 0.133 | – | – | – | – | – | 0.023 | – | – | – |
| | **235** | – | – | – | 0.450 | – | – | – | – | – | – | – | – |
| | 237 | – | – | 0.267 | 0.175 | – | – | – | – | 0.977 | 0.750 | – | – |
| | 239 | – | – | 0.267 | 0.025 | – | – | – | – | – | – | – | – |
| | **241** | – | – | – | – | – | – | – | – | – | – | 0.326 | – |
| | 250 | – | – | 0.300 | – | – | – | – | – | – | 0.200 | – | – |
| | **252** | – | – | 0.033 | – | – | – | – | – | – | – | – | – |
| | **255** | – | – | – | – | – | 0.065 | – | – | – | – | – | – |
| | 257 | 0.017 | – | – | – | 0.206 | 0.283 | – | – | – | – | – | – |
| | **267** | – | – | – | – | – | – | – | – | – | – | – | – |
| | 270 | – | – | – | – | 0.059 | 0.152 | – | – | – | – | – | – |
| | 272 | 0.850 | – | – | – | 0.294 | 0.044 | 0.346 | – | – | – | – | – |
| | 274 | 0.017 | – | – | – | 0.029 | – | – | – | – | – | – | – |
| | **319** | – | – | – | – | – | – | – | – | – | – | – | 0.118 |
| | **321** | – | – | – | – | – | – | – | – | – | – | – | 0.382 |
| | **323** | – | – | – | – | – | – | – | 0.053 | – | – | – | – |
| | **327** | – | – | – | – | – | – | – | – | – | – | 0.065 | – |
| | **331** | – | – | – | – | – | – | – | – | – | – | – | 0.206 |
| | **337** | – | – | – | – | 0.059 | – | – | – | – | – | – | – |
| | 339 | – | – | – | – | – | – | 0.654 | – | – | – | 0.022 | – |
| | **341** | – | – | – | – | – | – | – | – | – | – | 0.043 | – |
| | **343** | – | – | – | – | – | – | – | – | – | – | 0.022 | – |
| | **354** | – | – | – | – | – | 0.022 | – | – | – | – | – | – |

CA California, USA; TX Texas, USA; VC-1 La Tinaja, Vera Cruz; VC-2 Jaltipan, Vera Cruz; Na-LV Las Vivosas, Nayarit; Na-SB San Blas, Nayarit; Na-X Xalisco, Nayarit; Gu Guerrero; Ta Tabasco; Ch Chiapas; Ja Jalisco; Mi Michoacan (the latter are all samples from Mexico)

**Table 5** Conformity of microsatellite loci to Hardy–Weinburg equilibrium

| Locus | | CA (n = 30) | TX (n = 30) | VC-1 (n = 15) | VC-2 (n = 20) | Na-LV (n = 17) | Na-SB (n = 23) | Na-X (n = 13) | Gu (n = 19) | Ta (n = 22) | Ch (n = 10) | Ja (n = 23) | Mi (n = 17) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Pper*-B6 | $H_E$ | 0.430 | 0.128 | 0.481 | 0.358 | 0.579 | 0.510 | 0.077 | 0.152 | NA | 0.190 | 0.492 | 0.059 |
| | $H_O$ | 0.433 | 0.133 | 0.467 | 0.250 | 0.529 | 0.391 | 0.077 | 0.105 | NA | 0.200 | 0.565 | 0.059 |
| | $F_{IS}$ | −0.008 | −0.040 | 0.030 | 0.307 | 0.089 | 0.237 | 0.000 | 0.314 | NA | −0.059 | −0.153 | 0.000 |
| *Pper*-B7 | $H_E$ | 0.210 | NA | 0.559 | 0.753 | 0.116 | 0.162 | NA | 0.309 | 0.489 | 0.190 | 0.596 | 0.166 |
| | $H_O$ | 0.167 | NA | 0.533 | 0.800 | 0.118 | 0.174 | NA | 0.263 | 0.546 | 0.200 | 0.565 | 0.177 |
| | $F_{IS}$ | 0.208 | NA | 0.049 | −0.065 | −0.016 | −0.073 | NA | 0.151 | −0.118 | −0.059 | 0.053 | −0.067 |
| *Pper*-B106 | $H_E$ | 0.066 | 0.066 | 0.481 | 0.385 | 0.640 | 0.518 | 0.680 | 0.501 | NA | 0.526 | 0.611 | 0.451 |
| | $H_O$ | 0.000 | 0.067 | 0.600 | 0.300 | 0.706 | 0.522 | 1.000 | 0.421 | NA | 0.800 | 0.522 | 0.647 |
| | $F_{IS}$ | 1.000[a] | −0.018 | −0.260 | 0.225 | −0.107 | −0.008 | −0.500 | 0.163 | NA | −0.565 | **0.148** | −0.455 |
| *Pper*-B107 | $H_E$ | 0.494 | 0.608 | 0.497 | 0.583 | 0.437 | 0.650 | 0.409 | 0.235 | 0.172 | NA | 0.685 | 0.471 |
| | $H_O$ | 0.300 | 1.000 | 0.533 | 0.900 | 0.412 | 0.739 | 0.385 | 0.263 | 0.137 | NA | 0.304 | 0.000 |
| | $F_{IS}$ | 0.397[a] | **−0.664** | −0.077 | −0.565 | 0.059 | −0.140 | 0.063 | −0.125 | 0.213 | NA | **0.561**[a] | **1.000**[a] |
| *Pper*-B120 | $H_E$ | NA | NA | 0.460 | 0.501 | 0.513 | 0.125 | 0.323 | NA | NA | NA | 0.692 | NA |
| | $H_O$ | NA | NA | 0.000 | 0.350 | 0.471 | 0.130 | 0.231 | NA | NA | NA | 0.696 | NA |
| | $F_{IS}$ | NA | NA | **1.000**[a] | 0.307 | 0.086 | −0.048 | 0.294 | NA | NA | NA | **−0.006** | NA |
| *Pper*-C4 | $H_E$ | NA | NA | 0.248 | 0.447 | 0.259 | 0.394 | 0.409 | 0.053 | 0.439 | 0.395 | 0.294 | 0.337 |
| | $H_O$ | NA | NA | 0.133 | 0.400 | 0.294 | 0.522 | 0.539 | 0.053 | 0.318 | 0.300 | 0.261 | 0.294 |
| | $F_{IS}$ | NA | NA | 0.472 | 0.109 | −0.143 | −0.333 | −0.333 | 0.000 | 0.279 | 0.250 | 0.114 | 0.130 |
| *Pper*-C9 | $H_E$ | NA | NA | NA | NA | 0.451 | NA | 0.443 | NA | NA | NA | 0.501 | NA |
| | $H_O$ | NA | NA | NA | NA | 0.529 | NA | 0.462 | NA | NA | NA | 0.391 | NA |
| | $F_{IS}$ | NA | NA | NA | NA | −0.180 | NA | −0.044 | NA | NA | NA | 0.224 | NA |
| *Pper*-D9 | $H_E$ | 0.268 | NA | 0.775 | 0.660 | 0.797 | 0.717 | 0.471 | 0.102 | 0.046 | 0.416 | 0.705 | 0.733 |
| | $H_O$ | 0.233 | NA | 0.533 | 0.700 | 0.647 | 0.826 | 0.385 | 0.105 | 0.046 | 0.400 | 0.696 | 0.235 |
| | $F_{IS}$ | 0.131 | NA | 0.319 | −0.062 | **0.193** | −0.156 | 0.189 | −0.029 | 0.000 | 0.040 | 0.014 | **0.686**[a] |

**Table 5** continued

| Locus | CA (n = 30) | TX (n = 30) | VC-1 (n = 15) | VC-2 (n = 20) | Na-LV (n = 17) | Na-SB (n = 23) | Na-X (n = 13) | Gu (n = 19) | Ta (n = 22) | Ch (n = 10) | Ja (n = 23) | Mi (n = 17) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| All loci | | | | | | | | | | | | |
| $H_E$ | 0.183 | 0.100 | 0.437 | 0.461 | 0.474 | 0.385 | 0.352 | 0.169 | 0.143 | 0.215 | 0.572 | 0.277 |
| $H_O$ | 0.142 | 0.150 | 0.350 | 0.463 | 0.463 | 0.413 | 0.385 | 0.151 | 0.131 | 0.238 | 0.500 | 0.177 |
| $F_{IS}$ | 0.231 | **−0.510** | 0.205 | −0.004 | 0.023 | −0.076 | −0.098 | 0.107 | 0.090 | −0.114 | **0.128** | **0.370** |

Values in bold represent significant deviations after sequential Bonferroni correction

NA monomorphic; CA California, USA; TX Texas, USA; VC-1 La Tinaja, Vera Cruz; VC-2 Jaltipan, Vera Cruz; Na-LV Las Vivosas, Nayarit; Na-SB San Blas, Nayarit; Na-X Xalisco, Nayarit; Gu Guerrero; Ta Tabasco; Ch Chiapas; Ja Jalisco; Mi Michoacan (the latter are all samples from Mexico)

[a] Possible presence of a null allele (tested using MICRO-CHECKER v.2.2.3)

---

population that was most similar to the introduced California population was Na-LV (Table 6). The UPGMA dendrogram (Fig. 2) also placed the California population nearest to the Na-LV and Na-X populations in a group also containing Na-SB and Tabasco. The high levels of population differentiation were also detected in our AMOVA with differences among populations accounting for approximately half (51.85%) of total microsatellite variation (Table 7).

Bayesian clustering of multilocus genotypes, using $\Delta K$ (Evanno et al. 2005) to determine the true number of significant clusters (range tested $K = 1$–10, data not shown) revealed two major genetic clusters (Fig. 3). All individuals from California were assigned to a single cluster along with all individuals from the three Nayarit populations, with all individuals having membership coefficients close to 1.0 (mean ± SE; CA = 0.9959 ± 0.0001, Na-LV = 0.9937 ± 0.0003, Na-SB = 0.9927 ± 0.0007, Na-X = 0.9934 ± 0.0004). The majority of individuals from the Texas, Vera Cruz, Guererro, Chiapas, and Jalisco populations were assigned to the second cluster with similar membership coefficients. Only individuals from the Tabasco and Michoacán populations appeared as an admixture of the two clusters (Fig. 3).

## Discussion

The field of molecular population genetics offers a suite of markers and analysis methods that can be used to explore the invasion routes and most probable sources of introduced species from within large occupied ranges that may include both native and invaded regions. In this study, we used a combination of relatively conserved (mtDNA) and highly polymorphic (microsatellite) markers to characterize the genetic diversity of native and introduced populations of ALB, in an attempt to determine the most likely origin of the invasive population that entered California. Both types of marker revealed high levels of genetic structure in populations of ALB sampled from different geographic regions. The invasive California population was found to be most similar to populations from Las Vivosas and Xalisco, populations that are within close proximity to each other in the coastal hills of the state of Nayarit, along the tourist-popular Mexican Riviera.

**Table 6** Pairwise $F_{ST}$ (above diagonal) and Nei's unbiased genetic distances (below diagonal)

|       | CA     | TX     | VC-1   | VC-2   | Na-LV  | Na-SB  | Na-X   | Gu     | Ta     | Ch     | Ja     | Mi     |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| CA    | –      | 0.8233 | 0.6448 | 0.6172 | 0.3389 | 0.5476 | 0.5006 | 0.7166 | 0.7070 | 0.7237 | 0.5691 | 0.6294 |
| TX    | 1.4665 | –      | 0.5891 | 0.4615 | 0.6357 | 0.7114 | 0.7073 | 0.5905 | 0.8356 | 0.7550 | 0.4652 | 0.7030 |
| VC-1  | 1.2129 | 0.5074 |        | 0.1843 | 0.4014 | 0.3827 | 0.4801 | 0.3594 | 0.4380 | 0.1429 | 0.1541 | 0.4402 |
| VC-2  | 1.1864 | 0.3110 | 0.2195 |        | 0.3254 | 0.4032 | 0.3773 | 0.2947 | 0.5246 | 0.2830 | 0.1495 | 0.4009 |
| Na-LV | 0.2301 | 0.8094 | 0.8514 | 0.5620 |        | 0.2759 | 0.1061 | 0.5014 | 0.5402 | 0.4921 | 0.3325 | 0.3718 |
| Na-SB | 0.6053 | 1.2649 | 0.5664 | 0.6804 | 0.3363 |        | 0.4445 | 0.5650 | 0.4813 | 0.5002 | 0.3768 | 0.4409 |
| Na-X  | 0.3828 | 0.7636 | 0.9527 | 0.5328 | 0.0959 | 0.6428 |        | 0.6045 | 0.6487 | 0.5979 | 0.3712 | 0.5281 |
| Gu    | 0.7983 | 0.2358 | 0.2542 | 0.2006 | 0.6056 | 0.7191 | 0.6906 | –      | 0.6999 | 0.4725 | 0.3305 | 0.5154 |
| Ta    | 0.6609 | 1.1446 | 0.3256 | 0.6057 | 0.6450 | 0.3998 | 0.7657 | 0.5647 | –      | 0.5617 | 0.4901 | 0.6734 |
| Ch    | 0.9906 | 0.6179 | 0.0933 | 0.2535 | 0.8206 | 0.6500 | 0.9273 | 0.2326 | 0.2999 | –      | 0.2748 | 0.5409 |
| Ja    | 1.3372 | 0.4463 | 0.2251 | 0.2189 | 0.8267 | 0.8056 | 0.7998 | 0.3259 | 0.7383 | 0.3266 | –      | 0.4271 |
| Mi    | 0.6615 | 0.6440 | 0.5792 | 0.5141 | 0.4436 | 0.5171 | 0.7172 | 0.3658 | 0.7555 | 0.5236 | 0.8628 | –      |

All $F_{ST}$ estimates were significantly different from zero ($P < 0.01$ after sequential Bonferonni correction)

*CA* California, USA; *TX* Texas, USA; *VC-1* La Tinaja, Vera Cruz; *VC-2* Jaltipan, Vera Cruz; *Na-LV* Las Vivosas, Nayarit; *Na-SB* San Blas, Nayarit; *Na-X* Xalisco, Nayarit; *Gu* Guerrero; *Ta* Tabasco; *Ch* Chiapas; *Ja* Jalisco; *Mi* Michoacan (the latter are all samples from Mexico)
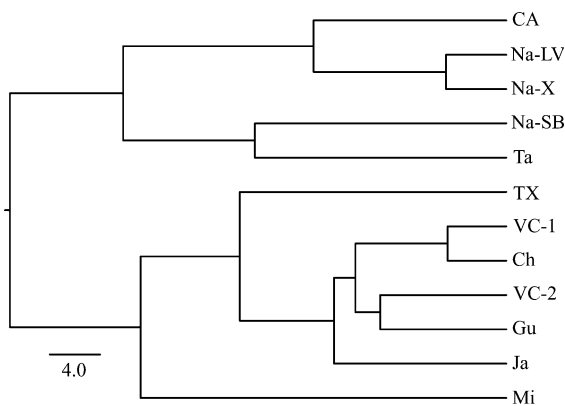


**Fig. 2** Dendrogram of relationships between 12 populations of *Pseudacysta perseae* constructed using a UPGMA method and based on Nei's unbiased genetic distance. *CA* California (USA); *TX* Texas (USA); *VC-1* La Tinaja, Vera Cruz; *VC-2* Jaltipan, Vera Cruz; *Na-LV* Las Vivosas, Nayarit; *Na-SB* = San Blas, Nayarit; *Na-X* Xalisco, Nayarit; *Gu* Guerrero; *Ta* Tabasco; *Ch* Chiapas; *Ja* Jalisco; *Mi* Michoacan (the latter are all Mexico)

## Mitochondrial diversity

Across its known range, COI sequences of ALB showed little genetic differentiation. Nine haplotypes were found among 429 sequenced individuals, but differences between these haplotypes were minor (Fig. 1). Indeed, only one haplotype differed from the others by more than 3 nucleotides. However, while nucleotide diversity was low, the geographic distribution of the nine haplotypes was highly structured, with a distinct east/west split. Sampled individuals from the California and Texas populations consisted entirely of a single haplotype; Hap-A (Fig. 1). Outside of the USA this haplotype was also the only one detected in our samples from the Mexican states of Guerrero, Chiapas and Tabasco, and was also the most prevalent haplotype in three of the four remaining states sampled along the Pacific Coast of Mexico (i.e. the states of Nayarit, Jalisco and Michoacán which are part of the tourist destination area known as the Mexican Riviera). Hap-A was also found in Vera Cruz, but here,

**Table 7** Analysis of molecular variance (AMOVA) for 12 populations of *Pseudacysta perseae*

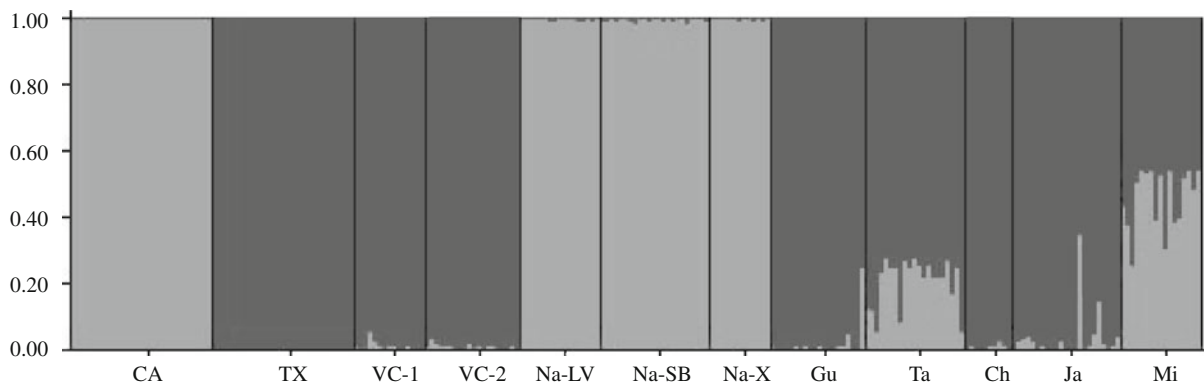| Source of variation | df  | Sum of squares | Variance components | Percentage of variation | P value |
|---------------------|-----|----------------|---------------------|-------------------------|---------|
| Among populations   | 11  | 577.784        | 1.2987 Va           | 51.85                   | <0.001  |
| Within populations  | 466 | 561.948        | 1.2059 Vb           | 48.15                   | <0.001  |

**Fig. 3** Bayesian clustering of the multilocus genotypes of 239 *Psuedacysta perseae* individuals. Each individual is denoted by a *narrow vertical bar* and its proportional membership in each of $K = 2$ clusters is represented by different *colors*. $K$ was determined according to Evanno et al. (2005). Geographic population IDs indicated on the x-axis: *CA* California (USA); *TX* Texas (USA); *VC-1* La Tinaja, Vera Cruz; *VC-2* Jaltipan, Vera Cruz; *Na-LV* Las Vivosas, Nayarit; *Na-SB* San Blas, Nayarit; *Na-X* Xalisco, Nayarit; *Gu* Guerrero; *Ta* Tabasco; *Ch* Chiapas; *Ja* Jalisco; *Mi* Michoacan (the latter are all Mexico)

a second haplotype which was unique to Vera Cruz (Hap-E), was slightly more common.

In contrast to the distribution of several haplotypes across the western part of the known range of ALB, a different but ubiquitous haplotype (Hap-G) was found among individuals sampled from Florida, Yucatan, French Guyana and throughout the Caribbean Islands. This "eastern" haplotype differed from the most common "western" haplotype by only one nucleotide, but Hap-G was not found in the western part of the range occupied by ALB. There are several potential explanations for this apparent geographic disparity. First, populations from the west may simply be geographically isolated from those in the east, such that admixture is absent or rare. However, the relatively close proximity of populations from the Yucatan peninsula with those from Chiapas and Tabasco in Mexico would argue against this hypothesis. A second explanation results from the number of individuals and number of locations sampled from the different regions. We only sampled a single population from Florida (Table 1) and it is possible that other haplotypes may have been detected had more samples been taken. However, this sample number issue is not a concern for the Caribbean region where we sampled multiple islands, or the Yucatan peninsula which was sampled at three different locations, and the results were consistent with those observed for Florida. A third explanation for the apparent lack of admixture, is that a degree of reproductive isolation may exist

between populations from the eastern and western range. Given the very small differences in mtDNA sequences and identical nature of their 28S-D2 sequences, it seems unlikely that any isolation results from the existence of cryptic species. However, there is evidence from a wide range of arthropods that such isolation can arise as a result of infection with microbes such as *Wolbachia*. *Wolbachia* are maternally inherited endosymbiotic bacteria that employ a variety of mechanisms to manipulate the reproduction of their hosts, in order to enhance their own transmission (Stouthamer et al. 1999). One such mechanism, cytoplasmic incompatibility (CI), occurs if *Wolbachia* modify the sperm of infected males in such a way that it prevents it from fertilizing an egg unless that is also infected with the same strain of *Wolbachia*. As a result, the sperm of males infected with a particular CI-*Wolbachia* strain cannot fertilize the eggs of females that are not infected with the same strain. Consequently, such females suffer reduced fitness relative to infected females (which can mate with any male), and a particular CI-*Wolbachia* strain spreads to fixation in a population (Werren and O'Neill 1997). Thus, if females from the western and eastern range of ALB harbor different strains of CI-*Wolbachia*, then any female immigrant (and her associated mtDNA haplotype) not carrying the correct CI-*Wolbachia* strain (e.g. a female from another population) would be quickly eliminated. *Wolbachia* infection is very common in ALB but our multi-locus sequence typing

shows that only a single *Wolbachia* strain is involved throughout the range of ALB. The identical nature of the *Wolbachia* was also evidenced by sequences of the wsp gene. Therefore, the present distribution of ALB mitochondrial haplotypes is unlikely to have resulted from mitochondrial sweeps caused by CI resulting from infection with *Wolbachia*. Furthermore, we found no evidence that any of our sampled populations were infected with *Cardinium* (Hunter et al. 2003) or *Rickettsia*; two other bacteria associated with manipulating host reproduction. At present the most plausible hypothesis to explain our findings is that the Eastern populations were founded by an ALB population with a single mitochondrial haplotype, which possibly reflects a population bottle neck during initial colonization.

The fact that the four most common mitochondrial haplotypes were all infected with the same *Wolbachia* strain is an interesting observation. Generally it is expected that cytoplasmically inherited microbes like *Wolbachia* will spread together with the mitochondrial type that was originally infected with the bacterium. Here we have a situation where the same *Wolbachia* strain is present in many different mitochondrial haplotypes. It is unlikely that the one *Wolbachia* type infected a single mitochondrial haplotype and subsequently over time the mitochondrial haplotypes accumulated mutations while the *Wolbachia* remained identical in the 2,375 bp that were sequenced for the MLST genes (and the 555 bp of the *wsp* gene). The alternative explanation is that this particular *Wolbachia* type also commonly transmits horizontally. How this horizontal transmission takes place is unknown. Several possible routes have been hypothesized. These can be divided into two mechanisms: introduction through parasitoids and parasites; and, introduction through sharing food sources. Several reports in the literature suggest horizontal transmission between insect species can take place through plants and interestingly, like ALB, the insects in these studies feed on plant cells using piercing and sucking mouthparts (Noda et al. 2001, Mitsuhashi et al. 2002 and Sintupachee et al. 2006).

The paucity of haplotypes in the eastern range of ALB is surprising. ALB was originally described in 1908, and for many years its distribution was thought to be restricted primarily to peninsular Florida. If Florida is truly part of the native range of ALB we might expect to see greater haplotype diversity; perhaps similar to that seen in populations along the Pacific coast of Mexico. We demonstrated that the presence of only a single haplotype throughout the Eastern range is unlikely to be the result of loss of variation following a symbiont-driven mitochondrial sweep. However, we may not need to invoke a role for a symbiont in such a mitochondrial sweep. There is a small amount of evidence that certain mitochondrial mutations may be adaptive and therefore have the potential to shape the phylogeography of populations (reviewed by Ballard and Melvin 2010). ALB has only become a serious pest in the past two decades and interestingly, the most damaging outbreaks of ALB appear to be restricted to the eastern part of its range (Mead and Peña 1991; Almaguel et al. 1999; Sandoval Cabrera and Cermeli 2005), and seemingly therefore, to a single haplotype (as shown in this study). Perhaps this haplotype (Hap-G) is particularly successful in the humid environment of Florida and the Caribbean, and its spread to fixation triggered the damaging ALB outbreaks that began in the 1990s.

The host plant, avocado, may have also played a role in selecting particular haplotypes. Historically three races of avocado have been recognized based on differences in morphological and taste characteristics (West Indian, Guatemalan, and Mexican). A recent genetic study provided strong support for these classical races, revealing substantial levels of differentiation between them (Chen et al. 2009). Commercial avocado production in the eastern and western parts of the range of ALB is based on varieties that originate from hybrid breeding of different races. The Hass variety accounts for the vast majority of production in California and Mexico and is a hybrid of Mexican and Guatemalan races (Chen et al. 2009). In contrast, most avocados grown commercially in Florida are hybrids of West Indian origin or hybrids between West Indian and Guatemalan races (Crane et al. 2007). Thus, the geographic distribution of ALB mtDNA haplotypes may be heavily influenced by adaptation of ALB populations to particular avocado varieties. Chen et al. (2009) also revealed previously undetected differentiation in wild avocados from central Mexico, with those growing at high altitude being genetically distinct from low altitude populations. Thus, host-plant adaptation may also account for the relatively distant mtDNA haplotype detected in the four ALB individuals collected around 1,800 m (Hap-I, Fig. 1).

Regardless of the reason for the east/west divide in the distribution of ALB mtDNA haplotypes, we used the fact that Hap-A was the only haplotype present in the invasive California population as justification for excluding regions from our microsatellite analyses (i.e., only populations where Hap-A was detected were analyzed). Thus microsatellite variation was not explored in populations from the eastern range of ALB. Similarly, although Hap-A was by far the most prevalent haplotype in surrounding states, it was completely absent in our samples from the Mexican state of Colima and from Guatemala. Therefore, these regions were also omitted from the microsatellite analyses.

## Microsatellite variation

Examination of microsatellite allele frequency data at eight loci revealed highly significant differentiation among the 12 populations studied. Pairwise estimates of $F_{ST}$ and Nei's unbiased genetic distances, even between closely neighboring populations, were very high, and AMOVA revealed that approximately half of the variation in our microsatellite data lay between populations. This was not surprising given the presence of many private alleles (43% of the total number of alleles), and the disjunct distribution of allele sizes (e.g., *Pper*-B120 and *Pper*-D9) and frequencies across all loci (Table 4). Together, these results suggest that admixture between ALB populations is rare, and that genetic drift and the emergence of unique alleles have played significant roles in producing highly genetically distinct populations. Similar to the mtDNA haplotypes, one explanation for this could be local adaptation of the ALB populations to what may be substantially genetically different avocado varieties (see above, Chen et al. 2009) or differing regional climates (i.e., Caribbean climates vs. high altitude drier conditions around Guanajuato, Mexico). These two factors, avocado variety and climate are not mutually exclusive as cultivars have been selected to perform well under prevailing climatic conditions. Unfortunately, in the present study we did not collect varietal information for host plants.

The presence of so many private alleles in our dataset may also indicate a weakness in this study. Measures of genetic diversity are sensitive to sample size and to account for the number of alleles per locus and accurately estimate heterozygosity, empirical tests

(coincidentally using eight microsatellite loci) have suggested that the number of individuals sampled in a population should be >20, and ideally >30 (Pruett and Winker 2008). In contrast to the ease with which ALB were found and collected throughout the Caribbean, ALB individuals were not abundant in any of the Mexican populations sampled (Hoddle et al. 2005). As a result, the numbers of individuals available for genotyping from these Mexican populations were relatively small ($n = 10$–23; Table 4). Thus, there is some doubt as to whether we were able to sample the true variation in each population. Of 32 private alleles, 21 occurred in their associated population at a frequency of 0.1 or less, and it seems likely that sampling a greater number of individuals may have detected at least some of these private alleles in neighboring populations.

Allele sizes at all eight microsatellite loci were not evenly separated as would be expected under a stepwise mutation model (SMM) (Slatkin 1995). This is well exemplified by considering the locus *Pper*-B120 (Table 4), which in the Texas population constitutes a fixed allele (308 bp), which is unique to that population, and which is 21 bp larger than the four remaining, equally spaced alleles (281–287 bp). Even larger size gaps appear between the alleles of *Pper*-D9 and it is quite certain that these disparate alleles must have arisen by some mutational process other than SMM. Alternatively, mis-priming and amplification of a different locus may explain this result. However, if this were the case, then when scoring allele sizes we might expect to find individuals for which there appeared to be more than two alleles. This was not the case, so we did not investigate these gaps further, and used measures of genetic differentiation that do not assume SMM (e.g., $F_{ST}$).

## Native range and origin of the California and other invasive populations

For much of its history, ALB was thought to be restricted primarily to peninsular Florida and therefore, Florida was assumed to be part of the native range of this pest. However, our results cast doubt on this assumption. We would typically expect to see the greatest amounts of intra-population genetic variation in areas that lie within the native range of an organism (Dlugosch and Parker 2008). Only a single mtDNA haplotype was present in our Florida sample. In

contrast, Jalisco and Nayarit (cumulative total across three populations) each with three mtDNA haplotypes, and Vera Cruz (cumulative total across two populations) with two, were the most diverse areas studied (Fig. 1). These levels of diversity were also reflected in nuclear loci with Jalisco and Nayarit both having 32 alleles across the 8 microsatellite loci, and Vera Cruz having 30 (Table 4). Furthermore, across the 8 loci there were no fixed alleles in Jalisco or Nayarit and only one in Vera Cruz. Although our small sample sizes for some locations make inferences somewhat questionable, it would appear that these three regions (and most likely the neighboring states of Guerrero and Michoacán) form at least part of the native range of ALB. This suggestion makes sense given that this area covers part of the native range of avocado. Whether the native range of ALB truly extends into Yucatan and Florida remains unclear since these populations were not subject to microsatellite genotyping.

In contrast to these relatively diverse populations, the two US populations examined were genetically poor. California and Texas individuals all possessed a single mtDNA variant, and across the eight microsatellite loci, total alleles numbered 17 and 13 respectively. Furthermore, three loci were fixed in California and five were fixed in Texas. Despite potential problems with the numbers of individuals sampled from some of the study populations, and the resultant high estimates of $F_{ST}$ and Nei's genetic distances, the California population was most similar to the population from Las Vivosas, Nayarit (Table 6; Fig. 2). Bayesian clustering of multilocus genotypes also grouped California individuals with those from the three Nayarit populations (Fig. 3). Membership coefficients for this cluster for individuals from these areas was close to 1. This can be interpreted as evidence that the invasive California population is derived from Nayarit, Mexico. In contrast, individuals from Texas clustered strongly with those from Vera Cruz, Guerrero, Chiapas and Jalisco suggesting that the Texas population may be derived from one of those populations, but that none of them are likely to have seeded the invasion of California. Only individuals from Michoacán and Tabasco showed intermediate membership coefficients, a result that in both cases may indicate admixture between a native and introduced population. Given the geographic distribution of the clusters, under this scenario it would appear more

likely that individuals had been transplanted from Nayarit to Michoacán and Tabasco rather than vice versa.

It is difficult to conclude much about the origin of the Caribbean populations. However, because they share a mtDNA haplotype, they are likely to have originated from Florida and/or the Yucatan Peninsula (or the same source as those populations if they are not native). Microsatellite data may reveal directional trends in the spread of ALB through the Caribbean islands. However, this was not a goal of the current project and these analyses were not undertaken. Microsatellite data may also shed light on the population from Colima, MX. The prevalent mtDNA haplotype in Colima, Hap-D, was detected in low numbers in neighboring states. Furthermore the haplotypes from those neighboring states were absent from Colima (Fig. 1). Interestingly, Hap-D was also the only haplotype observed in the sample from Guatemala. While the numbers of individuals sequenced from Colima ($n = 12$) and Guatemala ($n = 7$) are low (and therefore the similarity between the two may be superficial), the shared haplotype may be evidence that the Guatemala population is in fact an invasive one originating from a distinct Colima population, or vice versa.

Although the probable area of origin of the California population of ALB has been identified as Nayarit, Mexico, what is unresolved is how ALB was translocated from this area to San Diego, California. Two observations during sampling efforts along the Pacific Coast of Mexico allowed the development of a possible invasion scenario. This area, known as the Mexican Riviera, is extremely popular with American tourists who visit and camp in motorized recreation vehicles, and cruise ship traffic along this coast is very heavy. Also common in this area were roadside plant stands selling potted ornamental and fruit plants, including avocado seedlings, some of which were infested with ALB. Therefore, we tentatively hypothesize that ALB was accidentally moved on an infested avocado seedling that was purchased in Nayarit, imported illegally into Southern California, and planted in a backyard urban garden in San Diego.

Following the detection of ALB in California in 2004, populations did not develop into a significant new pest problem and detections in commercial avocado orchards have not been detected (Humeres et al. 2009a), possibly because this pest is vulnerable

to pesticides used to control avocado thrips, *S. perseae* (Byrne et al. 2010). However, ALB has not developed into a major pest of urban backyard avocados either, where pesticide use would be expected to be negligible. External factors that could hinder the spread of ALB include unfavorable climate and host plants. Southern California is not as humid as the coastal areas of Mexico where the source population likely originated, and winters are significantly colder too. Climate may also affect the ability of ALB to utilize a particular avocado variety. Although, ALB outbreaks have been observed on Hass avocado in the Dominican Republic (Hoddle unpublished), field studies in California have shown that Hass (the predominant cultivar in California) does not promote high population growth when compared to the highly preferred Bacon cultivar (Humeres et al. 2009a, b), a variety that is relatively uncommon in California. In the Dominican Republic, a favorable climate may help mitigate any adverse host plant effects. Alternatively, if the founding population of ALB in San Diego was small (as seems likely), genetic bottle-necking, drift, and subsequent inbreeding may have resulted in either a population that is not particularly pestiferous, or one that is undergoing a pronounced lag period as it adapts to its new environment. We did not look for statistical evidence of a bottleneck, since available methods (e.g., BOTTLENECK; Cornuet and Luikart 1996) typically require more than 10 polymorphic loci and large sample sizes to reliably detect such events (Luikart et al. 1998). However, reduced allelic diversity in the invading population may be interpreted as evidence that such an event has occurred.

While initial fears that ALB would become a serious pest problem in California avocado orchards have not yet materialized, this study has identified Nayarit, Mexico, as the most probable donor region for the population that invaded California. Thus, if deemed necessary, any effort to identify natural enemies of ALB for use in a classical biological control program against ALB in California should currently focus on Nayarit. However, this study has also revealed a distinct east/west split in ALB genotypes, and highlighted a potentially more pestiferous genotype 'lurking' in the eastern part of the studied range (i.e. Florida, Yucatan, and the Caribbean). As such, the identification of natural enemies from this region also, and continued genetic monitoring of ALB in California, would be pertinent so that if

ALB does at some point become a serious pest, we are better placed to identify and manage the threat quickly.

## References

Almaguel L, Blanco E, Suárez P, de la Torre P, Cáceres I, Nieves C, Márquez ME, Blanco L (1999) Control de la chinche del aguacate (*Pseudacysta perseae* (Heidemann)) en Ciudad de La Habana. Fitosanidad 3:69–74

Baldo L, Dunning Hotopp JC, Jolley KA, Bordenstein SR, Biber SA, Choudhury RR, Hayashi C, Maiden MCJ, Tettelin H, Werren JH (2006) Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. Appl Environ Microbiol 72:7098–7110

Ballard JWO, Melvin RG (2010) Linking the mitochondrial genotype to the organismal phenotype. Mol Ecol 19:1523–1539

Bender G (1993) A new mite problem on avocados. Calif Avocado Soc Yearb 77:73–77

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2008) GenBank. Nucleic Acids Res 36:D25–D30

Braig HR, Zhou W, Dobson S, O'Neill SL (1998) Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia*. J Bacteriol 180:2373–2378

Byrne FJ, Humeres EC, Urena AA, Hoddle MS, Morse JG (2010) Field evaluation of systemic imidacloprid for the management of avocado thrips and avocado lace bug in California avocado groves. Pest Manag Sci 66:1129–1136

Campbell BC, Steffen-Campbell JD, Werren JH (1993) Phylogeny of the *Nasonia* species complex (Hymenoptera: Pteromalidae) inferred from an rDNA internal transcribed spacer (ITS2) and 28 s rDNA sequences. Insect Mol Biol 2:225–237

Campbell B, Heraty JM, Rasplus J-Y, Chan K, Steffen-Campbell J, Babcock C (2000) Molecular systematics of the Chalcidoidea using 28 s-D2 rDNA. In: Austin A, Dowton M (eds) The Hymenoptera: evolution, biodiversity and biological control. CSIRO Publishing, Victoria, pp 57–71

Chen H, Morrell PL, Ashworth VETM, de la Cruz M, Clegg MT (2009) Tracing the geographic origin of major avocado cultivars. J Hered 100:56–65

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9:1657–1659

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

Crane JH, Balerdi CF, Maguire I (2007) Avocado growing in the Florida home landscape. Circular 1034, Horticultural Sciences Department, Florida Cooperative Extension Service,

Institute of Food and Agricultural Sciences, University of Florida

de la Torre P, Almaguel L, Blanco E (1999) Daños, distribución y enemigos naturales de la chinche de encaje del aguacate *Pseudacysta perseae* (Heidemann) (Hemiptera: Tingidae). Fitosanidad 3:69–74

Dlugosch KM, Parker M (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Mol Ecol 17:431–449

Earl DA (2010) Structure Harvester v0.56.4. http://taylor0.biology.ucla.edu/struct_harvest/. Accessed 28 Feb 2011

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinforma Online 1:47–50

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299

Gagné R, Peña JE, Acevedo FE (2008) A new Lestodiplosine (Diptera: Cecidomyiidae) preying on the avocado lace bug, *Pesudacysta perseae* (Hemiptera: Tingidae) in southern Florida. Fla Entomol 91:43–48

Gottlieb Y, Ghanim M, Chiel E, Gerling D, Portnoy V, Steinberg S, Tzuri G, Horowitz AR, Belausov E, Mozes-Daube N, Kontsedalov S, Gershon M, Gal S, Katzir N, Zchori-Fein E (2006) Identification and localization of a *Rickettsia* sp. in *Bemisia tabaci* (Homoptera: Aleyrodidae). Appl Environ Microbiol 72:3646–3652

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Goudet J (2001) FSTAT, version 2.9.3: a program to estimate and test gene diversities and fixation indices. Institute of Ecology, University of Lausanne, Switzerland. http://www2.unil.ch/popgen/softwares/fstat.htm. Accessed 28 Feb 2011

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48:361–372

Hall TA (1999) BioEdit: a user friendly biological sequence alignment and analysis program from Windows 95/98/NT. Nucleic Acids Symp 41:95–98

Hebert PDN, Cywinska A, Ball SL, deWard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Lond B 270:313–321

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc Natl Acad Sci USA 101:14812–14817

Heidemann O (1908) Two species of North American Tingitidae (Hemiptera: Heteroptera). Proc Entomol Soc Wash 10:103–108

Henry TJ, Peña JE, Long D, Acevedo F (2009) *Stethoconus praefectus* (Hemiptera: Miridae): first North American records of an old world plant bug predacious on avocado lace bug, *Pseudacysta perseae* (Hemiptera: Tingidae), in Florida. Proc Entomol Soc Wash 111:98–105

Hoddle MS (2004) Invasions of leaf feeding arthropods: why are so many new pests attacking California grown avocados? Calif Avocado Soc Yearb 87:65–81

Hoddle MS, Morse JG (2003) Avocado thrips biology and control. AvoResearch, California Avocado Commission, Irvine

Hoddle MS, Jetter KM, Morse JG (2003) The economic impact of *Scirtothrips perseae* Nakahara (Thysanoptera: Thripidae) on California avocado production. Crop Prot 22:485–493

Hoddle MS, Morse J, Stouthamer R, Humeres E, Jeong G, Roltsch W, Bender GS, Phillips P, Kellum D, Dowell R, Witney GW (2005) Avocado lace bug in California. Calif Avocado Soc Yearb 88:67–79

Humeres EC, Morse JG (2006) Resistance of avocado thrips (Thysanoptera: Thripidae) to sabadilla, a botanically derived bait. Pest Manag Sci 62:886–889

Humeres EC, Morse JG, Roltsch W, Hoddle MS (2009a) Detection surveys and population monitoring for *Pseudacysta perseae* on avocados in Southern California. Fla Entomol 92:382–385

Humeres EC, Morse JG, Stouthamer R, Roltsch W, Hoddle MS (2009b) Evaluation of natural enemies and insecticides for control of *Pseudacysta perseae* (Hemiptera: Tingidae) on avocados in Southern California. Fla Entomol 92:35–42

Hunter MS, Perlman SJ, Kelly SE (2003) A bacterial symbiont in the Bacterioidetes induces cytoplasmic incompatibility in the parasitoid wasp *Encarsia pergandiella*. Proc R Soc Lond B 270:2185–2190

Hurst GDD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proc R Soc Lond B 272:1525–1534

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Librado P, Rozas J (2009) DnaSP v.5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

Luikart G, Allendorf FW, Cornuet J-M, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered 89:238–247

McMurtry JA (1992) The role of exotic natural enemies in the biological control of insect and mite pests of avocado in California. In: Proceedings of II world avocado congress, pp 247–252

Mead F, Peña JE (1991) Avocado lace bug, *Pesudacysta perseae* (Hemiptera: Tingidae). Entomology Circular No. 346, Florida Department of Agriculture & Consumer Services, Division of Plant Industry

Mills NJ, Daane KM (2005) Biological and cultural controls … Nonpesticide alternatives can suppress crop pests. Cal Ag 59:23–28

Mitsuhashi W, Saiki T, Wei W, Kawakita H, Sato M (2002) Two novel strains of *Wolbachia* coexisting in both species of mulberry leafhoppers. Insect Mol Biol 11:577–584

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:582–590

Noda H, Koizumi Y, Zang Q, Deng K (2001) Infection density of *Wolbachia* and incompatibility lecel in two planthopper species, *Laodelphax striatellus* and *Sagatella furcifera*. Insect Biochem Mol Biol 31:727–737

Peña JE (2003) Pests of avocado in Florida. In: Proceedings of V world avocado congress, pp 487–494

Peña JE, Sundhari S, Hunsberger A, Duncan R, Schaffer B (1998) Monitoring, damage, natural enemies and control of avocado lace bug, *Pseudacysta perseae* (Hemiptera: Tingidae). Proc Fla State Hortic Soc 111:330–334

Peña JE, Triapitsyn SV, Long D, Evans GA, Roltsch W (2009) First record of *Erythmelus klopomor* (Hymenoptera: Mymaridae) as a parasitoid of the avocado lace bug, *Pseudacysta perseae* (Heteroptera: Tingidae). Fla Entomol 92: 394–395

Perlman SJ, Hunter MS, Zchori-Fein E (2006) The emerging diversity of *Rickettsia*. Proc R Soc Lond B 273:2097–2106

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Pruett CL, Winker K (2008) The effects of sample size on population genetic diversity estimates in song sparrows *Melospiza melodia*. J Avian Biol 39:252–256

Rambaut A (2009) FigTree, ver. 1.3.1 [online]. http://tree.bio.ed.ac.uk/software/figtree/. Accessed 21 Dec 2010

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

Rose M, Wooley JB (1984) Previously imported parasite may control invading whitefly. Calif Avocado Soc Yearb 68:127–131

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol Ecol Resour 8:103–106

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, pp 365–386

Rugman-Jones PF, Hoddle MS, Stouthamer R (2007) Population genetics of *Scirtothrips perseae*: tracing the origin of a recently introduced exotic pest of Californian avocado orchards, using mitochondrial and microsatellite DNA markers. Entomol Exp Appl 124:101–115

Sandoval Cabrera MF, Cermeli M (2005) Presencia de *Pseudacysta perseae* (Heidemann, 1908) (Insecta: Hemiptera: Tingidae) en Venezuela. Entomotropica 20:271–273

Sintupachee S, Milne JR, Poonchaisri S, Baimai V, Kittaypong P (2006) Closely related *Wolbachia* strains within the pumpkin arthropod community and the potential for horizontal transmission via the plant. Microb Ecol 51:294–301

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462

Song H, Buhay JE, Whiting MF, Crandall KA (2008) Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc Natl Acad Sci USA 105:13486–13491

Stouthamer R, Breeuwer JAJ, Hurst GDD (1999) *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. Annu Rev Microbiol 53:71–102

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data III. Cladogram estimation. Genetics 132:619–633

van Meer MMM, Witteveldt J, Stouthamer R (1999) Phylogeny of the arthropod endosymbiont *Wolbachia* based on the *wsp* gene. Insect Mol Biol 8:399–408

van Oosterhout C, Hutchinson WF, Willls DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Weeks AR, Velten R, Stouthamer R (2003) Incidence of a new sex-ratio-distorting endosymbiotic bacterium among arthropods. Proc R Soc Lond B 270:1857–1865

Werren JH, O'Neill SL (1997) The evolution of heritable symbionts. In: O'Neill SL, Hoffmann AA, Werren JH (eds) Influential passengers. Oxford University Press, New York, pp 1–41

Yeh FC, Boyle TJB (1997) Population genetic analysis of codominant and dominant markers and quantitative traits. Belgian J Bot 129:157