Ecology and Evolution

Open Access

Evidence of gene flow between sympatric populations of the Middle East-Asia Minor 1 and Mediterranean putative species of *Bemisia tabaci*

Abdessalem Tahiri[1], Fabien Halkett[2,3], Martine Granier[4], Gwénaelle Gueguen[5] & Michel Peterschmitt[4]

[1]Ecole Nationale d'Agriculture, BP S 40, Meknès, Morocco
[2]INRA, UMR1136 IAM, F-54500 Vandoeuvre-lès-Nancy, France
[3]Université de Lorraine, UMR1136 IAM, F-54280 Champenoux, France
[4]CIRAD, UMR BGPI, F-34398 Montpellier, France
[5]Université Lyon 1; CNRS; UMR 5558, F-69622 Villeurbanne, France

Abstract

Bemisia tabaci is a complex of putative species that exhibit a strong geographical pattern. Crossing experiments have revealed various degrees of reproductive isolation between these nascent species, ranging between fertile first-generation hybrids (F1) and no F1 at all. However, the relevance of these results under natural conditions is generally not known. The worldwide invasion of the putative species Middle East-Asia Minor 1 (MEAM1) has caused secondary contacts between allopatric species, which in turn provide an opportunity to detect potential hybrids in nature. A total of 346 female *B. tabaci* were collected in 2003 and 2005 in the North East of Morocco and assigned to MEAM1 (119), Mediterranean (Med) (225) and a new putative species (2) using mitochondrial cytochrome oxidase (mtCOI) gene sequences. MEAM1 and Med individuals were characterized at seven microsatellite loci. MEAM1 and Med were found to be sympatric in 11 of 12 samples (6 fields/year). As previously reported from Spain, MEAM1 frequency decreased over time. The genetic data are consistent with a recent introduction of MEAM1. A Bayesian clustering analysis (STRUCTURE) distinguished two groups, which were 100% consistent with the mtCOI groups. From several lines of evidence, two individuals were identified as hybrids. Assignment profiles using NEWHYBRIDS and allele composition indicated that they were not F1 hybrids. The results are discussed in relation to the secondary endosymbiont infection status determined on a sample of individuals, and the contrasting outcomes of the reported crossing experiments between MEAM1 and Med.

Introduction

The whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) (Fig. 1) is a pest and virus vector in all warm-to-hot climate regions. A strong geographical structure of *B. tabaci* has been detected using various genetic markers, particularly microsatellites (De Barro 2005), ITS1 and mitochondrial cytochrome oxidase I (mtCOI) (De Barro

et al. 2005; Boykin et al. 2007). The high genetic differentiation between populations and biological differences with respect to host plant and reproductive incompatibility has triggered an exciting debate as to whether *B. tabaci* is a complex species consisting of various biotypes or races, or a species complex (Brown et al. 1995; Frohlich et al. 1999; De Barro et al. 2005). Most recently, based on genetic and biological data, Dinsdale et al. (2010) and De Barro et al.

Figure 1. Adult *Bemisia tabaci* and eggs on cabbage (by courtesy of A. Franck (CIRAD), 2000).

(2011) concluded that *B. tabaci* is a species complex. These authors used mtCOI sequences to provide a rationale to classify previously identified individuals of *B. tabaci* into 11 groups containing 24 species. The species classification is consistent with crossing experiments in which mostly no reproductive compatibility could be detected (reviewed in Liu et al. [2012]). Although this classification needs to be confirmed with other genetic markers to establish a more reliable phylogeny, the terminology derived from this classification is used here and was preferred to the usual biotype terminology because the distinctive biological features that are required to distinguish biotypes were mostly missing.

Species in nature are often incompletely isolated for millions of years after their formation (Mallet 2005). As divergence between some putative species of *B. tabaci* is estimated to be within 3–4 million years (Delatte et al. 2005; De Barro et al. 2011), hybridization may still be expected. Consistently, hybrids were obtained in seven of the tested combinations of the putative species (reviewed in Liu et al. [2012]): Middle East-Asia Minor 1 × Mediterranean (MEAM1 × Med), Asia II-1 × Asia II-3, Asia II-1 × Asia II-7, MEAM1 × Asia II-7, MEAM1 × China 1, Sub-Saharan Africa 1 × Med (S-Afr1 × Med), S-Afr1 × S-Afr2. Hybridization between putative *B. tabaci* species is detectable by generation of female progeny because whiteflies exhibit haplodiploidy consisting of male production from unfertilized eggs and female production from fertilized eggs. The F1 progenies obtained with these inter-species crosses were dominated by males (Byrne et al. 1995; De Barro and Hart 2000; Maruthi et al. 2004; Omondi et al. 2005; Xu et al. 2010), whereas the sex ratio was nearly 50% with intra-species crosses performed as a control. The fertility of the F1 females was not formally tested except for crosses S-Afr1 × S-Afr2, in which an F2 generation was obtained (Maruthi et al. 2004), and for crosses MEAM1 × Med, in which F1 females were sterile (Sun et al. 2011). Interestingly, whereas crosses

between Chinese representatives of MEAM1 and Med species generated 0–2% females (Sun et al. 2011; Li et al. 2012), crosses between Israeli representatives did not generate any females (Elbaz et al. 2010). Different whitefly densities and durations of observation may explain this contrasting result, but potential differences between whitefly populations in the two studies cannot be ruled out (Sun et al. 2011). These crosses were carried out under laboratory conditions, and there are presently insufficient observations of potential crosses in natural conditions to estimate the importance of hybridization in the evolution of nascent species within the *B. tabaci* complex. The geographical spread of *B. tabaci* populations belonging to MEAM1 into many agro-ecosystems throughout the world has provided the opportunity to observe the various outcomes of secondary contacts between previously allopatric divergent populations. Besides the displacement scenario reported from the American continent (Brown et al. 1995) with complete reproductive isolation between the putative species MEAM1 and New World (Xu et al. 2010), hybrids were detected following secondary contact of MEAM1 species and indigenous species from Australia (Austr) using isozymes profiles (Gunning et al. 1997) and the Indian Ocean (IndOc) using microsatellite markers (Delatte et al. 2005). Unexpectedly, no hybrids were detected following secondary contact between MEAM1 and the indigenous species from Spain (Med) using 11 RAPD markers (Moya et al. 2001), although the genetic distance between Med and MEAM1 with respect to mtCOI sequences is lower than the distance between MEAM1 and Austr, and MEAM1 and IndOc (Delatte et al. 2005; Dinsdale et al. 2010), the pairs for which natural hybrids were detected. However, signatures that suggested possible hybridization between MEAM1 and Med were detected in North-American natural populations, although at very low frequency (McKenzie et al. 2012). Thus, from a very comprehensive sampling (4647 individuals analyzed from 517 collections), only 16 were suspected to be hybrids, 15 according to esterase zymogram assays, and 1 according to two microsatellite markers which proved relatively diagnostic for the two putative species. Consistent with field results, hybrids were obtained in experimental crosses with all three pairs of species (Ronda et al. 1999; De Barro and Hart 2000; Sun et al. 2011; H. Delatte pers. comm.) indicating some reproductive compatibility with all of them, including the MEAM1 and Med species pair.

Considering the relatively low genetic distance between Med and MEAM1 and the incomplete reproductive isolation, in this study we further investigated the expected gene flow between these two putative species under natural conditions. *B. tabaci* populations were sampled from the North East of Morocco, where representatives of the two species have been detected previously (Tahiri et al. 2006). A total of 346 female *B. tabaci* were collected in 2003 and 2005 and assigned to MEAM1 (119), Med (225), and a new

putative species (2) based on sequences of the mtCOI gene. MEAM1 and Med individuals were further analyzed using microsatellite markers. A Bayesian clustering analysis of the microsatellite profiles (STRUCTURE) distinguished two groups with 100% consistency with the mtCOI groups, except for two individuals identified as hybrids using complementary population genetic analyses. Hybrid detection is discussed in relation to the secondary endosymbiont infection status determined on a sample of MEAM1 and Med populations and the contrasting outcomes of reported crossing experiments between these species (Elbaz et al. 2010; Sun et al. 2011).

Materials and Methods

B. tabaci sampling and DNA extraction

B. tabaci individuals were collected in 2003 in three regions of the North East of Morocco, that is, from west to east, the regions of Nador, Berkane, and Oujda (Table 1). In each region, two tomato fields 5–10 km distant from each other were sampled, except in the region of Oujda where they were 30–40 km apart. Thirty individuals were collected from each field: the field was divided virtually into 30 equal plots and one female was collected from each plot. The same sampling was repeated in 2005 in the same field if it was planted again with tomato. If not, another tomato field in the neighborhood

was selected. *B. tabaci* samples were conserved in 100% ethanol at −20°C before DNA extraction (Delatte et al. 2005).

Distinction between MEAM1 and Med species using the mtCOI gene

A total of 347 whiteflies were tested for species identity. A set of mtCOI sequences of representatives of MEAM1 and Med species (formerly biotype B and Q, respectively) were aligned to design species-specific polymerase chain reaction (PCR) primers. The two or three last 3′ positions of the designed PCR primers were 100% complementary to the set of sequences of the representatives of one species and not complementary at any of these positions to the set of sequences of the other species. The pair of primers designed to specifically detect individuals of species MEAM1 was as follow; COIB 98 5′-GCT ATA TTG ACT ATT GGT ATT C-3′ and COIB 295c 5′-TCA AAG GCC AAG AGG CCT T-3′. PCR amplification with COIB primers was conducted in a final volume of 20 μL containing 2 μL of DNA extract, 6 mmol/L MgCl$_2$, 100 μmol/L of each dNTP, 0.5 μmol/L of each primer, and 1.2 U of *Taq* polymerase (Eurobio, Courtaboeuf, France). The amplification was carried out with the following parameters: a denaturation step at 95°C for 5 min, then 35 cycles at 94°C for 1 min, 53°C for 1 min 30 sec, 73°C for 1 min 30 sec, and a final cycle at 72°C

Table 1. Description of the populations of *Bemisia tabaci*, species identification using the cytochrome oxidase I (COI) gene and results of STRUCTURE analysis of three unassigned individuals.

Region	Location	Year of sampling	Population code (sample size)	Med[1]	MEAM1[1]	Individuals assigned with CO1 sequence[2] (putative species)	Potential hybrids: assignment coefficient (q) to the Med group, COI type
Oujda	Bouchtat	2003	**A**03 (26)	11	14	A2303 (Med)	
		2005	**A**05 (30)	24	5	A305 (Med)	
	Lbsara	2003	**B**03 (30)	22	4	B603 (Med)	
						B2303 (Med)	
						B2603 (New)	
						B2903 (Med)	
		2005	**B**05 (30)	28	0	B705 (Med)	B705: 0.6626, Med
						B2205 (Med)	B1105: 0.468, Med
Berkane	Zaouia	2003	**N**03 (30)	1	29		
		2005	**N**05 (30)	10	20		
	Essaidia	2003	**E**03 (28)	5	23		E2503: 0.2168, MEAM1
		2005	**E**05 (30)	24	6		
Nador	Bouareg M	2003	**M**03 (27)	25	2		
		2005	**M**05 (30)	26	3	M205 (New)	
	Bouareg K	2003	**K**03 (26)	24	1	K1003 (Med)	
		2005	**K**05 (30)	25	4	K405 (*T. vaporariorum*)	
Total			All (347)	225	111	11	3

[1]Assignment to Med and MEAM1 according to discriminating CO1 primers.
[2]Individuals that could not be assigned because the discriminating PCR test failed. Individuals were coded as follows: field code (A, B, N…) followed by the individual code (a number between 1 and 30) followed by the year of sampling (03 or 05).

for 5 min. The specificity of the primers was tested with DNA extracts of five MEAM1 individuals from Saidia (Morocco) including an individual for which a mtCOI sequence was previously generated (GenBank accession no. AJ517768) and of five Med individuals from Agadir (Morocco) including an individual for which a mtCOI sequence was previously generated (GenBank accession no. AM176573). A fragment of 197 bp was obtained with the control individuals of species MEAM1, and no fragment with control individuals of species Med. The pair of primers designed to specifically detect individuals of species Med was as follow, COIQ 380 5′-GAC ACT TAT TTT GTT GTT GCG-3′ and COIQ 796c 5′-CTC TTT AAA ACT GTG ATT AAG G-3′. PCR amplification with COIQ primers was conducted in a final volume of 20 μL containing 2 μL of DNA extract, 1.5 mmol/L $MgCl_2$, 100 μmol/L of each dNTP, 0.1 μmol/L of each primer, and 1 U of *Taq* polymerase (Eurobio). The amplification was carried out with the following parameters. A denaturation step at 95°C for 5 min, then 35 cycles at 94°C for 1 min, 61.2°C for 1 min, 73°C for 1 min, and a final cycle at 72°C for 5 min. Under these conditions, a fragment of 416 bp was obtained with the control individuals of the species Med, and no fragment with the control individuals of species MEAM1.

Sequence data of the mtCOI gene were generated for individuals for which the diagnostic PCR fragments were not clearly detected. An 817 bp mtCOI fragment was amplified and sequenced following cloning (Tahiri et al. 2006) except for two individuals (B705, K1003) for which a 587 bp sequence was determined directly from a PCR product (Dalmon et al. 2008). The sequences were compared to the sequences of representatives of the most related groups or species of whiteflies using the Optimal Alignment method of DNAMAN (version 5.0; Lynnon BioSoft, Québec, Canada). Sequences generated from representatives of MEAM1 and Med previously collected in Morocco (Tahiri et al. 2006) were included in the comparison. A phylogenetic tree was set up with a Jukes and Cantor distance matrix using the neighbor-joining method of DNAMAN (Saitou and Nei 1987). One thousand bootstrap iterations were performed. Sequences that were divergent from sequences of MEAM1 and Med representatives were compared to the 24 consensus sequences reported to be representative of 24 species at a 3.5% sequence divergence limit (Dinsdale et al. 2010). The percentage of nt divergence was calculated with the BioEdit software (Hall 1999).

Detection and molecular identification of endosymbionts

Secondary endosymbionts were detected with specific PCR primers in a sample of *B. tabaci* individuals representative of each identified species (Gueguen et al. 2010). To check for DNA extraction quality, all individuals were tested for the presence of *Portiera aleyrodidarum*, the obligate endosymbiont. The specific primers were targeted to the 16S rRNA gene for *P. aleyrodidarum*, *Hamiltonella*, *Cardinium*, and *Rickettsia*, the 23S rRNA gene for *Arsenophonus*, and the wsp gene for *Wolbachia*.

Microsatellite genotyping

A total of 344 individuals identified as MEAM1 or Med according to the mtCOI gene were genotyped using seven previously reported microsatellite loci: 11, 53, 145, 177, Bem23, BT4, BT159 (De Barro et al. 2003; Tsagkarakou and Roditakis 2003; Delatte et al. 2006; Dalmon et al. 2008). Two loci (BT4 and BT159) were designed from Med individuals, and five loci (11, 53, 145, 177, and Bem23) from MEAM1 individuals. As *B. tabaci* is a haplo-diploid species, only females were analyzed due to their diploid state. PCR primers, amplification conditions, and allele scoring were as described previously (Dalmon et al. 2008).

Genetic data analysis

Individual-based inferences on population structure

To assess the level of population structure, an individual-based clustering analysis was implemented using the software STRUCTURE 2.3.1 (Pritchard et al. 2000). Assuming panmixia and the population being at Hardy–Weinberg (HW) equilibrium for each locus, this software assigned individuals to K clusters (to be set), following an iterative approach. We used the admixture model where the individual's genome may be partitioned to match the frequency distribution of alleles among the K clusters. The reliability of the K clusters partition was evaluated for K varying from 1 to 6 using the ΔK estimator (Evanno et al. 2005). A total of 20 independent simulations were carried out for each K, and the length of the burn-in period and MCMC (Markov chain Monte Carlo) were set to 10,000 each. To detect putative hybrid individuals, the software was run again with the K determined to be the most reliable. Five independent simulations were done with a burn-in period of 60,000 iterations followed by 600,000 iterations.

Intra-population genetic diversity:

MICRO-CHECKER (Brookfield 1996; Van Oosterhout et al. 2004) was used to check microsatellite data for scoring errors and null alleles. Putative linkage disequilibrium was assessed between all pairs of loci using GENEPOP ([Raymond and Rousset 1995]; GENEPOP on the Web available

at http://genepop.curtin.edu.au/) following the Markov chain method with the default settings. *P*-values were combined for each locus pair across all populations using Fisher's method. In order to adjust the *P*-values for multiple tests for linkage disequilibrium, we used the false discovery rate (FDR) procedure (Benjamini and Hochberg 1995), which controls the proportion of significant results ("false discovery rate") instead of controlling the chance of making even a single error. The resulting adjusted *P*-values are called *Q*-values. This procedure is implemented in the R package QVALUE (Storey and Tibshirani 2003). Mean number of alleles per locus ($\bar{A}$), observed heterozygosities (H_O), and unbiased expected heterozygosities (H_E, calculated following Nei [1973]) were calculated using GENETIX 4.05.2 (Belkhir et al. 2004). To test for HW equilibrium within populations, Weir and Cockerham estimators of F_{IS} (Weir and Cockerham 1984) were calculated using FSTAT 2.9.3.2 (Goudet et al. 1996) for each locus and overall loci. The estimation of exact *P*-values of the difference of F_{IS} from 0 (departure from HW equilibrium) was tested with the Markov chain method of GENEPOP using the default settings.

Population differentiation tests

Pairwise F_{ST} estimates were obtained using GENEPOP. A hierarchical analysis of molecular variance (AMOVA) was performed, using HIERFSTAT (Goudet 2005), to test for geographical structure. This test was performed on eight Med populations using four hierarchical levels: individuals within populations, among populations within regions, among regions within years, and among years. Associated *P*-values were calculated with 10,000 permutations of the appropriate unit as advised in (Goudet 2005). A correspondence analysis (COA) was performed using GENETIX 4.05.2 (Belkhir et al. 2004) to visualize the major axes of genetic variation within the sample, and the position of the groups and the potential hybrid individuals detected by the STRUCTURE software along these axes.

Hybrid detection

We used the NEWHYBRIDS (1.1) software that implements a Bayesian assignment approach dedicated to the quantification of hybridization events from individual multilocus genotypes (Anderson and Thompson 2002). The strength of this method is that no pedigree information is required, allowing application to natural populations (Manel et al. 2005). Assignments were performed considering two hybridization schemes. First, we focused on the first generation of hybridization between the parental species (MEAM1 and Med), hence assuming only three classes (the two parental classes and the F1 hybrid).

Second, we considered a more complex scheme including the second generation of hybridization, even if the detection of backcrosses may require many more loci (almost 50) that we actually amplified (Vähä and Primmer 2006). This second scheme consisted of six classes (the two parental classes, F1 and F2 hybrids and the two types of backcrosses). We used the program without any prior information on parental allele frequencies. As the Markov Chains converged rapidly, we erased the initial sweeps after a short burn-in period of 5,000 iterations. Average memberships were computed over more than 600,000 iterations after burn-in. Five independent runs were performed for each hybridization scheme (3 or 6 classes) to check consistency of the assignments.

Results

Identification of genetic groups in the sampled whiteflies based on mtCOI sequences and microsatellite profiles

Using Med- and MEAM1-specific PCR primers targeted to the mtCOI gene, 225 individuals were identified as belonging to Med and 111 to MEAM1, whereas 11 exhibited nonconclusive profiles (Table 1). Nucleotide sequence data generated for these 11 individuals revealed that eight of them belong to the West Mediterranean Q1 group (Chu et al. 2008; Tsagkarakou et al. 2012) within Med, one belongs to a distinct species of whitefly, *Trialeurodes vaporariorum*, and two to a new genetic group of *B. tabaci* that is sister to the putative "Italy" species (Fig. 2). The sequenced mtCOI fragment of these two latter individuals shared 99.2% identity. The genetic distance between the representatives of this new group and each of the consensus sequences of the 24 putative species of *B. tabaci* was above the 3.5% divergence threshold characteristic of a new putative species (Dinsdale et al. 2010). It was also above the 3.5% divergence threshold with the putative species "Ru" recently detected in Italy (Parrella et al. 2012). The smallest divergence was obtained with the consensus of the Italy group, with 5.5–6.4% divergence. One individual of this new putative species was collected in 2003 in the region of Oujda and the other in 2005 in the region of Nador.

If individuals of Med and MEAM1 are sufficiently isolated regarding reproduction, mitochondrial-based differentiation should be supported by markers targeted to the nuclear genome. Of the 344 individuals of *B. tabaci* belonging to Med and MEAM1 groups according to mtCOI data (see above), 340 amplified for all seven microsatellite loci and four amplified for only six loci. Bayesian clustering analysis performed for *K* varying from 1 to 6 indicated that the likelihood of assignment
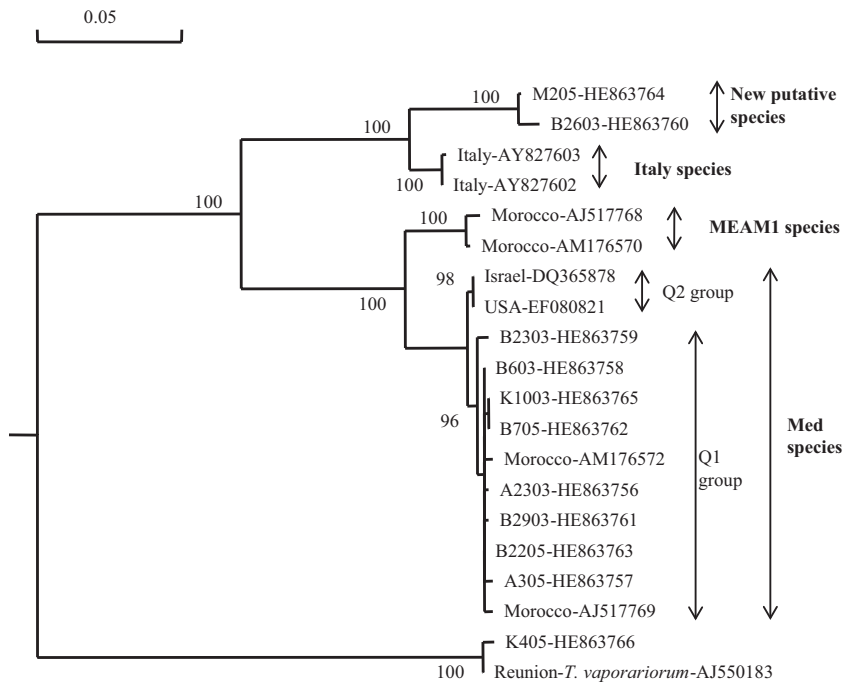
Figure 2. Rooted neighbor-joining tree based on a 587-nucleotide fragment of the mtCOI gene of *Bemisia tabaci*. The sequences generated in this study are identified by their accession number and the code number of the corresponding whitefly (see Table 1). The GenBank sequences are identified by their accession number and the country where the corresponding whitefly was collected. The Western Mediterranean Q1 group and the Eastern Mediterranean Q2 group are indicated according to previous reports (Chu et al. 2008). The scale measures the Jukes and Cantor distance between sequences. Numbers associated with nodes represent the percentage of 1,000 bootstrap iterations supporting the nodes; only percentages ≥95% are indicated.

increased from ln p(D) = −7997 for $K = 1$ to ln p(D) = −5611 for $K = 2$, where it reached a plateau at ln p(D) ≈ −5500 up to $K = 6$ (Table S1). Accordingly, the ΔK index was the highest for $K = 2$ ($\Delta K_2 = 690$), which meant that adding a cluster from $K = 2$ to $K = 3$ did not add much more to the likelihood of assignment and indicated that two was the most reliable estimation of K. All but three individuals were unambiguously assigned (assignment coefficient $q > 0.9$) to the two genetic clusters, which was 100% consistent with the mtCOI assignments (Fig. 3A). Given this consistency between mtCOI profile and genetic clustering, these two clusters are subsequently denoted as Med and MEAM1 species. The coefficients of assignment to the Med species (q) lie between 0.2 and 0.7 for the three unassigned individuals (Table 1). According to mtCOI gene sequences, two of them were of the Med type and one was the MEAM1 type.

Endosymbiont identification

Endosymbionts were detected in a sample of *B. tabaci* individuals representative of the three identified species (see above): 20 individuals belonging to Med, 18 belonging to MEAM1, and the two individuals of the new species. We also included the three individuals that were not assigned to MEAM1 and Med as defined by the STRUCTURE analysis ($0.2 < q < 0.8$) and the unique *T. vaporariorum* individual of the collection (Table 1). The obligate endosymbiont, *P. aleyrodidarum*, was positively detected by PCR for all 44 whiteflies tested, indicating that DNA extracts were suitable for PCR amplification. Three distinct endosymbiont species were detected in the 43 *B. tabaci* individuals: *Hamiltonella* (36/43), *Wolbachia* (6/43), and *Cardinium* (3/43) (Table 2). Five of the 43 *B. tabaci* individuals and the unique representative of *T. vaporariorum* were negative for all the secondary endosymbionts tested. There was no strict relationship between species identity of the individuals tested and the endosymbiont species they bore. The endosymbiont infection status did not differ significantly between Med and MEAM1 individuals (Fisher's exact test on the abundances of endosymbiont profiles, $P = 0.13$). *Hamiltonella* was the only secondary endosymbiont detected in MEAM1 individuals, whereas three distinct endosymbiont species (*Hamiltonella* plus *Cardinium* and *Wolbachia*) were detected in Med individuals (Table 2). Interestingly, the two Med individuals suspected to be hybrids were of the rare C group, which exhibits an infection with *Wolbachia* in addition to *Hamiltonella*. A specific endosymbiont profile was detected in the new species with both individuals positive only for *Wolbachia*.

Spatial distribution and evolution over time of individuals of MEAM1 and Med species

The abundance of individuals belonging to MEAM1 species compared to those belonging to Med species (assessed from the distribution of the 341 unambiguously
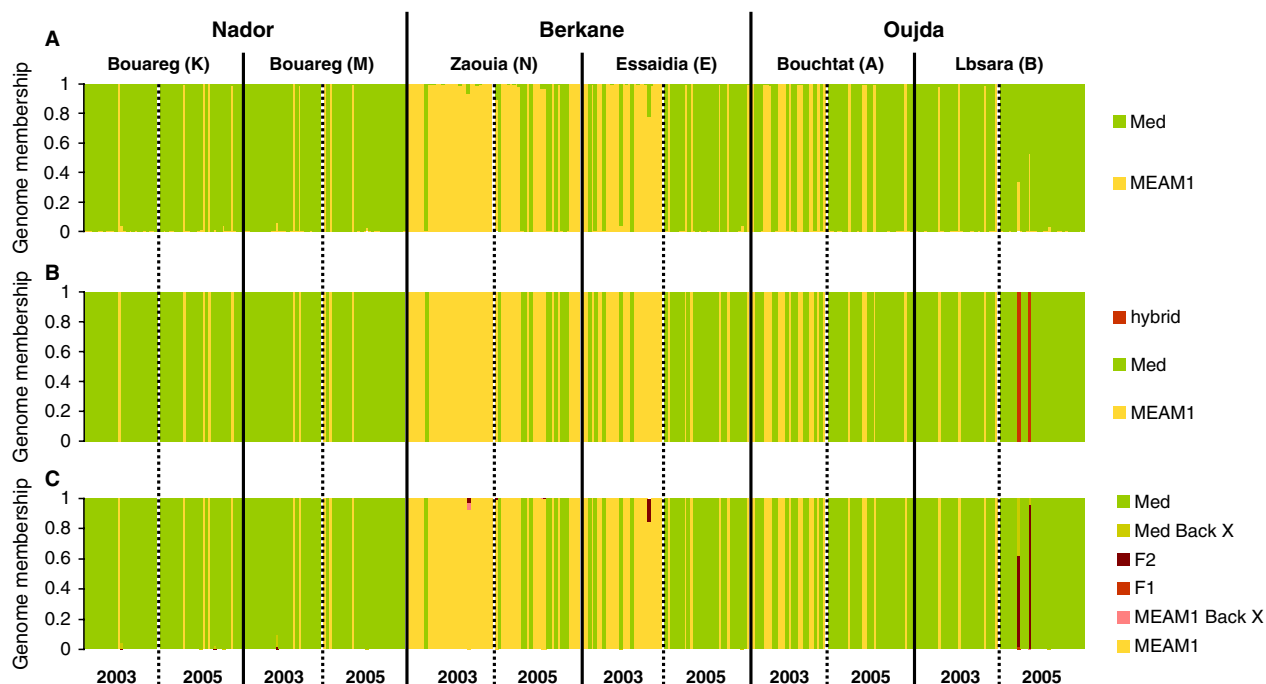
Figure 3. Assignment results of the 344 *Bemisia tabaci* individuals of the MEAM1 and Med species. (A) STRUCTURE results using the admixture model for *K* = 2 genetic groups. (B) NEWHYBRIDS results considering three classes: the two parental classes and the F1 hybrid. (C) NEWHYBRIDS results considering six classes: the two parental classes, F1 and F2 hybrids, and backcrosses. Individuals were sorted identically for all graphs, according to their assignment to Med and MEAM1 species and the detection of mixed ancestry. Each vertical line represents an individual whose genome is partitioned into *n* colored segments (with *n* being the number of classes). Color correspondence is given on the right side of each graph. Segment lengths are proportional to the individual's mean membership coefficients in each genetic group (averaged over five runs).

Table 2. Endosymbiont infection status of individual whiteflies using specific PCR primers.

mtCOI	Structure assignation	Number of individuals	Endosymbiont profile	Cardinium	Wolbachia	Hamiltonella	Rickettsia	Arsenophonus
Med	$q > 0.90$	12	A	-	-	+	-	-
	$q > 0.90$	3	B	+	-	+	-	-
	$q > 0.90$	2	C	-	+	+	-	-
	$q > 0.90$	3	D	-	-	-	-	-
	$q = 0.66$	1	C	-	+	+	-	-
	$q = 0.47$	1	C	-	+	+	-	-
MEAM1	$q \geq 0.90$	16	A	-	-	+	-	-
	$q > 0.90$	2	D	-	-	-	-	-
	$q = 0.78$	1	A	-	-	+	-	-
New species		2	E	-	+	-	-	-
T. vaporariorum		1	D	-	-	-	-	-

The results are presented according to the classification of the whiteflies as determined by mtCOI sequences and by the assignment coefficient to the two genetic groups distinguished by STRUCTURE.

assigned individuals) was found to differ significantly across sampled fields and years of collection (Fisher's exact test *P* < 0.0001). In the region of Berkane, a majority of the individuals were of the MEAM1 species for all but the Essaidia 2005 sampling (Table 1, Fig. 4). Conversely, in the region of Oujda, a majority of the

individuals belonged to the Med species for all samplings except the Bouchtat 2003 sampling, which displayed an even proportion of MEAM1 and Med individuals. The proportion of MEAM1 representatives decreased significantly from 2003 to 2005 in three fields: two fields in the region of Berkane (Zaouia [N] and Essaidia [E]) and in
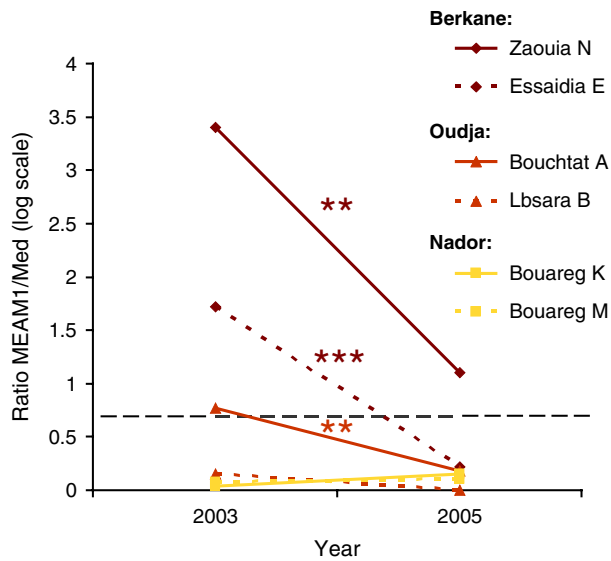
Figure 4. Temporal evolution of the MEAM1 to Med ratio in each field. The two fields from the same region are represented by the same color and same symbol (diamonds: Berkane, triangles: Oudja, squares: Nador). Significant differences in MEAM1 and Med counts between 2003 and 2005 are indicated by asterisks (Fisher's exact test, ***P < 0.001, **P < 0.01). The horizontal dashed line indicates the 1:1 ratio.

the Bouchtat (A) field in the region of Oujda (Fig. 4). This trend was not detected in the Nador region, or in the Lbsara (B) field in the region of Oujda, where most individuals belonged to the Med species (maximum of four MEAM1 individuals per sampling, Table 1).

Genetic characteristics of MEAM1 and Med species

Overall, strong differences were detected for each locus between individuals belonging to MEAM1 species and those belonging to Med species. The number of alleles over Med individuals ranged from 5 to 26 per locus, whereas the range was from 1 to 10 for MEAM1 individuals (Table 3). Significant heterozygote deficiencies were detected at loci 53, Bem23, and Bt4 for Med species and at all but one locus for MEAM1 species. Locus Bt159 was monomorphic in MEAM1 species. In all, individuals of MEAM1 species were less diverse and showed a larger heterozygote deficiency compared with individuals of Med species, despite five of the seven loci were designed from MEAM1 species.

Genetic analyses were subsequently carried out within four MEAM1 populations and 10 Med populations (Table 4). Each of these 14 populations was composed of at least 10 individuals collected in the same field and in

the same year. The three individuals that were unassigned with STRUCTURE were not included in the analysis. No significant linkage disequilibrium was detected across either Med or MEAM1 populations. The average allelic richness (Ar) per locus within Med populations ranged from 4.60 to 5.31 (Table 4) and was significantly higher than the average Ar within MEAM1 populations, which ranged from 2.43 to 3.31 (two-sided *P*-value = 0.002). The H_O values of Med populations ranged from 0.4505 ± 0.2460 to 0.5655 ± 0.2392 (Table 4) and were significantly higher than the H_O values of MEAM1 populations, which ranged from 0.1719 ± 0.1852 to 0.2286 ± 0.2430 (two-sided *P*-value = 0.002). The H_E values of Med populations ranged from 0.5291 ± 0.2323 to 0.6127 ± 0.2037 (Table 4) and were significantly higher than the H_E values of MEAM1 populations, which ranged from 0.2434 ± 0.3104 to 0.3061 ± 0.2718 (two-sided *P*-value = 0.002). For all populations, we observed $H_O < H_E$, which led to positive values of F_{IS}. F_{IS} values were significantly different from 0 for most populations. The heterozygote deficiency could not be attributed to any particular locus, as $H_O < H_E$ for all loci (Table 3); H_O was, however, equal to H_E with the four MEAM1 populations at locus 11 because it was monomorphic for three of the four populations, and at locus Bt159 because it was monomorphic for all of them. MICRO-CHECKER analysis revealed the presence of potential null alleles in most of the populations. Slight changes in allele size during PCR (stuttering) also detectable with MICRO-CHECKER were suspected in only two Med populations (Bouareg M sampled in 2003 and 2005) for locus Bt4.

Populations of the Med species were differentiated from populations of the MEAM1 species by very high F_{ST} values (0.431–0.512) (Table S2). However, genetic differentiation was low among Med populations (≤0.021) and MEAM1 populations (≤0.048). A hierarchical AMOVA was performed but only with Med populations from Oujda and Nador because the collections from all fields in these two regions and from both years contained at least 10 Med individuals (Table 1). The highest variance component was seen among individuals within fields (61.7%) (Table S3). All the other levels correspond to very low variance components; the inter-region within years component being negative. The only component to be significant was among individuals within fields.

To test for a genetic structure within MEAM1 or Med that would not depend on the geographic sample, Bayesian clustering analysis was performed within each putative species. The likelihood of assignment for *K* varying from 1 to 6 did not increase substantially for MEAM1: ln p(D) = −1049 for *K* = 1 to ln p(D) = −1065 for *K* = 6. The likelihood of assignment for *K* varying from 1 to 6 increased very slightly for Med and no obvious break in slope was detected; ln p(D) = −4491 for *K* = 1 to ln p(D) = −4804 for *K* = 6.

Table 3. Characteristics of the two genetic groups.

Locus	Med Allele range	Ar^1	$H_E{}^2$	$H_O{}^2$	$F_{IS}{}^3$	MEAM1 Allele range	Ar^1	$H_E{}^2$	$H_O{}^2$	$F_{IS}{}^3$
11	171–186	4.9	0.539	0.519	0.036	177–182	3.0	0.018	0.018	−0.002
53	133–179	17.1	0.904	0.785	0.131**	127–157	11	0.813	0.564	0.307***
145	169–179	5.7	0.336	0.326	0.028	177–220	6.0	0.184	0.162	0.121*
177	241–259	8.2	0.456	0.438	0.039	241–245	4	0.412	0.351	0.148***
Bem23	178–426	20.5	0.765	0.601	0.215***	175–222	8.0	0.462	0.225	0.514***
BT4	282–318	11.0	0.426	0.264	0.381***	286–305	6	0.241	0.155	0.36***
BT159	267–309	13.2	0.719	0.644	0.105	267	1	-	-	-
All		11.5	0.592	0.511	0.137***		5.6	0.304	0.211	0.309***

$*P < 0.05$; $**P < 0.01$; $***P < 0.001$.
[1] Ar determined according to a minimum of 110 diploid individuals using FSTAT application.
[2] H_O and H_E (Nei's estimation) as determined with GENETIX application.
[3] F_{IS} as determined with FSTAT application and $P(F_{IS})$ with GENEPOP application.

Table 4. Genetic characteristics of the populations.

Species	Year	Population	N	Ar	F_{IS}	$P(F_{IS})$	H_E	SE	H_O	SE
Med	2003	A03	12	5.00	0.191	0.084	0.5873	0.1513	0.5000	0.1735
		B03	25	4.78	0.096	0.007**	0.5877	0.2038	0.5429	0.1787
		M03	25	5.26	0.153	<0.001***	0.6127	0.2037	0.5314	0.2374
		K03	25	5.17	0.136	0.013*	0.5815	0.1982	0.5143	0.1938
	2005	A05	25	5.06	0.204	<0.001***	0.5815	0.2195	0.4743	0.1607
		B05	28	4.60	0.071	0.246	0.5291	0.2323	0.5011	0.1836
		N05	10	5.14	0.137	0.747	0.5779	0.1847	0.5286	0.1890
		E05	24	5.31	0.060	0.034*	0.5882	0.2214	0.5655	0.2392
		M05	26	5.09	0.208	<0.001***	0.5554	0.2332	0.4505	0.2460
		K05	25	4.83	0.149	0.005**	0.5578	0.1939	0.4857	0.1574
MEAM1	2003	A03	14	2.43	0.321	<0.001***	0.2434	0.3104	0.1735	0.2731
		N03	29	2.92	0.281	<0.001***	0.2880	0.2668	0.2118	0.2102
		E03	22	2.85	0.457	<0.001***	0.3061	0.2718	0.1719	0.1852
	2005	N05	20	3.31	0.225	0.033*	0.2861	0.2786	0.2286	0.2430

H_O, H_E (Nei's estimation) and standard error across loci (SE) as determined with GENETIX application. F_{IS} and Ar as determined with FSTAT application. $P(F_{IS})$ as determined with GENEPOP application after 10,000 permutation tests ($*P < 0.05$; $**P < 0.01$; $***P < 0.001$).

Detection and characterization of hybrid individuals

The three unassigned individuals detected with Bayesian clustering analysis (see above) may be hybrid individuals generated from crossings between MEAM1 and Med. To confirm their status, we first used a genetic COA that allows a graphical representation of the major axes of genetic differentiation within the whole collection. Although this analysis did not make use of the groups identified by the STRUCTURE software, there is a 100% correspondence between STRUCTURE assignments and the position of individuals along the first COA axis, which represents 9.55% of total genetic variation (Fig. 5). Moreover, two of the three putative hybrid individuals detected by STRUCTURE (B705, B1105) are clearly sitting apart from and between the two clusters, whereas the remaining one

(E2503) was associated with the MEAM1 cluster although closest to the Med cluster. These relative positions were consistent with the STRUCTURE assignment coefficients of these individuals, which were in the range of 0.35–0.45 for B705 and B1105 but close to 0.2 for E2503 (Table 1). The two putative hybrid individuals classified as Med according to mtCOI sequences (B705, B1105) were revealed to have alleles that were detected only with MEAM1 individuals (private allele), for example allele 222 at locus Bem23 (Table S4). Conversely, the putative hybrid individual E2503, classified as MEAM1 according to mtCOI sequences, was revealed to have an allele at locus 53 (141) that was private to Med individuals. Similarly, some of the other alleles of the putative Med hybrids were detected frequently in MEAM1 but rarely in Med (for example allele 267 at locus BT159) and vice versa for the MEAM1 hybrid. The fact that individual
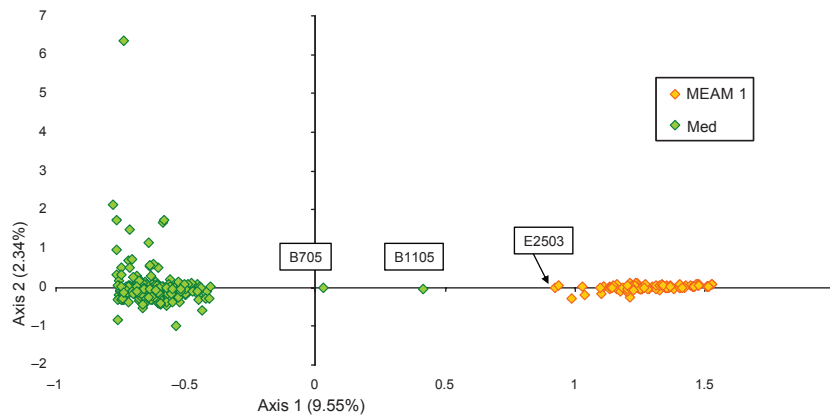
Figure 5. Genetic correspondence analysis of the 344 *Bemisia tabaci* individuals of the MEAM1 and Med species. First factorial plane of the genetic correspondence analysis based on the amplification profiles at seven microsatellite loci. Individuals that were previously identified as Med species according to mtCOI sequence data are indicated as green squares whereas those of MEAM1 species are indicated as yellow squares. The code numbers of the three individuals unassigned by Structure analysis are indicated.
[Correction added on 05 July 2013, after first online publication: The coloured squares of Med and MEAM1 populations have been switched. Med are in green and MEAM1 in yellow.]

B1105 was homozygous for a private allele at locus 11 and individual E2503 was homozygous for private alleles at loci 11, MS145, Bem23, and 53 (Table S4), suggests they are not F1 hybrids. Overall, the assignment profiles provided by NewHybrids were very close to those conferred by Structure analysis (Fig. 3), but helped to characterize further the hybrid status of these three individuals (B705, B1105, and E2503). Assuming three classes (two parental groups plus F1 hybrid) NewHybrids analysis inferred indisputably that two individuals (B705, B1105) were hybrids (Fig. 3B). The analysis based on a more complex hybridization scheme (Fig. 3C) revealed that individual B1105 would be a F2 hybrid, in accordance with its homozygous allelic status at several loci (Table S4). Classification of individual B705 is less clear (somehow intermediate between F2 and Med backcross), maybe because it would result from a third generation after initial hybridization. Likewise, individual E2503 was found to have a part of its genome assigned to the F2 hybrid class, but with the majority of its genome assigned to the MEAM1 species.

Discussion

Based on genetic and biological data it was concluded previously that *B. tabaci* is composed of at least 24 morphologically indistinguishable species (Dinsdale et al. 2010; De Barro et al. 2011). Although inter-species crossing experiments showed that several of these putative species, including MEAM1 and Med, can generate F1 hybrids (reviewed in Liu et al. [2012]), there are presently insufficient data to assess the evolutionary consequences of these crossings in nature, particularly when one of

these putative species invades a region occupied by another. Therefore, as MEAM1 has reportedly invaded the North East of Morocco, which was occupied by Med (Tahiri et al. 2006), we analyzed the genetic structure of 346 *B. tabaci* individuals sampled in 2003 and 2005 in this region. Our results confirmed that MEAM1 and Med were present in this region. Moreover, barring the field sampled in Lbsara in 2005 (B05), representatives of both species were detected within all fields, which is an indication that inter-species contacts were apparently not a limiting factor for potential inter-species crossings. With the exception of the three unassigned individuals, the mtCOI-based distinction between MEAM1 and Med individuals was 100% consistent with the two groups distinguished by Bayesian clustering analysis, revealing high reproductive isolation between MEAM1 and Med, and supporting their proposed separation as two distinct species (Dinsdale et al. 2010; De Barro et al. 2011). Among the three unassigned individuals, two were indisputably hybrids according to several lines of evidence: (i) their shared ancestry to the two species assessed with Structure; (ii) the fact that they mix private alleles from both species; and above all (iii) the results of the dedicated NewHybrids analysis, which clearly revealed these individuals to be the progenies of MEAM1-Med inter-species crossings.

Identification of a new putative species within the Bemisia tabaci species complex

Among the 346 *B. tabaci* individuals identified in this study, two did not belong to the previously reported MEAM1 and Med species. Neither did they belong to any of the 24 species distinguished according to the 3.5%

divergence threshold (Dinsdale et al. 2010). Consistently, no amplification product was detected with mtCOI primers designed to specifically detect MEAM1 and Med individuals. Members of the species "Italy", to which they are most closely related, were isolated on *Euphorbia characias*, and their progeny could not be reared on tomato – a plant species on which the adults did not survive for more than 10 days (Simon et al. 2003). It is therefore likely that the two sampled individuals of this new putative species were not generated on tomato plants but landed accidentally on the sampled tomato field. Further samplings are needed to detect plant species that allow progeny development of this newly identified putative species. The two representatives of the new putative species were positive only for *Wolbachia* – an infection status that is extremely rare in the Mediterranean but is easily detected in Cambodia (Chiel et al. 2007; Gueguen et al. 2010). Interestingly, the Italy species, to which this new putative species is most related, is sister to Asian species according to mtCOI comparisons (Boykin et al. 2007; Dinsdale et al. 2010). Other representatives of the new putative species in Morocco as well as representatives of the species "Italy" should be tested for their endosymbiont infection status to test if this particular infection with *Wolbachia* is inherited from a common ancestor of Asian and Mediterranean *B. tabaci* populations. Based on a global mtCOI-tree showing a sister relationship between the Italy species and Asia species, it was suggested that this phylogeographical inconsistency was due to human movement from Asia (Boykin et al. 2007). However, the detection of a second putative species from Italy – Ru – which is, together with "Italy", basal to the clade including the putative species from Asia and Australia, has suggested an alternative possible origin of the Asian/Australian populations of *B. tabaci* from ancient introduction of whiteflies from the European continent (Parrella et al. 2012). The detection of representatives of a putative species from Morocco that is most closely related to "Italy" is consistent with a Mediterranean origin of all these recently detected putative species related to Asian species. The fact that one representative of this new putative species was detected in 2003 in Oujda, the most eastern region of the sampling area, and the other in 2005 in Nador, the most western region (Table 1) suggests that this new putative species was not introduced recently into Morocco but might indeed be indigenous to Morocco. We therefore propose to name the new putative species identified in this study "Morocco".

MEAM1 was introduced recently in Morocco

The average Ar per locus in Med populations was significantly higher than those in MEAM1 populations (Tables 3 and 4), which is consistent with a more ancient settlement of Med in Morocco compared with MEAM1 as previously suggested according to the geographic distribution of MEAM1, which was, unlike Med, detected only in the North East of the country (Tahiri et al. 2006). The endosymbiont infection status determined for a sample of *B. tabaci* individuals distinguished MEAM1 and Med at the population level, as previously shown with other MEAM1 and Med samples (Gueguen et al. 2010). The unique endosymbiont species detected on MEAM1 individuals compared to the three distinct species on Med was consistent with the suggested recent introduction of MEAM1 in Morocco and with an older settlement of Med. MEAM1 was already detected in 2001 in the region of Nador and Berkane from several host species and was found to be the dominant species in the small sample analyzed (10 MEAM1/1Med) (Tahiri et al. 2006). Since 2001, the proportion of MEAM1 has apparently decreased, particularly between 2003 and 2005 on tomato (Table 1). Further samplings are needed to confirm or invalidate a possible regression of MEAM1. An opposite situation was detected in China and Australia where the relative frequency of the invaded MEAM1 has increased in comparison with the relative frequency of the indigenous putative species (Liu et al. 2007). The Moroccan results are apparently more consistent with the Spanish situation where after a period of coexistence with MEAM1, the indigenous Med became the only species detected in the Iberian Peninsula in surveys conducted after the year 2000 (Simon et al. 2007). MEAM1 was reported to have a higher fitness than Med under laboratory conditions (Pascual and Callejas 2004). The lower fitness of MEAM1 relative to Med in two distinct environments, that is, Spain (Simon et al. 2007) and apparently Morocco (this study), indicate that fitness tests carried out under laboratory conditions are not necessarily relevant to predicting the outcomes of competition in the more complex natural environment where for example insecticides may play crucial role in affecting the distribution and frequency of whitefly species across regions (Crowder et al. 2010).

Geographic structuring

As expected from putative species, the F_{ST} between MEAM1 and Med is high, up to 0.5. On the contrary, the F_{ST} between the geographic populations within each species was very low, which is indicative of relatively long distance migrations, even between the most distantly sampled regions of Oujda and Nador (~100 km apart). Consistently, most of the genetic diversity of the Med populations analyzed with AMOVA was detected within fields. Contrary to our expectations, the mountain

areas and the nonhost crops (cereals and citrus orchards) separating the three sampled regions did not result in high-population differentiation between fields. Similarly, high migration was previously inferred from the genetic structure between Med populations from France and mainland Greece (Dalmon et al. 2008; Tsagkarakou et al. 2012). Considerable migration was also detected between representatives of the putative species "Australia" collected in North and South Australia (De Barro 2005). However, in this latter study, the analysis of samples collected further north, in tropical regions, showed significant correlation between genetic and geographical distance, suggesting that migration is limited. It is therefore possible that the year-round availability of palatable hosts in tropical regions limits the need for migration.

Evidence of gene flow between MEAM1 and Med species

Bayesian clustering analysis (STRUCTURE and NEWHYBRIDS) and the graphical representation of the genetic COA did detect two natural hybrids that are the progeny of interspecies crosses between MEAM1 and Med. In their simulation study, Vähä and Primmer (2006) advocated using at least 12 loci for an accurate identification of hybrids from Bayesian clustering analysis when F_{ST} between parents equals 0.21 (48 loci with F_{ST} = 0.12, respectively). Given the high F_{ST} level between MEAM1 and Med species (F_{ST} = 0.5, Table S2) and the consistency of STRUCTURE and NEWHYBRIDS assignments [already pointed out by Vähä and Primmer (Vähä and Primmer 2006)], we are confident that individuals B705 and B1105 are hybrids although we characterized them at only seven loci. The case of individual E2503 is more puzzling, but as previously indicated (Vähä and Primmer 2006), the accurate detection of backcrosses would require far more loci (more than 50), otherwise they are often misclassified as purebred individuals. Our analysis could thus provide a minimal estimate of the actual gene flow between MEAM1 and Med species as it is biased toward the detection of first generation hybrids and has only limited inference power on backcross events (Vähä and Primmer 2006). The detection of hybrids between MEAM1 and Med in natural populations in Morocco is consistent with the 16 hybrids recently suspected between these putative species in natural North-American populations using esterase zymogram assays and two microsatellite markers (McKenzie et al. 2012). These results confirm that closely related species may occasionally produce hybrids (Mallet 2005; Liu et al. 2012) and are consistent with the hybrids previously obtained between MEAM1 and Med under experimental conditions (Ronda et al. 1999; Sun et al. 2011).

According to NEWHYBRIDS analysis, none of the hybrids were from an F1 generation, suggesting that F1 hybrid females were fertile within the natural populations of North East Morocco. Fertility of F1 females is likely to be rare, as suggested by experimental crosses between Chinese populations revealing that none of 57 F1 hybrid females tested were fertile (Sun et al. 2011). However, it may be possible that the Moroccan MEAM1 and Med populations were more compatible with generating fertile hybrids than the Chinese populations due, for example, to potentially better heterospecific protein interactions (Servedio 2001) or to a particular endosymbiont infection status of some individuals. It was indeed noted that the two putative hybrids derived from Med females belonged to a minor group of Med individuals that were positive for *Wolbachia* detection. However, it is not known how *Wolbachia* may have an effect on fertility, if any. It was also suggested that once a fertile F1 hybrid is generated in nature, the backcrossing to one of the parental putative species is apparently much more straightforward because the genetic gulf to be traversed is thought be halved (Mallet 2005). The two Med hybrids were detected in the only field where no MEAM1 representatives were found, which is apparently surprising but not unexpected because the hybrids were not from an F1 generation. Consistently, MEAM1 individuals were detected in the same field or at least in a neighboring field 2 years before.

The detection of fertile hybrids between MEAM1 and Med needs to be confirmed because it can presently not be excluded that alleles thought to be private to one species, and which reveal putative hybrids, may be detectable in other species but at an extremely low frequency. The available data suggest, however, that a limited gene flow between MEAM1 and Med still exists and that it may contribute to the evolution of the nascent Med and MEAM1 species by hybridization. Crossing experiments between Israeli representatives of MEAM1 and Med did not produce any hybrids (Elbaz et al. 2010) whereas F1 hybrids were obtained with Spanish and Chinese representatives (Ronda et al. 1999; Sun et al. 2011). As previously suggested (Sun et al. 2011), these contrasting results may be explained by potential differences between the whitefly populations used in the two studies. This explanation may be particularly relevant for the populations used for the Med parents. Indeed, the Israeli Med populations with which no hybrids were obtained belonged to the Eastern Mediterranean Q2 group (Elbaz et al. 2010), whereas the Spanish and the Chinese Med populations with which F1 hybrids were detected were from the Western Mediterranean Q1 group (Chu et al. 2008; Tsagkarakou et al. 2012). Consistently, the Moroccan Med populations in which putative hybrids were detected were from the Q1 group (Fig. 2), a group which

was also present in North America where hybrids were suspected (McKenzie et al. 2012). It is noteworthy that Q1 and Q2 could be distinguished not only by their mtCOI sequences but also by their endosymbiotic communities, including some species known to cause reproductive manipulations in arthropods: *Wolbachia*, *Rickettsia,* and *Arsenophonus* (Chiel et al. 2007; Gueguen et al. 2010). The fact that no hybrid was detected in field populations of MEAM1 and Med of the Q1 group in Tunisia may be due to the fact that spatial overlap between the two putative species is very rare (Saleh et al. 2012). It is worth noting that, under experimental conditions, the success of cross mating between MEAM1 and Med was indeed shown to depend on the distance between individuals according to the size of the cages (Li et al. 2012). The contrasting degrees of genetic isolation between Med and MEAM1 may also be explained by various durations of MEAM1 and Med secondary contact following allopatric speciation. The reinforcement process may be more advanced in Israel (no hybrids detected) (Elbaz et al. 2010) than in China and Morocco (hybrids detected) because contact between MEAM1 and Med is thought to have lasted for longer in Israel (at least since 1991 [Horowitz et al. 2003]) than in China (Chu et al. 2006) and Morocco (Tahiri et al. 2006), where the first simultaneous presence of the putative species was detected after 2000. To check this hypothesis, the degree of genetic isolation between naive MEAM1 and Med representatives that have never come into contact with each other should be determined and compared with that of representatives that have already experienced secondary contact for some time.

Although we provide evidence of a gene flow between Med and MEAM1 in field conditions, the ecological significance of this MEAM1 and Med hybridization is presently not known and should direct future research. However, testing the viability of natural hybrids is very challenging because it would need to be able to spot the rare hybrids with a nondestructive technique. The ecological consequences of hybridization may also be addressed by testing if such hybrids will still be detected in this area in the coming years.

Acknowledgments

References

Anderson, E. C., and E. A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217–1229.

Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. 2004. GENETIX 4.05.2, logiciel sous Windows™ pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier, France.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate – a practical and powerful approach tu multiple testing. J. R. Stat. Soc. Series B Methodol. 57:289–300.

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

Brookfield, J. F. Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Mol. Ecol. 5:453–455.

Brown, J. K., D. R. Frohlich, and R. C. Rosell. 1995. The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or a species complex? Annu. Rev. Entomol. 40:511–534.

Byrne, F. J., M. Cahill, I. Denholm, and A. L. Devonshire. 1995. Biochemical identification of interbreeding between B-type and non B-type strains of the tobacco whitefly *Bemisia tabaci*. Biochem. Genet. 33:13–23.

Chiel, E., Y. Gottlieb, E. Zchori-Fein, N. Mozes-Daube, N. Katzir, M. Inbar, et al. 2007. Biotype-dependent secondary symbiont communities in sympatric populations of *Bemisia tabaci*. Bull. Entomol. Res. 97:407–413.

Chu, D., Y. J. Zhang, J. K. Brown, B. Cong, B. Y. Xu, Q. J. Wu, et al. 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.

Chu, D., F. H. Wan, Y. L. Tao, G. X. Liu, Z. X. Fan, and Y. P. Bi. 2008. Genetic differentiation of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) biotype Q based on mitochondrial DNA markers. Insect Sci. 15:115–123.

Crowder, D. W., A. R. Horowitz, P. J. De Barro, S. S. Liu, A. M. Showalter, S. Kontsedalov, et al. 2010. Mating behaviour, life history and adaptation to insecticides determine species exclusion between whiteflies. J. Anim. Ecol. 79:563–570.

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. J. 2005. Genetic structure of the whitefly *Bemisia tabaci* in the Asia-Pacific region revealed using microsatellite markers. Mol. Ecol. 14:3695–3718.

De Barro, P. J., and P. J. Hart. 2000. Mating interactions between two biotypes of the whitefly, *Bemisia tabaci* (Hemiptera: Aleyrodidae) in Australia. Bull. Entomol. Res. 90:103–112.

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., J. W. H. Trueman, and D. R. Frohlich. 2005. *Bemisia argentifolii* is a race of *B. tabaci* (Hemiptera: Aleyrodidae): the molecular genetic differentiation of *B. tabaci* populations around the world. Bull. Entomol. Res. 95:193–203.

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.

Delatte, H., B. Reynaud, M. Granier, L. Thornary, J. M. Lett, R. Goldbach, et al. 2005. A new silverleaf-inducing biotype Ms of *Bemisia tabaci* (Hemiptera: Aleyrodidae) indigenous to the islands of the south-west Indian Ocean. Bull. Entomol. Res. 95:29–35.

Delatte, H., P. David, M. Granier, J. M. Lett, R. Goldbach, M. Peterschmitt, et al. 2006. Microsatellites reveal extensive geographical, ecological and genetic contacts between invasive and indigenous whitefly biotypes in an insular environment. Genet. Res. 87:109–124.

Dinsdale, A., L. Cook, C. Riginos, Y. M. Buckley, and P. De Barro. 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, and S. Morin. 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.

Frohlich, D. R., I. Torres-Jerez, I. D. Bedford, P. G. Markham, and J. K. Brown. 1999. A phylogeographical analysis of the *Bemisia tabaci* species complex based on mitochondrial DNA markers. Mol. Ecol. 8:1683–1691.

Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Mol. Ecol. Notes 5:184–186.

Goudet, J., M. Raymond, T. deMeeus, and F. Rousset. 1996. Testing differentiation in diploid populations. Genetics 144:1933–1940.

Gueguen, G., F. Vavre, O. Gnankine, M. Peterschmitt, D. Charif, E. Chiel, et al. 2010. Endosymbiont metacommunities, mtDNA diversity and the evolution of the *Bemisia tabaci* (Hemiptera: Aleyrodidae) species complex. Mol. Ecol. 19:4365–4376.

Gunning, R. V., F. J. Byrne, and A. L. Devonshire. 1997. Electrophoretic analysis of non-B and B-biotype *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) in Australia. Aust. J. Entomol. 36:245–249.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/ NT. Nucleic Acids. Symp. Ser. 41:95–98.

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.

Li, X. X., S. J. Li, X. Xue, M. Z. Ahmed, S. X. Ren, A. G. S. Cuthbertson, et al. 2012. The effects of space dimension and temperature on the cross mating of three cryptic species of the *Bemisia tabaci* complex in China. J. Integr. Agric. 11:197–205.

Liu, S. S., P. J. De Barro, J. Xu, J. B. Luan, L. S. Zang, Y. M. Ruan, et al. 2007. Asymmetric mating interactions drive widespread invasion and displacement in a whitefly. Science 318:1769–1772.

Liu, S. S., J. Colvin, and P. J. De Barro. 2012. Species concepts as applied to the whitefly *Bemisia tabaci* systematics: how many species are there? J. Integr. Agric. 11:176–186.

Mallet, J. 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol. 20:229–237.

Manel, S., O. E. Gaggiotti, and R. S. Waples. 2005. Assignment methods: matching biological questions techniques with appropriate. Trends Ecol. Evol. 20:136–142.

Maruthi, M. N., J. Colvin, R. M. Thwaites, G. K. Banks, G. Gibson, and S. E. Seal. 2004. Reproductive incompatibility and cytochrome oxidase I gene sequence variability amongst host-adapted and geographically separate *Bemisia tabaci* populations (Hemiptera: Aleyrodidae). Syst. Entomol. 29:560–568.

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

Moya, A., P. Guirao, D. Cifuentes, F. Beitia, and J. L. Cenis. 2001. Genetic diversity of Iberian populations of *Bemisia tabaci* (Hemiptera: Aleyrodidae) based on random amplified polymorphic DNA-polymerase chain reaction. Mol. Ecol. 10:891–897.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl Acad. Sci. USA 70:3321–3323.

Omondi, B. A., P. Sseruwagi, D. Obeng-Ofori, E. Y. Danquah, and R. A. Kyerematen. 2005. Mating interactions between okra and cassava biotypes of *Bemisia tabaci* (Homoptera: Aleyrodidae) on eggplant. Int. J. Trop. Insect Sci. 25:159–167.

Parrella, G., L. Scassillo, and M. Giorgini. 2012. Evidence for a new genetic variant in the *Bemisia tabaci* species complex and the prevalence of the biotype Q in southern Italy. J. Pest. Sci. 85:227–238.

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.

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

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

Ronda, M., A. Adan, D. Cifuentes, J. L. Cenis, and F. Beitia. 1999. Laboratory evidence of interbreeding between biotypes of Bemisia tabaci (Homoptera, Aleyrodidae) present in Spain. P. 83 *in* VII International Plant Virus Epidemiology Symposium. Almeria, Spain.

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

Saleh, D., A. Laarif, C. Clouet, and N. Gauthier. 2012. Spatial and host-plant partitioning between coexisting *Bemisia tabaci* cryptic species in Tunisia. Popul. Ecol. 54:261–274.

Servedio, M. R. 2001. Beyond reinforcement: the evolution of premating isolation by direct selection on preferences and postmating, prezygotic incompatibilities. Evolution 55:1909–1920.

Simon, B., J. L. Cenis, S. Demichelis, C. Rapisarda, P. Caciagli, and D. Bosco. 2003. Survey of *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotypes in Italy with the description of a new biotype (T) from *Euphorbia characias*. Bull. Entomol. Res. 93:259–264.

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

Storey, J. D., and R. Tibshirani. 2003. Statistical significance for genomewide studies. Proc. Natl Acad. Sci. USA 100:9440–9445.

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 behavioural evidence. Bull. Entomol. Res. 101:211–220.

Tahiri, A., A. Sekkat, A. Bennani, M. Granier, G. Delvare, and M. Peterschmitt. 2006. Distribution of tomato-infecting begomoviruses and *Bemisia tabaci* biotypes in Morocco. Ann. Appl. Biol. 149:175–186.

Tsagkarakou, A., and N. Roditakis. 2003. Isolation and characterization of microsatellite loci in *Bemisia tabaci* (Hemiptera: Aleyrodidae). Mol. Ecol. Notes 3:196–198.

Tsagkarakou, A., L. Mouton, J. B. Kristoffersen, E. Dokianakis, M. Grispou, and K. Bourtzis. 2012. Population genetic structure and secondary endosymbionts of Q *Bemisia tabaci*

(Hemiptera: Aleyrodidae) from Greece. Bull. Entomol. Res. 102:353–365.

Vähä, J. P., and C. R. Primmer. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Mol. Ecol. 15:63–72.

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.

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

Xu, J., P. J. De Barro, and S. S. Liu. 2010. Reproductive incompatibility among genetic groups of *Bemisia tabaci* supports the proposition that the whitefly is a cryptic species complex. Bull. Entomol. Res. 100:359–366.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Inference of the number of cluster(s) (K) that best explains the population structure of the *Bemisia tabaci* samples. For each tested value of K, we report the mean and standard deviation (SE) of the logarithm of the posterior probability distribution of the allele frequencies (X) given the number of clusters [Ln Pr(X|K)] and the ΔK index values (Evanno et al. 2005), both calculated over 20 simulation runs.

Table S2. Differentiation between populations.

Table S3. Hierarchical analysis of molecular variance.

Table S4. Allele composition of the three *Bemisia tabaci* individuals that were not assigned (assignment coefficient <0.8) to the two genetic groups as defined by STRUCTURE analysis. For each locus, the sizes of the alleles detected for each individual are indicated. If the alleles were detected only in one species and not in the other one (private allele), they are indicated in bold and the line below specifies if they were private to Med (**Md**) or to MEAM1 (**ME**). Alleles that are relatively more frequent in one species compared to the other one are in plain text and the line below specifies if it is more frequent in Med (Md) or MEAM1 (ME). Alleles similarly frequent in both species are indicated as "–" in the line below. "Null" stands for "null allele" which means that no amplification product was detected.