DOI: 10.1111/j.1570-7458.2007.00559.x

Population genetics of *Scirtothrips perseae*: tracing the origin of a recently introduced exotic pest of Californian avocado orchards, using mitochondrial and microsatellite DNA markers

P.F. Rugman-Jones, M.S. Hoddle & R. Stouthamer*
Department of Entomology, University of California, Riverside, CA 92521, USA

Accepted: 13 February 2007

Key words: avocado thrips, mtDNA, source population, bottleneck, gene flow, incursion management, invasive species, Thysanoptera, Thripidae

Abstract

The Californian avocado industry has recently been impacted by the establishment of three exotic arthropod pests that are native to Mexico and Central America. Establishment of all three pests is thought to have resulted from illegal movement of host plants (avocado) into California. To identify likely sources and routes of entry of such pest invasions, we examined the population genetics of the most recent invader, avocado thrips [*Scirtothrips perseae* Nakahara (Thysanoptera: Thripidae)], using mitochondrial DNA (mtDNA) and microsatellite markers. The mtDNA sequences revealed three geographically distinct and divergent lineages, of which the mtDNA haplotypes of Californian individuals were most closely related to populations in the center of the pest's native range. Analysis of allele frequencies at four microsatellite loci added resolution, indicating Coatepec Harinas, Mexico, as the most likely source of the Californian population. Statistically, we did not detect any bottleneck in population size associated with the invasion of California. However, estimates of the effective population size of the invading population suggest that a severe bottleneck occurred, indicating that the quantity of host-plant material entering California was small. Our findings implicate Coatepec Harinas, a large avocado germplasm and breeding center, as the most likely source of the introduced Californian population of *S. perseae* and as a likely source of previous and future avocado pest introductions. Efforts to identify natural enemies of *S. perseae* for biological control should focus on Coatepec Harinas and immediate surrounds. Moreover, identification of the source of invasive pests enables the establishment and enforcement of plant quarantine and free-trade protocols.

Introduction

Exotic invasive arthropod species pose a great threat to agriculture (Hallman & Schwalbe, 2002). They often have a significant detrimental impact on both the economic value of crops and the biodiversity of the agricultural ecosystems they invade (Pimentel et al., 2000; Reitz & Trumble, 2002). Free from natural enemies, exotic species often proliferate in their new environment causing extensive crop damage and displacing native species. A critical step in reconstructing the history of an invasion is identifying the geographical source of the exotic species. In turn, this may allow the identification of transport vectors and direct the search for appropriate biological control agents. The field of molecular genetics offers a powerful suite of tools for the identification of putative source populations (Luikart & England, 1999; Excoffier, 2004).

The avocado thrips, *Scirtothrips perseae* Nakahara (Thysanoptera: Thripidae), is a serious exotic pest of avocado, *Persea americana* Miller (Lauraceae), that was first detected attacking avocados around Port Hueneme in Ventura County, CA, USA, in July 1996. Around the same time, it was also detected some 160 km South at Irvine in Orange County, CA, USA. Following its discovery, *S. perseae* spread rapidly from these two focal points of invasion and today >99% of California avocado acreage is infested to some extent with this pest (Hoddle & Morse, 2003). The arrival of *S. perseae* had a significant negative economic impact on the California avocado industry. In California, this pest out-breaks in early spring and feeding

*Correspondence: E-mail: richard.stouthamer@ucr.edu

by adults and larvae can severely scar developing avocado fruit. Crop losses in 1998 were estimated at US$7.6–13.4 million (Jetter, 1999), and long-term losses have been estimated to be around $4.45–8.51 million per year (Hoddle et al., 2003). *Scirtothrips perseae* is the third invasive exotic pest of avocados to have established in California in recent years. Red-banded whitefly, *Tetraleurodes perseae* Nakahara, established in 1982 (Rose & Woolley, 1984), and persea mite, *Oligonychus perseae* Tuttle, Baker & Abatiello established in 1990 (Bender, 1993), both near the Port of San Diego. Both species were new to science at the time of their initial discovery in California, had been previously intercepted at border ports of entry before detection in California, and both are native to Mexico. At the time of its discovery in California, *S. perseae* was also undescribed, and traditional foreign exploration efforts have since shown that *S. perseae* is also native to Mexico, having a narrow range stretching throughout the avocado growing region between the state of Michoacan in Central Mexico and Central Guatemala (Hoddle et al., 2002a). The biology of these three pests also bears striking similarities that may indicate invasion pathways. All have a fundamental requirement for leaves in order to feed and reproduce (Hoddle et al., 1999; Hoddle & Soliman, 2001; Hoddle, 2002a). Consequently, the only way viable founding populations of these three pest species could have established in California was from infested avocado plants (or cuttings) smuggled into California from Mexico or Central America. None of these pests feeds or reproduces on mature avocado fruit, a product that is regularly smuggled into California from Mexico (CAC, 2004). Clearly, the routes of entry of avocado pests into California are unknown and there is a need to identify mechanisms of invasion and pinpoint the area of origin of avocado pests.

This study uses *S. perseae* as a model to investigate the source of invading exotic avocado pests using two types of molecular genetic markers. Within its native range, *S. perseae* only occurs at altitudes above 1500 m and its distribution is therefore fragmented, giving rise to geographically discrete populations. Typically, such fractured native populations exhibit high levels of genetic diversity within populations (relative to introduced populations), and so conserved markers are required for differentiation among populations. Furthermore, there is confusion over the taxonomic status of *S. perseae* and it is possible that what we currently regard as *S. perseae* may be a complex of several cryptic species (Johansen & Mojica-Guzman, 1998; Mound & zur Strassen, 2001). Identification and use of conserved molecular markers are required to circumvent taxonomic confusion and provide clarification of the specific status of what is currently defined as *S. perseae*. In contrast, as a result of population size bottlenecks and genetic drift in small founding populations, levels of genetic diversity are typically reduced, necessitating the use of highly polymorphic markers to retrieve meaningful population genetic information. In our search for the origin of the California population of *S. perseae*, we first used mitochondrial DNA (mtDNA) sequences to identify and exclude native populations that were highly divergent from the introduced populations. We then used highly polymorphic microsatellite loci to further resolve the relationships among genetically similar native and introduced populations.

Materials and methods

Collection

Following the discovery of *S. perseae* in California in July 1996, a coordinated collecting effort was made in the spring of 1997. Cooperators throughout California collected *S. perseae* from infested avocado orchards (Table 1). Collected thrips were preserved in 90% ethanol, shipped to the University of California at Riverside, CA, USA, and were stored at –20 °C until subject to genetic analyses in 2004. Foreign collections of *Scirtothrips* were made in Mexico, Guatemala, Honduras, Nicaragua, Costa Rica, and Panama between November 2003 and April 2004 (Figure 1; Table 1) from 'host' locations identified during more widespread collection efforts in Mexico, Central and South America, and the Caribbean over 1997–2000 (Hoddle et al., 2002a). Thrips were collected from commercial and non-commercial avocado orchards, roadside plants, and plants growing in private residences. Specimens were collected into 95% ethanol (see Hoddle et al., 2002a for field-collection methods), shipped to our laboratory in California, and were stored at –20 °C until analyzed. Finally, the current California population was sampled by field collections made in November 2004, in Orange, Santa Barbara, San Luis Obispo, and Ventura Counties (Table 1).

DNA extraction

Total DNA was isolated using a 'salting-out' protocol adapted from Sunnocks & Hales (1996). Each individual thrips was placed in a 0.5-ml low-binding microfuge tube containing 100 µl TNES (50 mM Tris, pH 7.5, 400 mM NaCl, 20 mM EDTA, and 0.5% SDS), and 1.7 µl Proteinase K (10 mg ml–1). Rather than grinding the specimen, the specimen's abdomen was simply pierced using a sterilized minuten pin. Tubes were incubated at 37 °C for 18 h, after which proteins were precipitated with the addition of 28 µl 5 M NaCl and 15 s of vigorous shaking. Proteins (and the specimen) were pelleted in a microfuge at 15 800 ***g*** for 5 min, and the supernatant was transferred to a new microfuge tube. DNA was precipitated from the supernatant by the

Table 1 Collection records for specimens of *Scirtothrips perseae* used in mitochondrial DNA (mtDNA) analysis

Collection location	Abbreviation	Specimen numbers	Collector	Date	GenBank[2]
Native range					
San Cristobal de las Casas, Chiapas, Mexico	MX-SC	MX-03-010, MX-03-017, MX-03-018	M. Hoddle	19/11/03	DQ075166–DQ075168
[1]San Andreas, Huayapam, Oaxaca, Mexico	MX-Hu	MX-03-040, MX-03-041, MX-03-042, MX-03-044, MX-03-045	M. Hoddle	22/11/03	DQ075142–DQ075146
[1]Chilapa de Diaz, Oaxaca, Mexico	MX-CD	MX-03-069, MX-03-070, MX-03-071	M. Hoddle	22/11/03	DQ075147–DQ075149
[1]Atlixco, Puebla, Mexico	MX-At	MX-03-080, MX-03-081, MX-03-082	M. Hoddle	24/11/03	DQ075158–DQ075160
[1]Tonatico, Mexico, Mexico	MX-To	MX-03-086, MX-03-215	M. Hoddle	25/11/03	DQ075150–DQ075151
[1]Coatepec-Harinas, Mexico, Mexico	MX-CH	MX-03-091, MX-03-096, MX-03-098, MX-03-100, MX-03-101	M. Hoddle	25/11/03	DQ075150–DQ075151 & DQ075161
Uruapan, Michoacan, Mexico	MX-Ur	MX-03-108, MX-03-118	M. Hoddle	26/11/03	DQ075162–DQ075163
Periban, Michoacan, Mexico	MX-Pe	MX-03-219, MX-03-220	M. Hoddle	28/11/03	DQ075164–DQ075165
Chimaltenango, Chimaltenango, Guatemala	GT	GT-04-015, GT-04-016, GT-04-017	M. Hoddle	08/03/04	DQ075169–DQ075171
California, USA					
Introduced range					
Irvine Ranch, Orange County	OC-97	US-97-CA/084, US-97-CA/085, US-97-CA/086	M. Hand	28/03/97	DQ075129–DQ075131
Somis, Ventura County	VC-97	US-97-CA/113, US-97-CA/114, US-97-CA/115, US-97-CA/116	M. Hand	31/03/97	DQ075132–DQ075134 & DQ075141
Santa Paula, Ventura County	CA-04	US-04-CA/004	P. Oevering	08/11/04	DQ075135
Somis, Ventura County	CA-04	US-04-CA/008	P. Oevering	08/11/04	DQ075156
Camarillo, Ventura County	CA-04	US-04-CA/013	P. Oevering	08/11/04	DQ075157
Oxnard, Ventura County	CA-04	US-04-CA/015	P. Oevering	08/11/04	DQ075136
South Coast Field Station, Irvine, Orange County	CA-04	US-04-CA/022	R. Vega	06/11/04	DQ075137
Nipomo, San Luis Obispo County	CA-04	US-04-CA/028	P. Oevering	08/11/04	DQ075138
San Luis Obispo, San Luis Obispo County	CA-04	US-04-CA/034	P. Oevering	08/11/04	DQ075172
Arroyo Verde, Ventura County	CA-04	US-04-CA/035	P. Oevering	08/11/04	DQ075173
Ojai, Ventura County	CA-04	US-04-CA/040	P. Oevering	09/11/04	DQ075174
Carpinteria, Santa Barbara County	CA-04	US-04-CA/051	P. Oevering	09/11/04	DQ075175
Santa Barbara, Santa Barbara County	CA-04	US-04-CA/054	P. Oevering	09/11/04	DQ075139
Goleta, Santa Barbara County	CA-04	US-04-CA/059	P. Oevering	09/11/04	DQ075140

[1]Native population included in the microsatellite analysis.
[2]GenBank: accession numbers for mtDNA sequences.

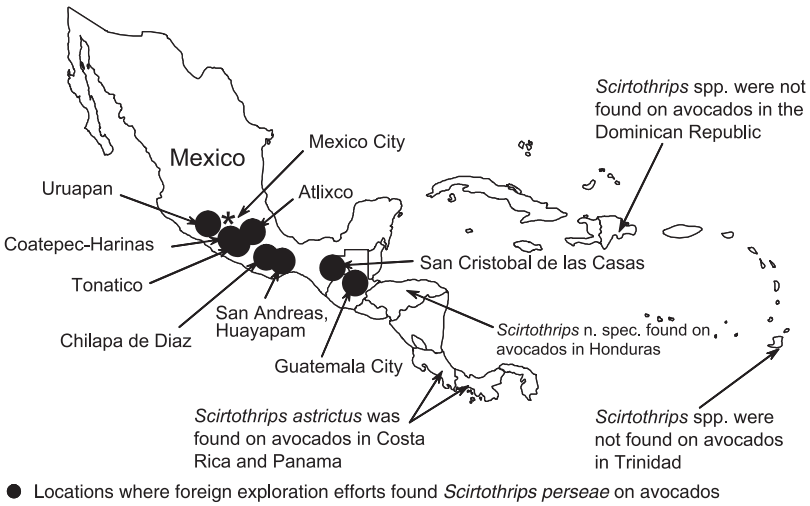

Figure 1 Geographic distribution of *Scirtothrips perseae* in Mexico and Central America (updated from Hoddle et al., 2002a).

addition of one volume of ice-cold 100% ethanol and incubation for 1 h at –20 °C. DNA was then pelleted in a cold (4 °C) microfuge at 15 800 ***g*** for 5 min, was washed in ice-cold 70% ethanol, air dried, and finally dissolved in 30 µl sterile distilled water. The advantage of this method is that it does not require destruction of the specimen. The thrips specimen was retrieved from the original microfuge tube by dissolving the pelleted proteins in 300 µl 50% ethanol for at least 1 h. The 50% ethanol was then replaced with 70% ethanol and the specimen stored at 4 °C for slide mounting. Specimens were cleared in 5% NaOH for 12 h, taken through an alcohol dehydration series, were placed in clove oil, and then slide mounted in balsam (Mound & Marullo, 1996). All slide-mounted specimens for this project have been deposited at the University of California Riverside Entomology Research Museum. All specimens were confirmed as *S. perseae* via two techniques: using traditional morphological characters (Nakahara, 1997) and a molecular method based on the size of the internal transcribed spacer regions 1 and 2 of rRNA (Rugman-Jones et al., 2006).

Mitochondrial DNA sequencing and analysis

The polymerase chain reaction (PCR) was used with the primer pair LCO1490 (5′-GGTCAACAAATCATAAAGA-TATTGG-3′) and HCO2198 (5′-TAAACTTCAGGG-TGACCAAAAAATCA-3′) to amplify a segment of the cytochrome *c* oxidase subunit I (COI) gene of mtDNA (Folmer et al., 1994). PCR was performed in 25 µl reactions containing 2 µl DNA template (concentration not determined), 1× PCR buffer (containing 2 mM $MgSO_4$; NEB Ipswich, MA, USA), 20 µM each dNTP, 2 mM $MgCl_2$, 1.25 µl BSA (NEB), 0.2 µM each primer, and 2 µl *Taq* polymerase (NEB). Amplification was performed in an Mastercycler 5331 (Eppendorf, Hamburg, Germany) programed for: an initial denaturing step of 1 min at 94 °C; followed by five cycles of 30 s at 94 °C, 1 min 30 s at 45 °C, and 1 min at 72 °C; followed by a further 35 cycles of 30 s at 94 °C, 1 min 30 s at 51 °C, and 1 min at 72 °C; and a final extension of 5 min at 72 °C. Amplified DNA was cleaned using the Wizard® PCR Preps DNA purification system (Promega, Madison, WI, USA) and was direct sequenced in both directions at the University of California Riverside Genomics Institute, Core Instrumentation Facility using an Applied BioSystems 3730 DNA analyzer with a Big-Dye® V3.1 kit (Applied BioSystems, Foster City, CA, USA). Raw sequences were trimmed, removing the primers, and were aligned manually in BioEdit v. 6.0.7 (Copyright© Hall 1997–2004). There were no gaps in the COI region, making alignment straightforward. Sequences are deposited in GenBank (Table 1). A haplotype network was constructed using TCS v. 1.18 (Clement et al., 2000), based on 673 bp of the COI region for 47 *S. perseae* specimens from 12 collection locations (Table 1), and genetic diversity was initially examined in terms of the average number of pairwise nucleotide differences (κ), within and between collecting locations, calculated using DnaSP v. 4.0 (Rozas et al., 2003). This revealed three very distinct and divergent mtDNA groups (see Results) for which haplotype diversity, nucleotide diversity (π), theta (θ_W) per site calculated from the number of segregating sites (Watterson, 1975), and κ were calculated, again using the DnaSP program. Pairwise estimates of F_{ST} were also obtained for these groupings using the distance method implemented in ARLEQUIN v. 2.000 (Schneider et al., 2000). The significance of the F_{ST} values was evaluated by permuting the haplotypes between groups as implemented in this program (number of permutations was set at 1000). Open reading frames were identified and sequences translated using the EMBOSS Transeq online software (http://www.ebi.ac.uk/emboss/transeq).

Due to large differences in the mitochondrial haplotypes (see Results), we also amplified a section of the conserved D2 domain of the large subunit rDNA from a representative sample of individuals, as an extra insurance of species identity. In a separate study (MS Hoddle, PF Rugman-Jones & R Stouthamer, unpubl), the D2 domain of 19 different *Scirtothrips* species showed zero intraspecific sequence divergence. However, pairwise comparisons of this region showed interspecific sequence divergence ranging from around 0.6 to 11.0% (mean ± SE 8.5% ± 0.2). Thus we looked for monophyly in the D2 domain as a reasonable qualitative confirmation that the *S. perseae* mitochondrial haplotypes were indeed a single species. A 539-bp section of the D2 domain was amplified in 25 μl reactions containing 2 μl DNA template, 1× PCR buffer, 1 μl BSA, 20 μM each dNTP, 0.2 μM each of the primers CF and CR (Campbell et al., 1993, 2000), and 2 μl *Taq* polymerase. The thermocycler was programed for an initial denaturing step of 3 min at 94 °C; followed by 30 cycles of 45 s at 94 °C, 30 s at 50 °C, and 1 min 30 s at 72 °C; and, a final extension of 30 min at 72 °C.

Microsatellite protocols

Scirtothrips perseae is haplodiploid and all our collections exhibited heavily female-biased sex ratios (PF Rugman-Jones, unpubl.). Therefore, to keep our analyses simple, haploid males were excluded. Data were collected for each of four microsatellite loci (*Sper05*, *Sper06*, *Sper18*, and *Sper19*), from a total of 221 females from seven populations: three introduced samples and four native samples (only five individuals were collected from Tonatico in Mexico and so because of its close proximity to Coatepec Harinas, these were integrated with the latter population). Primer sequences and amplification protocols are given in Rugman-Jones et al. (2005). For each individual specimen, an aliquot of amplified PCR products from all loci was combined in a final volume of 12 μl containing 4.0, 3.0, 5.0, and 1.0 μl of *Sper05*, *Sper06*, *Sper18*, and *Sper19* PCR products, respectively. Two microlitres of this pooled sample were mixed with an internal size standard, GeneScan®-500 [ROX]™ (Applied BioSystems), before separation on an ABI 3100® genetic analyzer (Applied BioSystems). Allele sizes were scored manually using GENEMAPPER® software v. 3.0 (Applied BioSystems). Samples were separated into batches of 96, each batch including four samples used as standards to ensure batch-to-batch consistency in sizing.

Microsatellite analyses

Genetic diversity within populations was assessed in terms of allele frequencies, expected heterozygosity (H_E), observed heterozygosity (H_O), and F_{IS} for each locus and across all loci. Allele frequencies and heterozygosity estimates were obtained using POPGENE 1.32 (Yeh & Boyle, 1997) and F_{IS} was calculated in FSTAT (Goudet, 2001). Deviations from Hardy–Weinberg equilibrium (HWE) were examined using a Markov chain method (Guo & Thompson, 1992) provided in GENEPOP 3.4 (Raymond & Rousset, 1995). P-values were estimated using the default parameters, 1000 dememorization steps, 100 batches, and 1000 iterations per batch. Sequential Bonferroni corrections were applied when multiple tests were performed (Sokal & Rohlf, 1995). Linkage disequilibrium was also investigated using Fisher's exact test as implemented in GENEPOP 3.4. Again, P-values were estimated using the Markov chain method with the default parameters, 1000 dememorization steps, 100 batches, and 1000 iterations per batch.

Genetic variation between populations was assessed by calculating multilocus estimates of F_{ST} for each pair of populations using FSTAT. F_{ST} was used, rather than R_{ST} (Slatkin, 1995), as it is considered a more reliable estimate of population differentiation when using relatively small data sets with fewer than 20 loci (Gaggiotti et al., 1999). Significance of the F_{ST} estimates was evaluated by permuting genotypes among samples (number of permutations was set at 1000) as implemented in the FSTAT program. This method does not assume HWE within populations. Sequential Bonferroni corrections were applied when multiple tests were performed. Nei's unbiased genetic distances (Nei, 1978) were also calculated and a dendrogram of relationships was constructed using the UPGMA method (unweighted pair group method with arithmetic mean) implemented in the program POPGENE 1.32. The resulting dendrogram was redrawn in TREEVIEW v. 1.6.6 (Page, 1996).

The distribution of microsatellite variation was estimated using the analysis of molecular variance (AMOVA; Excoffier et al., 1992) method implemented in ARLEQUIN v. 2.000. Genetic variation was partitioned into three levels: within populations, among populations, and among native and introduced population groups. The significance of population differentiation estimates was evaluated using the permutation method (1000 permutations) invoked in the ARLEQUIN program.

Origin of the Californian population. To trace the origin of the Californian population of *S. perseae* collected in Orange and Ventura Counties in 1997, we applied the assignment/exclusion test as implemented in the program GENECLASS v. 2.0 (Piry et al., 2004). This test uses Monte Carlo simulations to calculate the likelihood of encountering a particular multilocus genotype from an introduced population, in each of a series of potential source populations. The test was run with 10 000 simulated individuals, a population exclusion threshold of 0.01, and the Bayesian method of Paetkau et al. (2004). Such likelihood-based Bayesian methods

have been shown to perform better than frequency-based and distance methods even in cases where the assumptions of HWE are not met (see Cornuet et al., 1999).

Size of the invading population. We tested for evidence of a bottleneck in population size associated with the invasion of California by *S. perseae* using the program BOTTLENECK v. 1.2.02 (Cornuet & Luikart, 1996). This program utilizes the phenomenon that immediately following a 'bottleneck', allelic diversity is lost faster than heterozygosity and does not require HWE. The infinite alleles model (IAM) coupled with a Wilcoxson sign-rank test was implemented, minimizing the number of polymorphic loci required while maximizing the statistical power of the method. The locus *Sper 19* was excluded from the analysis as it is monomorphic in the introduced samples. The BOTTLENECK program is not an infallible method particularly when dealing with small numbers of loci (Cornuet & Luikart, 1996; Luikart & Cornuet, 1998), and for this reason, we also simply compared the number of alleles per locus and H_O between the native and introduced populations, using analyses of variance that included the locus identifier as a covariate (STATVIEW v. 5.0.1; SAS Institute Inc., Cary, NC, USA). Number of alleles per locus was natural log transformed to normalize its distribution prior to analysis.

Evidence of a population size bottleneck was also sought by calculating the effective population size N_e of the invading population, using the moments-based temporal method of Waples (1989), as implemented in the program NEESTIMATOR v. 1.3 (Peel et al., 2004). This method estimates N_e from allele frequency variation across generations (or populations), assuming no migration. If the number of generations between the samples is larger than one, the program calculates the harmonic mean of each generation's N_e. We assumed that allele frequencies in the population from Coatepec Harinas were representative of the ancestral source population from which the California population arose (see Results). Data for Orange County (OC-97) and Ventura County (VC-97) collections, made some 8–10 months after *S. perseae* was first detected in California, were combined to give the introduced population, with a generation gap of 10. This is a conservative estimate based on meteorological data for Irvine (in OC) and Port Hueneme (in VC) covering the period from 1 June 1996 to 31 May 2005, and degree-day models of *S. perseae* generation times (Hoddle, 2002b).

Results

Mitochondrial DNA sequences

We identified 79 polymorphic sites and 14 distinct haplotypes among the 47 mtDNA sequences (Figure 2).

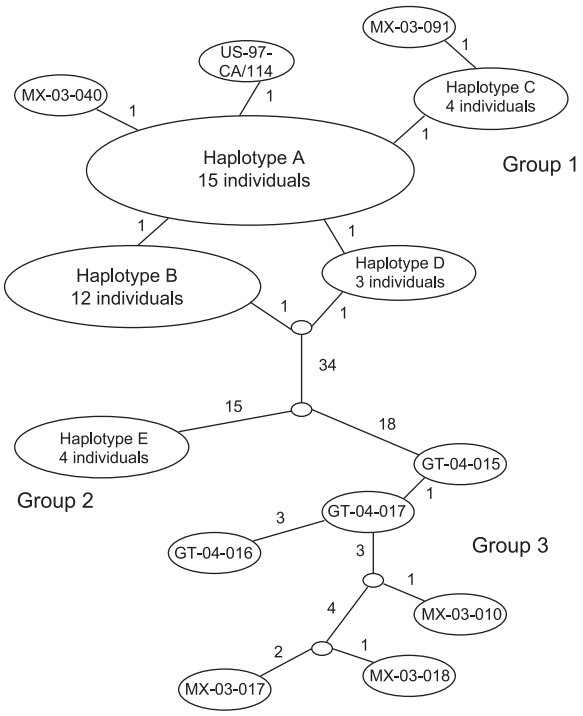

Figure 2 Haplotype network of *Scirtothrips perseae* mitochondrial DNA (mtDNA) sequence data constructed using TCS v. 1.18, revealing three distinct clusters of haplotypes (Groups 1–3). Haplotype A (MX-03-040, MX-03-041, MX-03-044, MX-03-045, MX-03-069, MX-03-070, MX-03-071, MX-03-086, MX-03-096, MX-03-098, MX-03-100, MX-03-101, MX-03-215, US-04-CA/008, and US-04-CA/013); haplotype B (US-97-CA/084, US-97-CA/085, US-97-CA/086, US-97-CA/113, US-97-CA/115, US-97-CA/116, US-04-CA/004, US-04-CA/015, US-04-CA/022, US-04-CA/028, US-04-CA/054, and US-04-CA/059); haplotype C (US-04-CA/034, US-04-CA/035, US-04-CA/040, and US-04-CA/051); haplotype D (MX-03-080, MX-03-081, and MX-03-082); haplotype E (MX-03-108, MX-03-118, MX-03-219, and MX-03-220).

These haplotypes fell into three very divergent mtDNA groups (Tables 2 and 3; Figure 2), and subsequent analyses were performed on these mtDNA groups. Group 1 included the two most common haplotypes (A and B) and included all individuals from the pest's introduced range in California and from Coatepec Harinas, Chilapa de Diaz, Atlixco, and Huayapam in the pest's native range in Mexico. The majority of California specimens fell into haplotype B; and we did not find an exact match in any of the native samples. However, haplotypes A and B differed from each other by only a single nucleotide substitution and all other individuals in this group had haplotypes only one or two steps removed from the most common haplotype. Within Group 1, the number of segregating (polymorphic) sites was seven,

Table 2 DNA sequence variation in a 673-bp segment of the cytochrome *c* oxidase subunit I (COI) region of mitochondrial DNA (mtDNA) of *Scirtothrips perseae*. Average number of pairwise differences within (diagonal element) and between collecting locations (below diagonal). For population abbreviations, see Table 1

	OC-97	VC-97	CA-04	MX-Hu	MX-CD	MX-At	MX-To	MX-CH	MX-Ur	MX-Pe	MX-SC	GT
OC-97	0.000											
VC-97	0.000	1.500										
CA-04	0.318	0.068	1.030									
MX-Hu	1.000	0.500	0.318	0.400								
MX-CD	1.000	0.500	0.318	0.000	0.000							
MX-At	2.000	1.500	1.318	1.000	1.000	0.000						
MX-To	1.000	0.500	0.318	0.000	0.000	1.000	0.000					
MX-CH	1.000	0.500	0.185	0.000	0.000	1.000	0.000	0.800				
MX-Ur	50.000	50.000	50.318	51.000	51.000	50.000	51.000	51.000	0.000			
MX-Pe	50.000	50.000	50.318	51.000	51.000	50.000	51.000	51.000	0.000	0.000		
MX-SC	52.333	51.833	52.652	53.333	53.333	53.000	53.333	52.933	35.667	35.667	5.333	
GT	52.333	51.833	52.652	53.333	53.333	52.333	53.333	52.933	33.333	33.333	3.667	2.667

Table 3 Pairwise F_{ST} estimates (above the diagonal) and their associated significance values after sequential Bonferroni correction (below the diagonal) between the three mitochondrial DNA (mtDNA) groups (see Figure 2), based on a 673-bp section of the cytochrome *c* oxidase subunit I (COI) gene region of *Scirtothrips perseae*

	Group 1	Group 2	Group 3
Group 1		0.981	0.968
Group 2	<0.001		0.895
Group 3	<0.001	0.005	

haplotype diversity was 0.730 ± 0.047 (mean ± SD), nucleotide diversity (π) was 0.00158 ± 0.00022, θ_W per site calculated from the number of segregating sites was 0.00249, and the average number of pairwise nucleotide differences (κ) was 1.060. Group 2 consisted of a single haplotype (haplotype E) including all individuals from Michoacan, Mexico. Hence, all measures of sequence polymorphism were 0. Group 3 was the most diverse of our groupings (no two individuals shared the same haplotype) and included all individuals from San Cristobal de las Casas and Chimaltenango (either side of the Mexico/Guatemala border, respectively). The number of segregating sites was 14, haplotype diversity was 1.000 ± 0.096, π was 0.00921 ± 0.00159, θ_W per site was 0.00911, and κ was 6.200. Sequences of the D2 domain from randomly sampled individuals were identical across mitochondrial haplotypes (GenBank accession numbers: DQ075176–DQ075188).

Despite the high degree of sequence polymorphism between groups, synonymous changes constituted the overwhelming majority of the 79 substitutions in the 673-bp stretch of COI. Only one nucleotide substitution (within Group 1), at position 529, resulted in an amino acid substitution with isoleucine being replaced by leucine in haplotype C and specimen MX-03-091 (Figure 2).

Microsatellite variation

The four microsatellite loci had a total of 22 alleles in the seven populations studied, and the number of alleles was highest in native populations (Table 4). We found no evidence for linkage disequilibrium ($P<0.05$) between loci in any single population or across all populations, and we therefore assume that the loci are independent. Only two of the 28 locus*population combinations showed significant deviation from HWE (MX-Hu, locus *Sper05*; MX-Pu, locus *Sper05*; see Table 1). In each case, significant deviations were the result of an excess of homozygotes as indicated by the positive inbreeding coefficient F_{IS} (Table 5). The presence of a null allele could have contributed to these departures from HWE, but they may also be a result of the modest sample sizes for these populations, which may have lowered the probability of sampling all possible genotypes at that locus.

Pairwise multilocus estimates of F_{ST} revealed significant differentiation between the introduced and native populations and between each native population pair (Table 6). However, F_{ST} values for two of the three pairwise comparisons between the introduced samples were not significantly different from 0 (the third was only marginally significant). The lowest computed F_{ST} value (0.0021) was between the original introduced samples from OC and VC.

Nei's unbiased genetic distance (Nei, 1978) between populations ranged from 0.0022 to 0.2923 with those populations from the introduced range showing the least differentiation and being more closely related to the population from Coatepec Harinas than any other source population (Table 6). The UPGMA dendrogram indicated

Table 4 Allele frequencies at four microsatellite loci for each population of *Scirtothrips perseae*

Locus	Size (bp)	CA-04	VC-97	OC-97	MX-CD	MX-CH[1]	MX-Hu	MX-Pu
Sper05	170	–	–	–	–	0.013	–	–
	171	0.041	0.020	0.040	0.031	0.316	0.062	0.031
	172	0.008	–	–	–	–	–	–
	173	0.697	0.714	0.620	0.438	0.329	0.156	0.438
	174	0.131	0.051	0.080	0.406	0.013	0.188	0.312
	176	0.123	0.194	0.120	0.031	0.211	0.281	0.062
	178	–	0.020	0.140	0.062	0.013	0.031	0.094
	180	–	–	–	–	0.039	0.062	–
	182	–	–	–	0.031	0.026	0.094	–
	183	–	–	–	–	–	–	0.031
	184	–	–	–	–	–	0.062	0.031
	186	–	–	–	–	0.039	0.062	–
Sper06	193	–	–	–	–	–	0.031	–
	194	–	–	–	–	0.013	0.062	–
	195	0.287	0.357	0.360	0.719	0.303	0.719	0.031
	196	0.713	0.643	0.640	0.281	0.684	0.156	0.969
	197	–	–	–	–	–	0.031	–
Sper18	293	0.762	0.633	0.680	0.906	0.711	1.000	0.750
	295	0.238	0.367	0.320	0.094	0.289	–	0.250
Sper19	269	–	–	–	–	–	0.031	–
	270	–	–	–	–	0.013	–	–
	271	1.000	1.000	1.000	1.000	0.987	0.969	1.000

[1]MX-CH incorporates individuals from Coatepec Harinas and Tonatico, Mexico.
Sample sizes: CA-04, n = 61; VC-97, n = 49; OC-97, n = 25; MX-CD, n = 16; MX-CH, n = 38; MX-Hu, n = 16; MX-Pu, n = 16.

Table 5 Test of conformity of the microsatellite loci to Hardy–Weinberg equilibrium (HWE) in each population of *Scirtothrips perseae*

Locus	CA-04 (n = 61)	VC-97 (n = 49)	OC-97 (n = 25)	MX-CD (n = 16)	MX-CH (n = 38)	MX-Hu (n = 16)	MX-Pu (n = 16)
Sper05							
H_E	0.485	0.453	0.585	0.657	0.753	0.863	0.718
H_O	0.443	0.449	0.520	0.625	0.658	0.563	0.313
F_{IS}	0.087	0.010	0.114	0.051	0.128	**0.356**	**0.573**
Sper06							
H_E	0.413	0.464	0.470	0.417	0.446	0.468	0.063
H_O	0.344	0.429	0.480	0.313	0.395	0.438	0.063
F_{IS}	0.167	0.077	–0.021	0.257	0.116	0.067	0.000
Sper18							
H_E	0.365	0.470	0.444	0.175	0.417	NA	0.387
H_O	0.443	0.612	0.560	0.066	0.368	NA	0.250
F_{IS}	–0.213	–0.308	–0.268	0.651	0.118	NA	0.362
Sper19							
H_E	NA	NA	NA	NA	0.026	0.063	NA
H_O	NA	NA	NA	NA	0.026	0.063	NA
F_{IS}	NA	NA	NA	NA	0.000	0.000	NA
All loci							
H_E	0.316	0.347	0.375	0.313	0.411	0.348	0.292
H_O	0.307	0.372	0.390	0.250	0.362	0.266	0.156
F_{IS}	0.026	–0.075	–0.041	0.205	0.120	**0.243**	**0.473**

NA, monomorphic.
Values in bold represent significant deviations after sequential Bonferroni correction.

Table 6 Pairwise F_{ST} estimates (above the diagonal) and Nei's unbiased genetic distances (below the diagonal) based on multilocus microsatellite genotypes between seven populations of *Scirtothrips perseae*

	CA-04	VC-97	OC-97	MX-CD	MX-CH	MX-Hu	MX-Pu
CA-04		**0.0122**	0.0055	0.1704	0.0698	0.2597	0.0712
VC-97	0.0080		**0.0021**	0.1784	0.0693	0.2544	0.1165
OC-97	0.0052	0.0022		0.1352	0.0438	0.2110	0.0806
MX-CD	0.1044	0.1212	0.0941		0.1684	0.0520	0.2704
MX-CH	0.0429	0.0482	0.0362	0.1362		0.1785	0.0888
MX-Hu	0.1906	0.2059	0.1744	0.0388	0.1572		0.3381
MX-Pu	0.0404	0.0720	0.0532	0.1895	0.0643	0.2923	

F_{ST} values in bold are not significantly different from 0 after sequential Bonferroni correction.

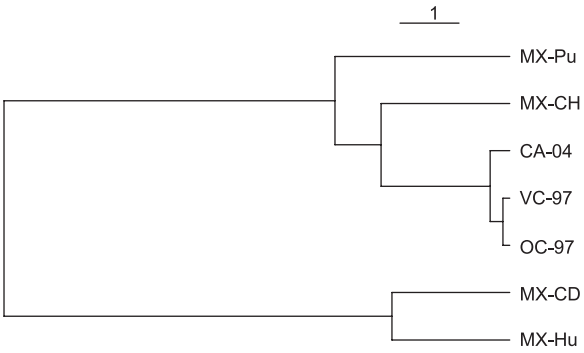

Figure 3 Dendrogram of relationships between native and introduced populations of *Scirtothrips perseae* based on Nei's unbiased genetic distance. For abbreviations, see Table 1.

the existence of two groups with the Atlixco and Chilapa de Diaz populations together being distinct from the remaining populations (Figure 3).

Analysis of molecular variance (Table 7) revealed that the greatest amount of total microsatellite variation (88.48%) was accounted for by differences among individuals within populations. A much smaller but significant amount of variation (8.66%) was found among populations within groups, but variation between native and introduced groups was not significant.

With an exclusion threshold of 0.01, our assignment/exclusion test did not allow us to designate a single native population as the source of any of the introduced individuals. Indeed, only six individuals were excluded from any single population; two from Huayapam and four from Atlixco. However, removing the exclusion threshold allows individuals to be assigned to the population with the highest probability (Table 8). On this basis, the probability of an introduced individual originating from Coatepec Harinas was significantly higher than the probabilities that it originated from any of the other native populations (Friedman test: $\chi^2 = 29.32$, d.f. = 3, $P<0.001$; Table 8).

Table 7 Analysis of molecular variance for seven populations of *Scirtothrips perseae* from its native (Mexico) and introduced (California) range

	F-statistics	Percentage of variation	d.f.	P
Among groups	$F_{CT} = 0.029$	2.86	1	0.097
Among populations within groups	$F_{SC} = 0.089$	8.66	5	<0.001
Within populations	$F_{ST} = 0.115$	88.48	435	<0.001

Total diversity is partitioned among the three levels being compared.
F_{CT} is the fixation index of the native and introduced population groups relative to the total combined population.
F_{SC} is the fixation index of individual populations relative to the population groups.

Table 8 Assignment of *Scirtothrips perseae* individuals collected from Ventura County and Orange County in 1997 to their potential source populations. Individuals are assigned to the population with the highest probability, but the other populations could not be excluded as a source of those individuals

Collected from	MX-CD	MX-CH	MX-Hu	MX-Pu
Ventura County	6 (0.213)	28 (0.518)	1 (0.200)	13 (0.310)
Orange County	3 (0.241)	12 (0.425)	3 (0.204)	7 (0.277)
Combined	9 (0.223)	40 (0.487)	4 (0.201)	20 (0.299)

Values in parentheses are mean probabilities (calculated from all individuals).

However, it must be stressed that we can not definitively rule out the other populations as potential sources.

The program BOTTLENECK failed to detect any population with the significant heterozygote excess associated with a strong bottleneck in population size. The BOTTLENECK program utilizes the phenomenon that immediately

following a 'bottleneck', allelic diversity is lost faster than heterozygosity due to the rapid loss of rare alleles, which have little effect on heterozygosity (Cornuet & Luikart, 1996). Simple analyses of variance comparing allelic diversity between introduced and native populations also failed to detect any significant reduction in the number of alleles per locus ($F_{1,17}$ = 2.342, P = 0.144) or increase in observed heterozygosity H_O ($F_{1,17}$ = 2.650, P = 0.216). However, the number of alleles at the two most polymorphic loci clearly appears to be reduced in the introduced samples (CA-04, VC-97, and OC-97; Table 4), and the lost alleles corresponded to those that occurred at the lowest frequencies in the native populations (i.e., the rarest alleles) as would be expected following a population bottleneck (Cornuet & Luikart, 1996).

The harmonic mean effective population size N_e of the invading Californian population was estimated as 64.8 (95% confidence limits: 20.4–167.6). Estimates of N_e are typically smaller than the actual effective population size and this discrepancy is likely to increase given the female-biased sex ratio inherent in *S. perseae* populations (Frankham, 1995; Hoddle, 2002b). Conversely, our estimate of N_e is the harmonic mean over 10 generations and therefore, assuming population expansion, is likely to be higher than the true N_e at the time *S. perseae* arrived in California. Assuming a gap of only one generation yields an estimate for N_e of 6.5 (95% confidence limits: 2.0–16.8).

Discussion

Using a combination of conserved and highly polymorphic genetic markers, we have been able to identify that the California population of the exotic pest *S. perseae* was most likely the result of a single invasion event, initiated by a small number of individuals arriving from Coatepec Harinas, Mexico. Our measures of population subdivision among native and introduced samples showed significant population structure at both mtDNA and microsatellite markers.

Variation in mtDNA sequences

Across its native range, *S. perseae* exhibited unusually high genetic mtDNA differentiation, with three distinct haplotype lineages. These lineages appear to correlate well with the geographic location of each sampled population within the native distribution. The haplotypes of all sampled individuals from Coatepec Harinas, Tonatico, Chilapa de Diaz, Atlixco, and San Andreas, Huayapam (neighboring areas in the center of the insect's native range; Figure 1) differed by only 1–3 nucleotide substitutions. However, individuals from the state of Michoacan to the North and around the Mexico/Guatemala border to the South differed from the first group by 7.4 and 7.9%, respectively, at the sequence level. The Michoacan group contained only a single haplotype whereas the latter group displayed the highest level of within-group sequence divergence (~0.6%). These two groups were also about 5.1% divergent from one another. Surprisingly, the haplotypes of the introduced California specimens were not identical to any of the native samples. However, they were clearly related to a single mtDNA lineage, Group 1 (Figure 2), and all differed by only a single nucleotide substitution from the most common native haplotype (haplotype A) in this group. Given the vast differences between the California samples and the other two mtDNA lineages, it appears that the California population of *S. perseae* originated from one (or more) of the native populations contained within Group 1 (Figure 2). While it is possible that we failed to sample the 'real' source population, our failure to detect identical haplotypes between the native and introduced samples is probably a result of our relatively small sample sizes and/or problems with the genetic relatedness of the individuals sampled.

Divergence between the three mtDNA lineages of *S. perseae* was exceptionally high (5.1–7.9%). Similar studies across the native and introduced ranges of other haplodiploid taxa have revealed maximum levels of intraspecific mtDNA sequence divergence of around 1–2% (Tsutsui et al., 2001; Gyllenstrand & Seppä, 2003; Sanetra & Crozier, 2003; Hufbauer et al., 2004). Perhaps more surprisingly, in the present study there appears to be no geographic mixing of *S. perseae* individuals from the different mtDNA lineages. Again, this may be a reflection of our small sample sizes, particularly in Groups 2 and 3 with only four and six individuals, respectively. However, in a preliminary study using our microsatellite loci, a larger sample from Groups 2 and 3 (n = 15 and n = 26, respectively) consistently failed to amplify at the largely monomorphic locus *Sper19* (see Rugman-Jones et al., 2005), suggesting that mtDNA differences may similarly be consistent throughout these groups (PF Rugman-Jones, unpubl.).

There are several evolutionary explanations for this pattern of mtDNA differentiation. First, populations within each mtDNA lineage may be geographically isolated from populations in other lineages, such that immigration is non-existent or very rare. Although *S. perseae* adults are relatively weak dispersers (Hoddle et al., 2002b), this seems unlikely because the host plant, avocado, is commercially moved within Mexico from plant nurseries to orchards (MS Hoddle, pers. obs.). It is highly unlikely that all transported plants are free from *S. perseae* eggs, larvae, or even adults. However, most are treated with extremely powerful insecticides such as methyl parathion (MS Hoddle, pers. obs.) and this may greatly limit spread from commercial producers and suppliers. Alternatively, the level of divergence

may suggest that the different mtDNA lineages are in fact different species. A recent revision of the genus *Scirtothrips* in Mexico listed seven species from avocado (Johanson & Mojica-Guzman, 1998), five of which were newly described. The validity of these descriptions has been challenged, as species designations were made based on morphological characters that are known to be highly variable within *Scirtothrips* species (Mound & zur Strassen, 2001). This idea that cryptic species exist is opposed by the fact that a 539-bp section of the D2 domain of rDNA was identical across all specimens studied. In a separate study, pairwise comparisons of the section of the D2 domain of 19 different *Scirtothrips* species showed sequence divergence ranging from around 0.6–11.0% (mean ± SE 8.5% ± 0.2; MS Hoddle, PF Rugman-Jones & R Stouthamer, unpubl.).

While the three mtDNA lineages may not be separate species, the levels of sequence divergence suggest that some degree of reproductive isolation exists between the lineages, and has existed for some considerable time. A third explanation for the pattern we have observed is offered by the endosymbiotic bacterium *Wolbachia*. These bacteria infect a wide variety of arthropods and filarial nematodes and manipulate the reproduction of their host, enhancing their own transmission, by a variety of mechanisms, including cytoplasmic incompatibility (CI) (Werren & O'Neill, 1997; Stouthamer et al., 1999). Under CI, *Wolbachia* modify the sperm of infected males in such a way that prevents fertilization from occurring unless the egg (being fertilized) is also infected with the same *Wolbachia* strain (Werren & O'Neill, 1997). Thus, incompatibility can occur between the sperm of infected males and the eggs of uninfected females, or between the sperm of individuals infected with one strain of *Wolbachia* and the eggs of individuals infected with a different strain. As a result, uninfected individuals suffer reduced fitness and the *Wolbachia* infection spreads to fixation in a population and, if different *Wolbachia* strains infect different populations, leads to reproductive isolation between populations. Initial data suggest that such a scenario may be active in the *S. perseae* mtDNA lineages. A small sample of only four individuals from each lineage suggests that Group 1 is partially infected with one strain of *Wolbachia*, a different *Wolbachia* strain has spread to fixation in Group 2, and Group 3 appears to be free from infection (PF Rugman-Jones, unpubl., GenBank accession numbers: DQ075189–075198). Thus, any immigrant into a population with a different mtDNA lineage would be strongly selected against. Such a system has been suggested to account for relatively high levels of interpopulation mtDNA divergence in the fire ant *Solenopsis invicta* (Shoemaker et al., 2003). It is surprising that Group 3 appears to have remained distinct (and free from *Wolbachia* infection), as in theory only a single *Wolbachia*-infected immigrant from one of the other two mtDNA lineages would be needed to begin the erosion of this distinction. Perhaps some other bacterial endosymbiont is preventing the introduction of the other haplotype lineages (Zchori-Fein et al., 2004).

Wolbachia infection may also account for the different levels of haplotype diversity within the three mtDNA lineages. *Wolbachia* infections are commonly associated with a reduction or complete loss of mtDNA haplotype diversity within a population (Turelli et al., 1992). Following the infection of a female in a population, the spread of CI *Wolbachia* entails a concomitant spread of all maternally inherited organelles, including mitochondria, found in the initially infected female (Turelli & Hoffmann, 1991). Thus, one particular mtDNA haplotype sweeps through a population, via hitchhiking, in association with the selectively driven sweep of *Wolbachia* (Turelli et al., 1992). Such an event may have taken place relatively recently in Group 2, accounting for the presence of only a single mtDNA haplotype. Such an event may also have taken place in Group 1 but long enough ago for a degree of haplotype diversity to have been restored following the sweep. Alternatively, the *Wolbachia* in this population may no longer cause CI allowing the haplotype diversity to grow. Similarly, the lack of infection in Group 3 may explain why this group exhibits the highest levels of mtDNA diversity.

Whether or not the three mtDNA lineages represent different species or isolated races, justification for sampling only native populations from Group 1 in our microsatellite study comes from the fact that each of the other two groups (Figure 2) harbors only haplotypes that are extremely divergent from the haplotypes in Group 1 and hence from those detected in California. To test whether these lineages constitute species would require a series of crossing experiments in which all pairwise population crosses are performed. If the crosses fail, the lineages are reproductively isolated. If the *Wolbachia* infection is then cured by treatment with antibiotics (Breeuwer & Werren, 1990) (i.e., the potential isolating mechanism is removed) and the crosses still fail, we will have good evidence that these mtDNA lineages do indeed represent different species.

Variation in microsatellite loci

Differences in allele frequencies provide a tool to uncover patterns of population differentiation. The present study shows high levels of genetic variation in both introduced and native samples of *S. perseae* (Tables 4 and 5). Microsatellite allele frequency data revealed significant population differentiation among the native populations and between each native population and each introduced population (Table 6). However, by far the majority of genetic variation was among individuals within populations, and overall

there was no difference between the introduced and native populations as a whole (Table 7). The three introduced samples were essentially genetically identical (Table 6). The fact that the CA-04 sample is marginally differentiated from one of the two original introduced samples (OC-97) is probably the result of the evolution of a unique allele at the *Sper05* locus in the CA-04 population (Table 4). The similarity between the introduced samples suggests that the California population was seeded by a single introduction from a single source population. Estimates of F_{ST} and Nei's genetic distances agree in identifying the native population from Coatepec Harinas as the most genetically similar to the introduced populations (Table 6), and this is reflected in the UPGMA dendrogram (Figure 3). The similarity between *S. perseae* from California and Coatepec Harinas was also supported by the results of our assignment/exclusion test (Cornuet et al., 1999). Statistically, California individuals were more likely to originate from Coatepec Harinas than any of the other native populations tested, although we were unable to definitively exclude the possibility that the other native populations could be possible sources. As it is rarely possible to sample all the genetic subdivisions in the native range of a species, this finding may indicate that we failed to sample the 'true' source population and Coatepec Harinas is simply the closest match, of those populations sampled, to the 'true' source population. However, even if Coatepec Harinas is the 'true' source population, our finding is perhaps not surprising as human movement of the host plant is likely to foster extensive genetic homogenization through gene flow among native populations (or at least those which share a common mtDNA lineage; see above). That said, the identification of Coatepec Harinas as the most likely source of the California population fits well with our knowledge of the native populations. Coatepec Harinas is the site of a large avocado germplasm bank and breeding station, the Fundación Salvador Sánchez Colín-CICTAMEX S.C., which is regularly visited by foreign researchers and growers (MS Hoddle, pers. obs.). In contrast, the other populations we identified and sampled lie in areas that are much less likely to draw attention from international travelers with interest in avocado breeding and novel cultivars.

The colonization of a new area by an introduced species is often associated with a loss of genetic variation as a result of population size bottlenecks (i.e., a small number of founders) and genetic drift in the founding population. This is typically followed by a reduction in heterozygosity. However, recently bottlenecked populations typically exhibit higher than usual levels of heterozygosity as rare alleles are quickly lost from the bottlenecked population without affecting heterozygosity (Cornuet & Luikart, 1996). *Scirtothrips perseae* was first detected, in relatively small numbers, attacking avocados at two distant locations in California in July 1996, after which it quickly spread across almost the entire avocado growing region of California (Hoddle & Morse, 2003). Despite this knowledge, the present study provided no statistical evidence that the introduced California population had undergone a recent bottleneck. While invasions without a bottleneck may occur (e.g., Holland, 2001), there is good reason to doubt the validity of this finding. The BOTTLENECK analysis we employed (Cornuet & Luikart, 1996) typically requires more than 10 polymorphic loci and sample sizes of around 30 to have a reasonable chance of detecting a bottleneck (Luikart et al., 1998). In the present study the sample size of each population reflects the numbers of *S. perseae* specimens we were able to collect at that location, and while our mean sample size was 31.6, four of our seven samples contained only 16 individuals each. Furthermore, the number of loci we were able to identify, and hence use, for this species was low (Rugman-Jones et al., 2005). The low number of loci is also likely to account for why we did not detect any significant drop in the number of alleles per locus using analysis of variance. Examination of Table 4 clearly suggests that the rare alleles present at three of the four loci (*Sper05*, *Sper06*, and *Sper19*) in the native populations have been lost in the California samples, as predicted following a bottleneck. Furthermore, the remaining locus (*Sper18*), which had only two alleles across all native and introduced samples, shows signs of heterozygote excess, evident in the negative F_{IS} values for each of the California samples (Table 5).

Perhaps the best evidence we have of a population bottleneck is our estimate of effective population size for the original invading population. Using a method that assumes no migration, our estimate of the harmonic mean N_e over the first 10 generations following the introduction was 64.8 (confidence limits: 20.4–167.6). Thus, accounting for the rapid population expansion that has occurred in California (Hoddle & Morse, 2003), the number of genetically distinct breeding females that initiated the California (i.e., generation 0) population is likely to be smaller than our harmonic mean value. While such estimates of N_e are typically smaller than the actual effective population size (i.e., the number of breeding females; Frankham, 1995), our estimate is in line with a severe population bottleneck. Care should be taken interpreting our estimate as we have inferred starting allele frequencies from a population that may or may not be the 'true' source population, i.e., Coatepec Harinas may just be the population sampled that is genetically closest to an unsampled 'true' source population. However, while this is likely to introduce a degree of error into our estimation of N_e, we believe our model provides a reasonable 'ballpark' estimate for the size of the invading population.

Allowing for genetic relatedness of the invaders and a female-biased sex ratio, our estimate of N_e also suggests that the total number of thrips initiating the invasion of California was quite small. A single avocado leaf may contain more than 30 eggs (Hoddle, 2002a) and *S. perseae* females typically lay 20–40 eggs during their short lives (Hoddle, 2002b). Given that *S. perseae* are weak fliers, we may expect that females do not disperse far during the oviposition period (i.e., a single female lays all her eggs into just a few adjacent leaves). If we take for granted that *S. perseae* arrived in California as an incidental passenger on smuggled avocado plants, then from our estimate of N_e, we may predict that the number of host plants smuggled into California was also small. Indeed, it is quite possible that the material entering California constituted no more than a branch or two that were quickly grafted onto California rootstock. Regardless of whether the graft was successful, such material may support eggs (laid within leaves) long enough to produce viable larvae (the oviposition to egg hatch stage in *S. perseae* is around 10 days at favorable temperatures; Hoddle, 2002b) that could crawl onto neighboring live host material. Further, smuggled plant material may support thrips larvae and adults. These motile stages are usually very small, extremely thigmotactic, and low numbers are almost impossible to detect without magnification equipment. Work on *Sericothrips staphylinus*, a thrips used for the biological control of a noxious weed in New Zealand, has demonstrated that a third of carefully managed releases of just 10 adult thrips into a favorable environment resulted in establishment and proliferation (Memmott et al., 1998).

Implications for pest management

The use of molecular genetics to reconstruct the history of agricultural pest invasions and to identify the source of introduced populations is a major step forward in the management of such pests. For instance, such information may facilitate the identification of effective biological control agents for introduced species. Previous exploration efforts to identify biological control agents for *S. perseae* have been largely unproductive (Hoddle et al., 2002a). However, the present study has identified a cluster of populations in the center of the native range of *S. perseae* as being genetically similar to the California population. Within this cluster, Coatepec Harinas has been identified as the most likely source of the California population. In theory, natural predators and parasitoids from Coatepec Harinas should be better adapted to California *S. perseae* than such biological control agents from other native regions (e.g., Guatemala) and future efforts to identify efficacious host-specific natural enemies should focus on this area.

The establishment of *S. perseae* in California avocado orchards was preceded by the establishment of two other avocado pests, *T. perseae* and *Oligonychus perseae*, which are also native to Mexico and have very similar biological requirements to *S. perseae*. Another phytophagous thrips, *Neohydatothrips burungae*, which is not yet present in California, is also widespread in Mexico and has been identified as a potential serious avocado pest should it become established in California (Hoddle et al., 2002a). In identifying Coatepec Harinas as the source of the California population of *S. perseae*, we may have uncovered a common route of entry into California for other resident avocado pests. Identification of the source of invasive pests enables government agencies to establish and enforce fair scientifically based plant quarantine and free-trade protocols.

Acknowledgements

Financial support for this work came from the University of California Exotic/Invasive Pests and Diseases Research Program. We are indebted to Phil A. Phillips for assistance in collecting thrips in Mexico and Central America; Laurence A. Mound for assistance with specimen verification; and the staff of the Core Instrumentation Facility of the University of California Riverside Genomics Institute for assistance with sequencing and genotyping work. We also thank Andrew Weeks for his efforts during the early stages of microsatellite isolation and our colleagues listed in Table 1, whose collecting efforts in California have contributed greatly to this study.

References

Bender G (1993) A new mite problem on avocados. California Avocado Society Yearbook 77: 73–77.

Breeuwer JAJ & Werren JH (1990) Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. Nature 346: 558–560.

California Avocado Commission (2004) Mexican Avocados in Amercia: The Epic Battle to Protect California Growers. California Avocado Commission, Irvine, CA, USA.

Campbell B, Heraty JM, Rasplus JY, Chan K, Steffen-Campbell J & Babcock C (2000) Molecular systematics of the Chalcidoidea using 28s-D2 rDNA. The Hymenoptera: Evolution, Biodiversity and Biological Control (ed. by A Austin & M Dowton), pp. 57–71. CSIRO Publishing, Victoria, Australia.

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 28s rDNA sequences. Insect Molecular Biology 2: 255–237.

Clement MD, Posada D & Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657–1660.

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.

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

Excoffier L (2004) Special issue: analytical methods in phylogeography and genetic structure. Molecular Ecology 13: 727.

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.

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. Molecular Marine Biology and Biotechnology 3: 294–299.

Frankham R (1995) Effective population size/adult population size ratios in wildlife: a review. Genetical Research 66: 95–107.

Gaggiotti OE, Lange O, Rassmann K & Gliddon C (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. Molecular Ecology 8: 1513–1520.

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. Available at: http://www.unil.ch/izea/softwares/fstat.html.

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

Gyllenstrand N & Seppä P (2003) Conservation genetics of the wood ant, *Formica lugubris*, in a fragmented landscape. Molecular Ecology 12: 2931–2940.

Hallman GJ & Schwalbe CP (2002) Invasive Arthropods in Agriculture: Problems and Solutions. Science Publishers Inc, Enfield, NH, USA.

Hoddle MS (2002a) Oviposition preference of *Scirtothrips perseae* Nakahara (Thysanoptera: Thripidae) in southern California avocado orchards. Pan Pacific Entomologist 78: 177–183.

Hoddle MS (2002b) Developmental and reproductive biology of *Scirtothrips perseae* (Thysanoptera: Thripidae): a new avocado pest in California. Bulletin of Entomological Research 92: 279–285.

Hoddle MS, Aponte O, Kerguelen V & Heraty J (1999) Biological control of *Oligonychus perseae* (Acari: Tetranychidae) on avocado. I. Evaluating release timings, recovery, and efficacy of six commercially available phytoseiids. International Journal of Acarology 25: 211–219.

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

Hoddle MS & Morse JG (2003) Avocado Thrips Biology and Control. California Avocado Commission, Irvine, CA, USA.

Hoddle MS, Nakahara S & Phillips PA (2002a) Foreign exploration for *Scirtothrips perseae* Nakahara (Thysanoptera: Thripidae) and associated natural enemies on avocado (*Persea americana* Miller). Biological Control 24: 251–265.

Hoddle MS, Robinson L & Morgan D (2002b) Attraction of thrips (Thysanoptera: Thripidae and Aeolothripidae) to colored sticky cards in a California avocado orchard. Crop Protection 21: 383–388.

Hoddle MS & Soliman GN (2001) Developmental and reproductive biology of the red-banded whitefly, *Tetraleurodes perseae* Nakahara (Homoptera: Aleyrodidae). Subtropical Fruit News 8: 15–18.

Holland BS (2001) Invasion without a bottleneck: microsatellite variation in natural and invasive populations of the brown mussel *Perna perna* (L). Marine Biotechnology 3: 407–415.

Hufbauer RA, Bogdanowicz SM & Harrison RG (2004) The population genetics of a biological control introduction: mitochondrial DNA and microsatellite variation in native and introduced populations of *Aphidius ervi*, a parasitoid wasp. Molecular Ecology 13: 337–348.

Jetter K (1999) Case studies: citrus canker; avocado thrips and mites. Exotic Pests and Diseases: Biology, Economics, Public Policy (ed. by RH Coppock & M Kreith), pp. 124–129. University of California Agricultural Issues Center, Davis, CA, USA.

Johansen RM & Mojica-Guzman A (1998) The genus *Scirtothrips* Shull, 1909 (Thysanoptera: Thripidae, Sericothripini), in Mexico. Folia Entomologica Mexicana 104: 23–108.

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

Luikart G & Cornuet J-M (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology 12: 228–237.

Luikart G & England PR (1999) Statistical analysis of microsatellite DNA data. Trends in Ecology and Evolution 14: 253–256.

Memmott J, Fowler SV & Hill RL (1998) The effect of release size on the probability of establishment of biological control agents: gorse thrips (*Sericothrips staphylinus*) released against gorse (*Ulex europaeus*) in New Zealand. Biocontrol Science and Technology 8: 103–115.

Mound LA & Marullo R (1996) The Thrips of Central and South America: An Introduction. International Memoirs on Entomology 6: 1–487.

Mound LA & zur Strassen R (2001) The genus *Scirtothrips* (Thysanoptera: Thripidae) in Mexico: a critique of the review by Johansen & Mojica-Guzman (1998). Folia Entomologica Mexicana 40: 133–141.

Nakahara S (1997) *Scirtothrips perseae* (Thysanoptera: Thripidae), a new species infesting avocado in southern California. Insecta Mundi 11: 189–192.

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

Paetkau D, Slade R, Burden M & Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology 13: 55–65.

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12: 357–358.

Peel D, Ovenden JR & Peel SL (2004) NeEstimator: Software for Estimating Effective Population Size, version 1.3. Queensland Government, Department of Primary Industries and Fisheries, Deception Bay, Queensland, Australia.

Pimentel D, Lach L, Zuniga R & Morrison D (2000) Environmental and economic costs of nonindigenous species in the United States. Bioscience 50: 53–65.

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L & Estoup A (2004) GeneClass2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity 95: 536–539.

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

Reitz SR & Trumble JT (2002) Competitive displacement among insects and arachnids. Annual Review of Entomology 47: 435–465.

Rose M & Woolley JB (1984) Previously imported parasite may control invading whitefly. California Avocado Society Yearbook 68: 127–131.

Rozas J, Sanchez-DelBarrio JC, Messeguer X & Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496–2497.

Rugman-Jones PF, Hoddle MS, Mound LA & Stouthamer R (2006) Molecular identification key for the pest species of *Scirtothrips* (Thysanoptera: Thripidae). Journal of Economic Entomology 99: 1813–1819.

Rugman-Jones PF, Week AR, Hoddle MS & Stouthamer R (2005) Isolation and characterization of microsatellite loci in the avocado thrips *Scirtothrips perseae* (Thysanoptera: Thripidae). Molecular Ecology Notes 5: 644–646.

Sanetra M & Crozier RH (2003) Patterns of population subdivision and gene flow in the ant *Nothomyrmecia macrops* reflected in microsatellite and mitochondrial DNA markers. Molecular Ecology 12: 2281–2295.

Schneider S, Roessl D & Excoffier L (2000) Arlequin, version 2.000: A Software for Population Genetics Data Analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

Shoemaker DD, Keller G & Ross KG (2003) Effects of *Wolbachia* on mtDNA variation in two fire ant species. Molecular Ecology 12: 1757–1771.

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

Sokal RR & Rohlf FJ (1995) Biometry. The Principles and Practice of Statistics in Biology Research. WH Freeman, New York, NY, USA.

Stouthamer R, Breeuwer JAJ & Hurst GDD (1999) *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. Annual Review of Microbiology 53: 71–102.

Sunnocks P & Hales DF (1996) Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in Aphids of the genus *Sitobion* (Hemiptera: Aphididae). Molecular Biology and Evolution 13: 510–524.

Tsutsui ND, Suarez AV, Holway DA & Case TJ (2001) Relationships among native and introduced populations of the Argentine ant (*Linepithema humile*) and the source of introduced populations. Molecular Ecology 10: 2151–2161.

Turelli M & Hoffmann AA (1991) Rapid spread of an inherited incompatibility factor in California *Drosophila*. Nature 353: 440–442.

Turelli M, Hoffmann AA & McKechnie SW (1992) Dynamics of cytoplasmic incompatibility and mtDNA variation in natural *Drosophila simulans* populations. Genetics 132: 713–723.

Waples RS (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics 121: 379–391.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. Theoretical Population Biology 7: 256–276.

Werren JH & O'Neill SL (1997) The evolution of heritable symbionts. Influential Passengers: Inherited Microorganisms and Arthropod Reproduction (ed. by SL O'Neill, AA Hoffmann & JH Werren), pp. 1–41. Oxford University Press, New York, NY, USA.

Yeh FC & Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belgian Journal of Botany 129: 157.

Zchori-Fein E, Perlman SJ, Kelly SE, Katzir N & Hunter MS (2004) Characterization of a 'Bacteroidetes' symbiont in *Encarsia* wasps (Hymenoptera: Aphelinidae): proposal of '*Candidatus Cardinium hertigii*'. International Journal of Systematic and Evolutionary Microbiology 54: 961–968.