Population Genetics of Invasive *Bemisia tabaci* (Hemiptera: Aleyrodidae) Cryptic Species in the United States Based on Microsatellite Markers

AARON M. DICKEY,[1] LANCE S. OSBORNE,[2] ROBERT G. SHATTERS, JR.,[1] PAULA M. HALL,[2] AND CINDY L. MCKENZIE[1,3]

J. Econ. Entomol. 106(3): 1355–1364 (2013); DOI: http://dx.doi.org/10.1603/EC12512

ABSTRACT The *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) cryptic species complex of whiteflies contains two species, MEAM1 and MED, that are highly invasive in supportive climates the world over. In the United States, MEAM1 occurs both in the field and in the greenhouse, but MED is only found in the greenhouse. To make inferences about the population structure of both species, and the origin and recent spread of MED within the United States, 987 MEAM1 whiteflies and 340 MED whiteflies were genotyped at six and seven microsatellite loci, respectively, for population genetic analyses. Major results of the study are 1) MED exhibits more population structure and genetic differentiation than MEAM1, 2) nuclear microsatellite markers exhibit a high degree of concordance with mitochondrial markers recovering a major east–west phylogeographic break within MED, 3) both eastern and western MED are found throughout the continental United States and eastern MED is present in Hawaii, and 4) MEAM1 contains two greenhouse U.S. populations significantly differentiated from other U.S. MEAM1. The results suggest that MED was introduced into the United States on at least three occasions and rapidly spread throughout the United States, showing no discernible differentiation across 7,000 km. The results further suggest that there is an enhanced role of the protected agricultural environment in promoting genetic differentiation in both invasive B. tabaci cryptic species.

KEY WORDS *Bemisia tabaci,* invasive species, population genetics

The sweetpotato whitefly, *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae), is a cryptic species complex containing at least 24 morphologically indistinguishable species (Dinsdale et al. 2010, De Barro et al. 2011). Two highly polyphagous members of the complex, one originating in the Middle East Asia Minor (MEAM1) and the second originating in the Mediterranean (MED) (De Barro et al. 2011), have become globally invasive crop pests (Oliveira et al. 2001, Perring 2001, Ren et al. 2001, Horowitz et al. 2003, Viscarret et al. 2003, Chu et al. 2006, Ueda and Brown 2006, McKenzie et al. 2012). The native ranges of both cryptic species may extend into sub-Saharan Africa (De Barro and Ahmed 2011, De Barro et al. 2011), and MEAM1 now occurs in >50 countries and MED in 10 countries (De Barro et al. 2011). Damage estimates exceed a billion U.S. dollars over the past 30 yr in cash crops such as cotton, Gossypium hirsutum L., and plant diseases vectored by MEAM1 and MED can be limiting factors to crop production in developing countries (Oliveira et al. 2001).

In the United States, MED was first documented in 2004 (Dennehy et al. 2010) and has since established as a pest in greenhouses (McKenzie et al. 2009, 2012). Importantly, MED in the United States has not yet escaped protected agriculture to establish a population in the field. However, MEAM1 has been established in the United States for nearly 30 yr (McKenzie et al. 2009) and is a pest in both the field and in greenhouses where it can co-occur with MED (McKenzie et al. 2012).

MEAM1 and MED are sister taxa exhibiting 3% sequence divergence in the mitochondrial cytochrome oxidase subunit 1 (CO1) gene (Boykin et al. 2007, Dinsdale et al. 2010, De Barro et al. 2011) and 0.8% sequence divergence in coding regions of the transcriptome (Wang et al. 2011). Most whiteflies, including MEAM1 and MED, are haplodiploid with males produced parthenogenetically and females produced sexually. Reproductive isolation between MEAM1 and MED is nearly complete and consists of both prezygotic and postzygotic boundaries (Elbaz et al. 2010, Sun et al. 2011). Elbaz et al. (2010) monitored

Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. The USDA is an equal opportunity provider and employer.

[1] USDAûARS, U.S. Horticultural Research Laboratory, 2001 South Rock Rd., Fort Pierce, FL 34945.

[2] Mid-Florida Research & Education Center, University of Florida, 2725 Binion Rd., Apopka, FL 32703.

[3] Corresponding author, e-mail: cindy.mckenzie@ars.usda.gov.

0022-0493/13/1355–1364$04.00/0 © 2013 Entomological Society of America

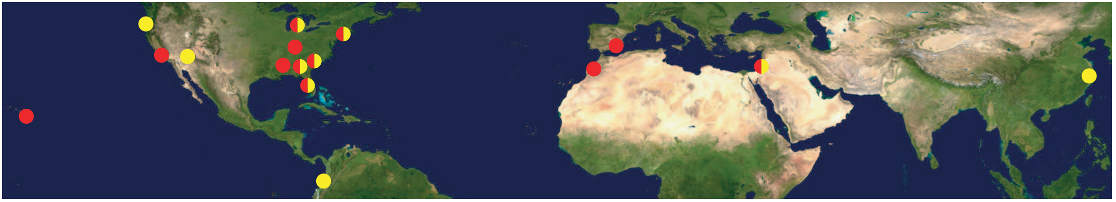


Fig. 1. Countries and U.S. states from which B. tabaci cryptic species MEAM1 (yellow) and MED (red) were sampled (also see Table 1).

interspecific pairs of whiteflies and found that all behaviors leading up to copulation occurred but witnessed no copulation events and subsequently found no evidence of sperm transfer indicating prezygotic reproductive isolation. Sun et al. (2011) was able to produce a small number of female interspecific hybrids (0–2%) when caging 10 males with 10 females, but F1s produced no progeny, indicating postzygotic reproductive isolation via hybrid sterility (Sun et al. 2011). In the field, evidence of hybridization is very rare and hybrids do not persist in the environment (McKenzie et al. 2012). In addition, nearly all putative cases of field hybridization based on nuclear microsatellites were the result of homoplasy and rejected upon allele sequencing (McKenzie et al. 2012; C.L.M., unpublished data).

Important biological differences have been noted between the two cryptic species. Although both vector a large number plant viruses, MED was found to be superior at transmitting tomato yellow leaf-curl virus, leading to rapid spread of the virus in China (Pan et al. 2012). MED also has been found to have elevated resistance to neonicatinoid insecticides in the field (Dennehy et al. 2010, Wang et al. 2010) and under selection pressure in the laboratory (Horowitz et al. 2005). In contrast, MEAM1 is more fecund, has a lower rate of nymphal mortality, and a higher rate of population increase than MED (Pascual and Callejas 2004, Crowder et al. 2010a). When MEAM1 and MED are caged together, MEAM1 females maintain both the frequency of copulation and the sex ratio of their progeny, whereas MED females mate less and produce fewer females due to reproductive interference from MEAM1 (Pascual and Callejas 2004; Crowder et al. 2010a,b). It is predicted that MEAM1 will generally outcompete and displace MED (Crowder et al. 2010a) based on this behavioral flexibility; however, MED has been rapidly displacing MEAM1 in field populations in China, likely due to high levels of neonicitinoid resistance (Wang et al. 2010). This mechanism also may explain the co-occurrence of MED and MEAM1 in greenhouse populations in the United States.

During the course of a nationwide survey of invasive MEAM1 and MED, several thousand whiteflies were identified via mitochondrial CO1 sequencing (McKenzie et al. 2012). A subset of these individuals plus samples from other countries and from laboratory colonies (Fig. 1; Table 1) were genotyped at six (MEAM1) and seven (MED) microsatellite loci to 1) test concordance among nuclear and mitochondrial genetic markers, 2) determine the population structure within each invasive cryptic species in the United States, and 3) infer possible invasion pathways for the more recently introduced MED cryptic species.

Materials and Methods

Whiteflies were solicited from vegetable and ornamental growers, crop consultants, industry representatives, international collaborators, and state and local agencies (Fig. 1; Table 1). DNA from female whiteflies was extracted and B. tabaci individuals were identified to cryptic species based on the sequence of a 700–800-bp mitochondrial CO1 gene fragment; methods, and primers in McKenzie et al. (2009, 2012) and Shatters et al. (2009). The samples identified in this way from both cryptic species were further genotyped at six microsatellite markers designated Bem6, Bem11, Bem15, Bem23, Bem25, and Bem31 (De Barro et al. 2003). MED samples were genotyped at a seventh microsatellite locus, Bem21 (De Barro et al. 2003). These microsatellites were developed before the designation of B. tabaci as a cryptic species complex, and some were designed using individuals from other cryptic species; however, most have been used previously for MEAM1 and MED (De Barro et al. 2003; Simon et al. 2007; Dalmon et al. 2008; McKenzie et al. 2009, 2012). This is the first study using Bem15 and Bem21 with MED. Allele fragments were fluorescently labeled, polymerase chain reaction (PCR) amplified, analyzed using a 3730xl sequencer (Applied Biosystems, Carlsbad, CA), and sized using GeneMapper 4.0 (Applied Biosystems). A detailed extraction protocol, PCR specifications, and Bem6 and Bem23 allele frequency data for U.S. field collections have been provided previously (McKenzie et al. 2009, 2012).

Duplicate genotypes were generated for 55 MEAM1 individuals and 203 MED individuals with a mismatch error rate of 1.2% (MEAM1) and 0.1% (MED) across all loci. Alleles were treated as missing data in cases of amplification failure, three or more peaks in the electropherogram, peaks of insufficient height, or a conflict between duplicates (Table 2). Accessions retained for population genetics analyses each had a minimum of 10 individuals (Table 1) and individuals retained amplified at (50% of loci. An accession here refers to a collection of whiteflies from a given site at a set time point as some sites were sampled on multiple occasions.

Table 1. State/country and host data for 103 accessions of two *B. tabaci* cryptic species

State or country	Host (if known)	*n*	State/ country	Host (if known)	*n*	State/ country	Host (if known)	*n*
MEAM1 accessions (76)								
AZ (f)	Cauliflower	12	FL (f)	Tomato	12	GA (f)	Collards	12
AZ (f)	Cauliflower	12	FL (f)	Tomato	12	GA (f)	Collards	12
AZ (g)	Poinsettia	12	FL (f)	Tomato	12	GA (f)	Collards	11
FL (f)	Basil	12	FL (f)	Tomato	12	GA (f)	Cucumber	11
FL	Bush bean	12 (l)	FL (f)	Tomato	12	GA (f)	Cucumber	12
FL (g)	Collards	14	FL (f)	Tomato	12	GA (f)	Eggplant	12
FL (g)	Daisy	12	FL (f)	Tomato	12	GA (f)	Eggplant	12
FL (g)	*Lantana*	12	FL (f)	Tomato	12	GA (f)	Eggplant	12
FL	Melon/*Tropaeolum*	12 (l)	FL (f)	Tomato/cantaloupe/ watermelon	12	GA (f)	Eggplant	12
FL (g)	Muskmelon	10	FL (f)	Tomato	12	GA (g)	Poinsettia	12
FL (f, g)	Pepper	12	FL (f)	Tomato	12	GA (f)	Snap beans	10
FL (f)	Pepper	12	FL (f)	Tomato	12	GA (f)	Squash	12
FL (g)	Poinsettia	13	FL (g)	Tomato	12	GA (f)	Squash	12
FL (g)	Poinsettia	13	FL (f)	Tomato	12	GA (f)	Tomato	11
FL (g)	Poinsettia	10	FL (f)	Tomato	12	GA (f)	Tomato	12
FL (g)	Poinsettia	12	FL (f)	Tomato	12	MI	Poinsettia	24 (l)
FL (g)	*Ruellia*	12	FL (f)	Tomato	12	NY (g)	Poinsettia	12
FL (g)	Sweet potato	12	FL (f)	Tomato/bean	11	OR (g)	*Begonia*	29
FL (f)	Sweet potato	12	FL (f)	Tomato	12	SC (f)	Collards/cantaloupe	12
FL (f)	Sweet potato	13	FL (f)	Tomato	12	China		15
FL (f)	Tomato	12	FL (f)	Tomato	12	Colombia	Beans	24 (l)
FL (f)	Tomato	12	FL	Tomato	12	Colombia	Cassava	21 (l)
FL (f)	Tomato	10	FL (f)	Tomato	12	Israel	Cotton	32
FL (f)	Tomato	11	FL (g)		10	Israel	Sweet Potato	21 (l)
FL (f)	Tomato	11	FL (g)		12	Israel		16
			FL (g)		12			987
MED accessions (27)								
AL (g)	Ornamentals	15	GA (g)	Ornamentals	12	MI	Poinsettia	12 (l)
CA (g)	Daisy	12	GA (g)	Poinsettia	14	NY (g)	Poinsettia	11
CA (g)	Daisy	24	GA (g)	*Salvia*	12	NY (g)	Poinsettia	10
FL (g)	Mint	11	GA (g)	*Salvia*	10	SC (g)	*Hibiscus*	12
FL (g)	*Hibiscus*	11	GA (g)	Poinsettia	10	Israel	Cotton	13
FL (g)	*Hibiscus*	11	HI (g)	Daisy	19	Morocco		12
FL (g)	*Hibiscus*	12	HI (g)	*Hibiscus*	11	Morocco		12
FL (g)	*Hibiscus*	11	KY (g)	Ornamentals	18	Spain	Tobacco	12
FL (g)	*Hibiscus*	10	MI	Poinsettia	13(l)	Spain		10 (l)
								340

Within each state/country, there are often multiple collection sites. (l), laboratory colonies; (g), greenhouse collected (United States only); and (f), field collected (United States only).

Population Genetics Statistics. Within each cryptic species, the number of alleles per locus and allelic richness standardized to the minimum sample size were calculated for each microsatellite in FSTAT version 2.93 (Goudet 1995). FSTAT was used to calculate gene diversity (HE averaged over loci), inbreeding (FIS averaged over loci), and the number of polymorphic loci for each accession. Each accession was tested for the presence of null alleles using Micro-checker version 2.2.3 (van Oosterhout et al. 2004) and departure from Hardy–Weinberg equilibrium (HWE) was assessed with exact tests in GenePop version 4 (Rousset 2008). Each accession showing evidence of heterozygote deficiency also was tested for inbreeding in INEst (Chybicki and Burczyk 2009), while accounting for null alleles. Each accession also was tested for linkage disequilibrium (nonindependence of loci) for each locus pair by using log-likelihood ratio tests in GenePop.

Table 2. Locus statistics and percentage of missing data for seven microsatellites in two cryptic species of *B. tabaci*

Locus	MEAM1[a]			MED[a]		
	n_A	*A*	% missing	n_A	*A*	% missing
Bem6	2	1.55	<1	4	1.42	2
Bem11	6	2.25	5	4	1.82	15
Bem15	11	4.92	1	9	1.52	<1
Bem21				32	4.64	2
Bem23	5	1.05	<1	8	2.11	1
Bem25	12	4.32	1	14	3.47	6
Bem31	4	1.23	0	4	2.19	<1

[a] n_A, *number of alleles*, *A*, allelic richness.

Determining Population Structure and Number of Populations. The number of populations (k) represented by each data set, MEAM1 and MED, was determined using the allele frequency contingency table permutation method described by Waples and Gaggiotti (2006) implemented in GenePop with $\alpha = 0.01$. In this method, if α exceeds 0.01 for any population pair, the two populations are consolidated. This method was shown to be more reliable in recovering k than individual-based methods (e.g., Pritchard et al. 2000, Corander and Marttinen 2006) with simulated data, especially when examining a small number of loci

or populations with high levels of migration (Waples and Gaggiotti 2006). To compare this strategy for determining k with a more commonly used individual-based approach, the data also were analyzed in STRUCTURE 2.3 (Pritchard et al. 2000) by using the alleles correlated model (Falush et al. 2003) with sampling location included (Hubisz et al. 2009). The model was run for 100,000 generations with a burn-in of 10,000 generations 20 times for each k from k = 1 to k = 10. The Evanno method (Evanno et al. 2005) was then implemented in STRUCTURE HARVESTER 0.6.93 (Dent and vonHoldt 2012), and the data were summarized for the optimal value of k by using CLUMPP 1.1 (Jakobsson and Rosenberg 2007) and Distruct 1.1 (Rosenberg 2004).

Principal coordinates analyses of each data set were conducted using GenAlEx 6.41 (Peakall and Smouse 2006) with missing data interpolated. Population structure was then visualized as polygons enclosing the populations inferred using the contingency table permutation method relative to the two major axes of molecular variation. To quantify the degree of allele fixation among inferred populations, pairwise FST values were estimated within each cryptic species, accounting for null alleles where necessary using FreeNA (Chapuis and Estoup 2007). After adjusting allele frequencies to account for the presence of null alleles, FreeNA estimates FST after Weir (1996). Confidence intervals around pairwise FST estimates were calculated using 10,000 bootstrap pseudoreplicates over loci. FST estimates are valuable in that they allow comparison with other studies due to their long history of use. However, because FST values tend to be biased downward when using highly variable microsatellites (Meirmans and Hedrick 2011), pairwise GST″ also were used calculated after Meirmans and Hedrick (2011) with the unbiased estimators of HT and HS from FSTAT. GST″ is an unbiased estimate of F′ST standardized by its maximum possible value (Meirmans and Hedrick 2011).

Comparing Microsatellite Genotype and Mitochondrial CO1 Haplotypes for MED. Previous reports documented an association of MED mitochondrial CO1 haplotype with allele frequencies at putative diagnostic microsatellite loci Bem6 and Bem23 (McKenzie et al. 2009, 2012). To assess whether this pattern was robust to a larger nuclear data set, genotype eigenvalues for each whitefly also were visualized in a principal coordinates graph according to their mitochondrial haplotype.

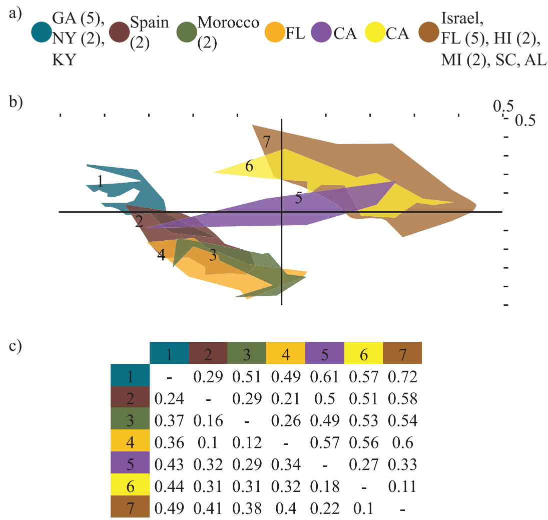
Results

Population Genetics Statistics

Allelic richness ranged from 1.05 to 4.92 (Table 2). Gene diversity within each accession was fairly low and ranged from 0.14 to 0.4 for MEAM1 (mean HE = 0.31) and from 0.09 to 0.52 for MED (mean HE = 0.30). Within MED, accessions outside the United States had higher gene diversity (mean HE of five accessions = 0.44) than accessions within the United States (mean HE of 22 accessions = 0.27) (permutation test; P = 0.002). The number of polymorphic loci per accession ranged from 2 to 6 for MEAM1 (mean = 3.53) and from 2 to 7 for MED (mean = 4.78) (Supp Table 1 [online only]).

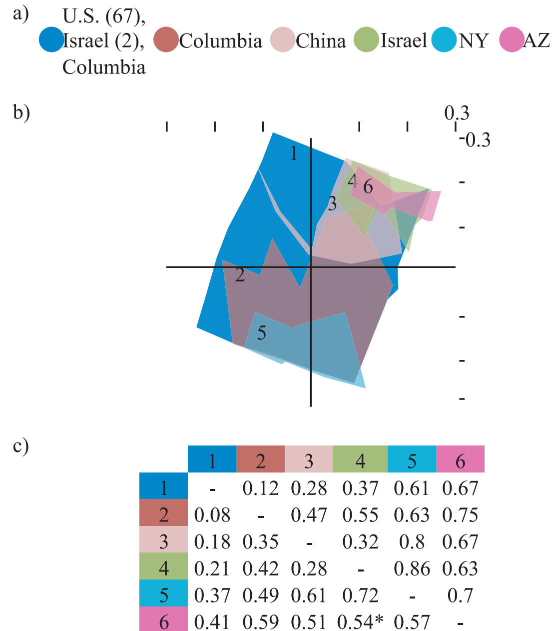
Micro-checker indicated the presence of null alleles at a single locus in 8% of MEAM1 accessions and 22% of MED accessions. There was evidence for null alleles at two loci in 7% of MED accessions. Bem6 (both cryptic species) and Bem21 (MED) accounted for 10 of 16 cases of null alleles. Five of six MEAM1 accessions and eight of eight MED accessions showing evidence of null alleles also deviated significantly from HWE ($P < 0.05$) (Supp Table 1 [online only]). Accessions with evidence of null alleles tended to have positive inbreeding coefficients (mean FIS of six MEAM1 accessions = 0.06, mean FIS of eight MED accessions = 0.26), whereas accessions with no evidence of null alleles tended to have slightly negative inbreeding coefficients (mean FIS of 70 MEAM1 accessions = −0.1, mean FIS of 14 MED accessions = −0.08) (Supp Table 1 [online only]). None of the accessions had a significantly positive inbreeding coefficient after nulls were accounted for. A significant excess of heterozygotes was found in 17 (Bem11), four (Bem15), and one (Bem25) MEAM1 accession and two (Bem11), one (Bem21), and one (Bem25) MED accessions ($P < 0.05$); 81% of MEAM1 accessions and 74% of MED accessions were in HW equilibrium. Violations of HWE in MED were always the result of heterozygote deficit, whereas violations in MEAM1 were just as likely to result from heterozygote excess as from heterozygote deficit. Two accessions from each cryptic species showed linkage disequilibrium between a single locus pair. This is at least an order of magnitude fewer significant comparisons than would be expected by chance (5%). Population genetic summary statistics for each accession are provided in Supp Table 1 [online only].

Estimating Number of Populations. Seven populations of B. tabaci were inferred for MED (Fig. 2) by using the contingency table permutation method of Waples and Gaggiotti (2006) with $\alpha = 0.01$. One population contained accessions from Georgia (five accessions), New York (two accessions), and Kentucky (population 1 in Fig. 2). A second population contained accessions from Florida hibiscus (five accessions); Hawaii (two accessions); Michigan (two accessions); and South Carolina, Alabama, and Israel (population 7 in Fig. 2). The remaining U.S. accessions, two from California and one from Florida (host: mint) were each resolved as distinct populations (populations 4–6 in Fig. 2) as were accessions from Spain (two) and Morocco (two) (populations 2 and 3, respectively, in Fig. 2). Principal coordinates axes 1 and 2, respectively, accounted for 56 and 13% of the genetic variation found in MED. Visualizing genotype eigenvalues on these two axes showed two major clusters of populations (Fig. 2b). One cluster contained populations 1–4 (Spain/Morocco-like) and one cluster contained populations 6 and 7 (Israel-like) (Fig. 2b). One of the populations from California contained



	1	2	3	4	5	6	7
1	-	0.29	0.51	0.49	0.61	0.57	0.72
2	0.24	-	0.29	0.21	0.5	0.51	0.58
3	0.37	0.16	-	0.26	0.49	0.53	0.54
4	0.36	0.1	0.12	-	0.57	0.56	0.6
5	0.43	0.32	0.29	0.34	-	0.27	0.33
6	0.44	0.31	0.31	0.32	0.18	-	0.11
7	0.49	0.41	0.38	0.4	0.22	0.1	-

Fig. 2. (a) Seven populations of B. tabaci MED inferred using pairwise contingency table permutation. (b) Populations visualized as shapes enclosing the eigenvalues of individual genotypes plotted on the first and second principal coordinates axes of molecular variation. (c) Pairwise fixation indices FST (below diagonal) and GST″ (above diagonal) among inferred populations. Color coding of populations is consistent throughout. Parentheses denote the number of accessions from a given state/country in the inferred population.



	1	2	3	4	5	6
1	-	0.12	0.28	0.37	0.61	0.67
2	0.08	-	0.47	0.55	0.63	0.75
3	0.18	0.35	-	0.32	0.8	0.67
4	0.21	0.42	0.28	-	0.86	0.63
5	0.37	0.49	0.61	0.72	-	0.7
6	0.41	0.59	0.51	0.54*	0.57	-

Fig. 3. (a) Populations of B. tabaci MEAM1 inferred using pairwise contingency table permutation. (b) Populations visualized as shapes enclosing the eigenvalues of individual genotypes plotted on the first and second principal coordinates axes of molecular variation. (c) Pairwise fixation indices FST (below diagonal) and GST″ (above diagonal) among inferred populations. Color coding of populations is consistent throughout. Parentheses denote the number of accessions from a given state/country in the inferred population. The asterisk denotes an FST value whose confidence interval overlaps zero when bootstrapping over loci.

genotypes in both clusters and genotypes intermediate between both clusters (population 5 in Fig. 2b). Analysis with STRUCTURE recovered two populations consistent with the two major clusters of populations seen using principal coordinates analysis with California (population 5 in Fig. 2b) intermediate (Supp Fig. 1 [online only]). Of the seven populations recovered with contingency table permutation, only population 1 had no evidence of null alleles, so all pairwise FST values were calculated using the ENA correction in FreeNA (Chapuis and Estoup 2007). Pairwise FST estimates ranged from 0.1 to 0.49 (Fig. 2c). None of the FST confidence intervals overlapped zero. Pairwise GST″ were on average 56% higher than their corresponding FST (Fig. 2c).

Six populations of B. tabaci were inferred for MEAM1 by using contingency table permutation (Fig. 3). Within MEAM1, 71 of 76 accessions comprised a single large population for which the hypothesis of panmixia could not be refuted. This population contained 67 U.S. accessions, two accessions from Israel, and one from Colombia (population 1 in Fig. 3). Individual accessions from Colombia, China, Israel, New York, and Arizona were each resolved as separate populations (populations 2–6 in Fig. 3). Analysis with STRUCTURE recovered two populations, one population made up largely of individuals from the New York and Arizona accessions recovered as separate populations by contingency table permutation method, and the second population containing all other accessions (Supp Fig. 1). Principal coordinates analysis for MEAM 1 was less informative than for MED with axes 1 and 2 accounting for 26 and 24% of the genetic variation, respectively (Fig. 3b). Pairwise FST estimates ranged from 0.08 to 0.72. Fourteen of 15 estimates had 95% confidence limits (CL) not overlapping zero (Fig. 3c). Pairwise GST″ were on average 38% higher than their corresponding FST (Fig. 3c).

Concordance of Nuclear and Mitochondrial Data for MED. Grouping the nuclear genotype eigenvalues according to mitochondrial CO1 haplotype indicated concordance between marker systems (Table 3 and Fig. 4). CO1 haplotypes are labeled Q1-Q4 consistent with previous reports (McKenzie et al. 2009, 2012) and are associated with the eastern Mediterranean (haplotype Q1; Israel-like in Fig. 2b) and western Mediterranean (haplotypes Q2-Q4, Spain and Morocco-like in Fig. 2b). Haplotype Q4 was only found in Morocco. All but one individual with haplotype Q1 clustered together and all but two individuals with haplotypes Q2-Q4 clustered together. Furthermore, haplotypes Q2, Q3, and Q4 are not evenly dispersed but form distinct clusters of their own within the western Mediterranean cluster (Fig. 4). Three individuals were of mixed ancestry, showing discordance between markers. One individual with the Q1 CO1 haplotype had a Q2-Q4 associated microsatellite ge-

Table 3. Single-nucleotide polymorphisms delimiting four *B. tabaci* MED mitochondrial CO1 haplotypes

Position[a]	49	88	211	283	457	481	502	541	613	661	710
Consensus sequence	C	C	C	C	C	C	C	T	T	C	T
Q1	*	*	T	*	*	T	T	C	*	T	C
Q2	*	*	*	*	*	*	*	*	*	*	*
Q3	*	G	*	*	*	*	*	*	*	*	*
Q4	T	T	*	T	T	*	*	*	C	T	*

[a] Position is relative to GenBank accession HQ198789; position 0 is the 3′ terminal residue of the C1-J-2195 universal CO1 primer (Simon et al. 1994).

notype and two individuals with the Q3 CO1 haplotype had a Q1-associated microsatellite genotype (Fig. 4). All three mixed individuals were from California (population 5 in Fig. 2), the only U.S. accession containing Q1, Q2, and Q3 CO1 haplotypes.

Discussion

Both STRUCTURE analysis and principal coordinates analysis of B. tabaci MED microsatellites recovered two major clusters of nuclear genotypes (Fig. 2), indicating the presence of two major MED genetic groups with origins in the western Mediterranean (populations 4–6 in Fig. 2) and eastern Mediterranean (populations 1 and 2 in Fig. 2) (also see Fig. 1 and Supp Fig. 1 [online only]). Grouping the nuclear genotype eigenvalues according to mitochondrial CO1 haplotype suggested that the two marker systems were largely concordant (Table 3; Fig. 4). The western MED and eastern MED groups are supported here by both microsatellites and mitochondrial CO1 sequences and have been supported previously based solely on mitochondrial CO1 sequences (Tsagkarakou et al. 2007, De Barro and Ahmed 2011). This east–west MED clustering contrasts somewhat with the microsatellite-based results of Simon et al. (2007) who found greater differentiation between populations from Spain and Morocco. Their FST estimate between Spain and Morocco (0.30) was almost double ours and outside the 95% confidence interval (CI) of our estimate (0.16; 95% CI, 0.12–0.18). The broad confidence limits of some of our estimates combined with the low number of markers used in both studies highlight the need for a larger number of molecular markers in resolving the phylogeographic breaks within MED in greater detail. In California, where both eastern and western MED coexist, we uncovered evidence of introgression between eastern and western MED causing discordance between nuclear and mitochondrial markers for three individuals. These results suggest that within the first 3 yr after the discovery of MED in the United States, only limited gene flow had taken place between eastern and western MED whiteflies. The insular greenhouse environment that presently constrains MED in the United States likely contributes to continued isolation and population structuring between eastern and western MED whitefly populations as only 8% of MED collections since 2005 have had both eastern and western MED mitochondrial haplotypes occurring together (McKenzie et al. 2012).

It is possible that MED in the United States was founded by a single hyperdiverse population, possibly in California, with subsequent bottlenecks accompanying expansion within the United States. This, however, seems unlikely because the population from California containing both eastern and western MED whiteflies and the population from Florida mint are isolated along principal coordinate two (populations 5 and 4, respectively, in Fig. 2b). It is more likely that the diverse California population represents a secondary sink population and not the source of initial colonization. Because MED in the United States is only found in greenhouses, migration and colonization are probably mediated only by movement of infested

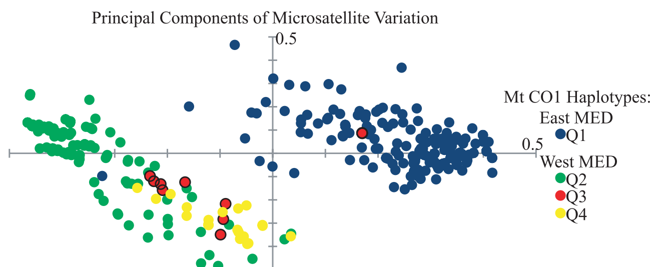


Fig. 4. B. tabaci MED microsatellite genotypes plotted on the first and second principal coordinates axes of molecular variation color coded by mitochondrial CO1 haplotype. Generally the two marker systems are concordant. One individual has and eastern Mediterranean mitochondrial haplotype (Q1) in a western Mediterranean nuclear background, and two individuals have a western Mediterranean mitochondrial haplotype (Q3) in an eastern Mediterranean nuclear background. All three discordant individuals were from California (population 5 in Fig. 2). Haplotype Q4 is only found in Morocco. Two singleton western Mediterranean CO1 haplotypes are not shown. Single-nucleotide polymorphisms delimiting mitochondrial haplotypes are presented in Table 3.

plant material shipments between greenhouses rather than through active flight. The California population may have persisted long enough to receive migrants in this way and to integrate genetic diversity from multiple sources. However, this introgression among eastern and western MED groups seems anomalous, and colonization and extinction dynamics without such genetic integration is more likely the norm among U.S. commercial greenhouses.

The Florida, Hawaii, Alabama, South Carolina, and Michigan accessions originated in the eastern Mediterranean and Israel is a plausible source of this introduction because all 11 of these U.S. accessions were consolidated to a single population with Israel by the contingency table permutation analysis (population 7 in Fig. 2). Nonetheless, an Israeli origin of eastern Mediterranean MED should be corroborated by additional genetic markers. Populations 5 and 6 in Fig. 2 also contain majority ancestry from the eastern Mediterranean. Although western MED has been the dominant invasive MED group globally, the United States, Mexico, and Canada are the only countries outside its native range that eastern MED has been detected (De Barro and Ahmed 2011, McKenzie et al. 2012). MED populations 1 (Georgia/New York/Kentucky) and 4 (Florida mint) originated somewhere in the western Mediterranean and likely represent independent introductions. Previous reports have supported multiple introductions of western MED into the United States (McKenzie et al. 2009, 2012) based on mitochondrial data and two microsatellite loci, and this interpretation is supported by this larger nuclear data set. It is unclear based on these results whether the introductions follow neatly with the mitochondrial haplotypes designated Q2 and Q3; however, it should be noted that the haplotypes Q2, Q3, and Q4 are not distributed evenly but themselves form subclusters within western MED (Fig. 4). This clustering suggests additional concordance between nuclear and mitochondrial markers within western MED that could be explored in more detail with greater sampling. Such a correlation has previously been proposed using two microsatellite loci (McKenzie et al. 2012). Given such a high degree of MED genetic diversity present in the United States, impetus should be given to describing insecticide resistance profiles in a population genetic context that takes into account the phylogeographic origin of the particular MED population in question as insecticide resistance traits may be differentially expressed in eastern and western MED.

Population 6 in Fig. 2 is an accession of whiteflies from California, which showed evidence of hybridization with MEAM1 by using esterase enzyme analysis (F. J. Byrne, personal communication). The enzyme analysis is destructive and so individuals subject to enzyme analysis could not be subsequently genotyped. Even so, none of the 24 whiteflies genotyped from this putatively hybridizing population showed any evidence of being interspecific hybrids between MEAM1 and MED based on lack of allele sharing at Bem6 and Bem23 (McKenzie et al. 2012). Another population from Michigan also had a large proportion of hybrids by using the esterase method (F. J. Byrne, personal communication). This population was the source material of a quarantined lab colony (Table 1 both MEAM1 and MED) and also showed no evidence of hybridization (McKenzie et al. 2012). All 49 whiteflies from this site possessed both nuclear and mitochondrial molecular markers unambiguously identifying them as either MED or MEAM1.

MEAM1, in general, shows a much higher degree of genetic homogeneity across the United States (less population substructure) than does MED. Even though 63 of the 71 accessions in the large MEAM1 population came from just two southern U.S. states, Georgia and Florida, this large population also contained populations from Arizona and Oregon in the western United States and Michigan in the northern United States (population 1 in Fig. 2). Although two field-collected accessions of Arizona MEAM1 were resolved within the single large U.S. population, the greenhouse-collected accession from Arizona poinsettia was resolved as a separate population (population 6 in Fig. 2). MEAM1 in Arizona field populations was the subject of a fitness-driven sweep of Rickettsia infection from 2000 to 2006 (Himler et al. 2011), and it would be interesting to know if this sweep also occurred in Arizona greenhouses if mixing among field and greenhouse populations is limited. The accession from New York (population 5 in Fig. 2) was also from a poinsettia greenhouse and both the New York and Arizona greenhouse populations had a high frequency of the much rarer 223 allele at the Bem6 locus. Given that both of these MEAM1 accessions resolved as distinct populations were from protected agricultural environments and given the high degree of population structure recovered within MED, which only occurs in greenhouses in the United States, it seems plausible that the isolation afforded by greenhouses may be one of the factors maintaining genetic differentiation in U.S. Bemisia populations. This is a conjecture that is further supported by the population structure found for greenhouse MED in France (Dalmon et al. 2008).

Because MED is restricted to greenhouses in the United States, a reduction in genetic diversity is predicted for all populations, not just for those from laboratory colonies because all populations have restricted gene flow opportunities. In fact, reduced genetic diversity was found for U.S. MED relative to MED collected in the native range (Supp Table 1 [online only]), and this could be partly due to ongoing genetic isolation imposed by the restriction of MED to greenhouses in addition to the bottleneck effect expected for a newly invasive species (Hufbauer and Roderick 2005). Consistent with the prediction of lower genetic diversity in protected agriculture, the accessions of MEAM1 that were resolved as differentiated populations did have some of the lowest levels of gene diversity (New York, HE = 0.2; Arizona, HE = 0.18) (Supp Table 1 [online only]). However, U.S. greenhouse-collected MEAM1 tended to show a similar degree of gene diversity (mean HE of 21 accessions = 0.30) to MEAM1 collected in the field (mean HE of 48 accessions = 0.33), inconsistent with this

prediction. This inconsistency could result from the low number of genetic markers used but could also reflect ongoing significant gene flow between field and greenhouse MEAM1 populations, at least within Florida and Georgia where temperature, host availability, and humidity differences among environments are less likely to constrain whitefly movement from the greenhouse into the field.

The strategy for identifying the number of populations in a genetic data set is greatly limited by the number of loci used, and this is true for both individual based methods (e.g., Pritchard et al. 2000, Corander and Marttinen 2006) as well as the population based method used in this study (Waples and Gaggiotti 2006). However, in this study the Waples and Gaggiotti method resolved four more differentiated populations of MEAM1 and five more differentiated populations of MED than did the individual based assignment program STRUCTURE (Pritchard et al. 2000, Evanno et al. 2005) (Supp Fig. 1 [online only]). This is consistent with the simulation results of Waples and Gaggiotti (2006) indicating a greater power of the contingency table permutation method to detect population structure when low numbers of loci are used. This is also consistent with the results of Kalinowski (2011) who provide evidence that the Hardy–Weinberg population detection algorithm of STRUCTURE tends to be conservative. STRUCTURE, despite recovering a smaller number of populations than the contingency table permutation method, produced results that were largely consistent with other analyses with the exception of the uninformative principal coordinate analysis of MEAM1 (Fig. 3b). STRUCTURE consolidated greenhouse collected accessions from Arizona and New York into the same population (Supp Fig. 1 [online only]) and in contrast, principal coordinate analysis isolated them completely along axis 2 (Fig. 3b). In the absence of more data and a phylogenetic analysis, neither result should be used to infer the evolutionary history and relatedness of these greenhouse populations at the present time (see Kalinowski 2011).

With the small number of loci used in this study, only the population pairs exhibiting very high levels of differentiation among them could be resolved (Figs. 1c and 2c). Interpretation of such results should be done cautiously as increasing the number of loci (especially those that are highly polymorphic) would resolve additional population structure in the United States for both MEAM1 and MED. Rather than interpreting groups of accessions as panmictic based on the Waples and Gaggiotti method, these groups result from the failure to reject panmixia given sampling constraints. For example, it is unlikely that U.S. MEAM1 (established in the United States for ~30 yr) is panmictic with populations in Colombia and Israel. Instead, this failure to reject panmixia is probably a byproduct of low genetic sampling, and we predict that these populations would be resolved as differentiated with the addition of more loci.

In another example, panmixia could not be rejected for 11 U.S. MED accessions from Florida to Hawaii (~7,500 km) (population 7 in Fig. 2). Although some of this failure is certainly the result of low molecular marker number, adding data from five loci to accessions from Michigan (two) and Hawaii (two), did not resolve further population structure and panmixia was still not rejected among these four accessions (data not shown). Continued failure to reject panmixia with 12 microsatellite loci for whitefly collections separated by ~7,000 km leads to increased confidence that these four collections were in very recent contact and highlights the gene flow opportunities that anthropogenic trade of greenhouse ornamental plants provides for invasive MED whiteflies within the United States.

Acknowledgments

We are grateful to Holly Nance, Michel Slotman, and one anonymous reviewer for helpful suggestions to earlier versions of this manuscript. This research was partially funded by the Floriculture and Nursery Research Initiative.

References Cited

Boykin, L. M., R. G. Shatters, Jr., R. C. Rosell, C. L. McKenzie, R. Bagnall, P. De Barro, and D. R. Frohlich. 2007. Global relationships of Bemisia tabaci (Hemiptera: Aleyrodidae) revealed using bayesian analysis of mitochondrial COI DNA sequences. Mol. Phylogenet. Evol. 44: 1306–1319.

Chapuis, M. P., and A. Estoup. 2007. Microsatellite null alleles and estimation of population differentiation. Mol. Biol. Evol. 24: 621–631.

Chu, D., Y. J. Zhang, J. K. Brown, B. Cong, B. Xu, Q. Wu, and G. Zhu. 2006. The introduction of the exotic Q biotype of Bemisia tabaci from the Mediterranean region into China on ornamental crops. Fla. Entomol. 89: 168–174.

Chybicki, I., and J. Burczyk. 2009. Simultaneous estimation of null alleles and inbreeding coefficients. Heredity 100: 106–113.

Corander, J., and P. Marttinen. 2006. Bayesian identification of admixture events using multi-locus molecular markers. Mol. Ecol. 15: 2833–2843.

Crowder, D. W., A. R. Horowitz, P. J. De Barro, S. S. Liu, A. M. Showalter, S. Kontsedalov, V. Khasdan, A. Shargal, J. Liu, Y. Carrière, et al. 2010a. Mating behaviour, life history and adaptation to insecticides determine species exclusion between whiteflies. J. of Animal Ecology. 79: 563–570.

Crowder, D. W., M. I. Sitvarin, and Y. Carrière. 2010b. Plasticity in mating behaviour drives asymmetric reproductive interference in whiteflies. Anim. Behav. 79: 579–587.

Dalmon, A., F. Halkett, M. Granier, H. Delatte, and M. Peterschmitt. 2008. Genetic structure of the invasive pest Bemisia tabaci: evidence of limited but persistent genetic differentiation in glasshouse populations. Heredity 100: 316–325.

De Barro, P., and M. Z. Ahmed. 2011. Genetic networking of the Bemisia tabaci cryptic species complex reveals pattern of biological invasions. PLoS ONE 6: e25579.

De Barro, P. J., K. D. Scott, G. C. Graham, C. L. Lange, and M. K. Schutze. 2003. Isolation and characterization of microsatellite loci in Bemisia tabaci. Mol. Ecol Notes 3: 40–43.

De Barro, P. J., S. S. Liu, L. M. Boykin, and A. B. Dinsdale. 2011. Bemisia tabaci: a statement of species status. Annu. Rev. Entomol. 56: 1–19.

Dennehy, T. J., B. A. Decain, V. S. Harpold, M. Zaborac, S. Morin, J. A. Fabrick, R. L. Nichols, J. K. Brown, F. J. Byrne, and L. I. Xianchun. 2010. Extraordinary resistance to insecticides reveals exotic Q biotype of Bemisia tabaci in the New World. J. Econ. Entomol. 103: 2174–2185.

Dent, E. A., and B. M. vonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE outpup and implementing the Evanno method. Conserv. Genet. Resour. 4: 359–361.

Dinsdale, A., L. Cook, C. Riginos, Y. M. Buckley, P. De Barro, and A. Dinsdale. 2010. Refined global analysis of Bemisia tabaci (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae) mitochondrial cytochrome oxidase 1 to identify species level genetic boundaries. Ann. Entomol. Soc. Am. 103: 196–208.

Elbaz, M., N. Lahav, S. Morin, and M. Elbaz. 2010. Evidence for pre-zygotic reproductive barrier between the B and Q biotypes of Bemisia tabaci (Hemiptera: Aleyrodidae). Bull. Entomol. Res. 100: 581–590.

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

Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567–1587.

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

Himler, A. G., T. Adachi-Hagimori, J. E. Bergen, A. Kozuch, S. E. Kelly, B. E. Tabashnik, E. Chiel, V. E. Duckworth, T. J. Dennehy, E. Zchori-Fein, et al. 2011. Rapid spread of a bacterial symbiont in an invasive whitefly is driven by fitness benefits and female bias. Science 332: 254–256.

Horowitz, A. R., I. Denholm, K. Gorman, J. L. Cenis, S. Kontsedalov, and I. Ishaaya. 2003. Biotype Q of Bemisia tabaci identified in Israel. Phytoparasitica 31: 94–98.

Horowitz, A. R., S. Kontsedalov, V. Khasdan, and I. Ishaaya. 2005. Biotypes B and Q of Bemisia tabaci and their relevance to neonicotinoid and pyriproxyfen resistance. Arch. Insect Biochem. Physiol. 58: 216–225.

Hubisz, M. J., D. Falush, M. Stephens, and J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resour. 9: 1322–1332.

Hufbauer, R. A., and G. K. Roderick. 2005. Microevolution in biological control: mechanisms, patterns, and processes. Biol. Control 35: 227–239.

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

Kalinowski, S. T. 2011. The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity 106: 625–632.

McKenzie, C. L., G. Hodges, L. S. Osborne, F. J. Byrne, and R. G. Shatters, Jr. 2009. Distribution of Bemisia tabaci (Hemiptera: Aleyrodidae) biotypes in Florida-investigating the Q invasion. J. Econ. Entomol. 102: 670–676.

McKenzie, C. L., J. Bethke, F. J. Byrne, J. Chamberlin, T. J. Dennehy, A. M. Dickey, D. Gilrein, P. Hall, S. Luswig, R. Oetting, et al. 2012. Distribution of Bemisia tabaci (Hemiptera: Aleyrodidae) biotypes in North America after the Q invasion. J. Econ. Entomol. 105: 753–766.

Meirmans, P. G., and P. W. Hedrick. 2011. Assessing population structure: FST and related measures. Mol. Ecol. Resour. 11: 5–18.

Oliveira, M., T. Henneberry, and P. Anderson. 2001. History, current status, and collaborative research projects for Bemisia tabaci. Crop Prot. 20: 709–723.

Pan, H., D. Chu, W. Yan, Q. Su, B. Liu, S. Wang, Q. Wu, W. Xie, X. Jiao, R. Li, et al. 2012. Rapid spread of tomato yellow leaf curl virus in China is aided differentially by two invasive whiteflies. PLoS One 7: e34817.

Pascual, S., and C. Callejas. 2004. Intra- and interspecific competition between biotypes B and Q of Bemisia tabaci (Hemiptera: Aleyrodidae) from Spain. Bull. Entomol. Res. 94: 369–375.

Peakall, R., and P. E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6: 288–295.

Perring, T. M. 2001. The Bemisia tabaci species complex. Crop Prot. 20: 725–737.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Ren, S., Z. Wang, B. Qui, and Y. Xiao. 2001. The pest status of Bemisia tabaci in China and non-chemical control strategies. Insect Sci. 8: 279–288.

Rosenberg, N. A. 2004. Distruct: a program for the graphical display of population structure. Mol. Ecol. Notes 4: 137–138.

Rousset, F. 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol. Ecol. Resour. 8: 103–106.

Shatters, R. G., Jr., C. A. Powell, L. M. Boykin, H. Liansheng, and C. L. McKenzie. 2009. Improved DNA barcoding method for Bemisia tabaci and related Aleyrodidae: development of universal and Bemisia tabaci biotype-specific mitochondrial cytochrome c oxidase I polymerase chain reaction primers. J. Econ. Entomol. 102: 750–758.

Simon, B., J. L. Cenis, and P. De La Rua. 2007. Distribution patters of the Q and B biotypes of Bemisia tabaci in the Mediterranean Basin based on microsatellite variation. Entomol. Exp. Appl. 124: 327–336.

Simon, C., F. Frati, B. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87: 651–701.

Sun, D. B., J. Xu, J. B. Luan, and S. S. Liu. 2011. Reproductive incompatibility between the B and Q biotypes of the whitefly Bemisia tabaci in China: genetic and behavioral evidence. Bull. Entomol. Res. 101: 211–220.

Tsagkarakou, A., C. S. Tsigenopoulos, K. Gorman, J. Lagnel, and I. D. Bedford. 2007. Biotype status and genetic polymorphism of the whitefly Bemisia tabaci (Hemiptera: Aleyrodidae) in Greece: mitochondrial DNA and microsatellites. Bull. Entomol. Res. 97: 29–40.

Ueda, S., and J. K. Brown. 2006. First report of the Q biotype of Bemisia tabaci in Japan by mitochondrial cytochrome oxidase I sequence analysis. Phytoparasitica 34: 405–411.

van Oosterhout, C., W. F. Hutchinson, D.P.M. Wills, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Notes 4: 535–538.

Viscarret, M. M., I. Torres-Jerez, E. Agostini de Manero, S. N. Lopez, E. E. Botto, and J. K. Brown. 2003. Mitochondrial DNA evidence for a distinct new world group of Bemicia tabaci (Gennadius) (Hemiptera: Aleyrodidae) indigenous

to Argentina and Bolivia, and presence of the old world B biotype in Argentina. Ann. Entomol. Soc. Am. 96: 65–72.

Wang, X. W., J. B. Luan, J. M. Li, Y. L. Su, J. Xia, and S. S. Liu. 2011. Transcriptome analysis and comparison reveal divergence between two invasive whitefly cryptic species. BMC Genomics 12: 458.

Wang, Z., H. Yan, Y. Yang, and Y. Wu. 2010. Biotype and insecticide resistance status of the whitefly Bemisia tabaci from China. Pest Manag. Sci. 66: 1360–1366.

Waples, R. S., and O. Gaggiotti. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol. Ecol. 15: 1419–1439.

Weir, B. S. 1996. Genetic data analysis II. Sinauer, Sunderland, MA.

Received 19 December 2012; accepted 7 February 2013.