Genet. Res., Camb. (2006), **87**, *pp.* 109–124. © 2006 Cambridge University Press
doi:10.1017/S0016672306008135 Printed in the United Kingdom

Microsatellites reveal extensive geographical, ecological and genetic contacts between invasive and indigenous whitefly biotypes in an insular environment

H. DELATTE[1,4]*, P. DAVID[2], M. GRANIER[3], J. M. LETT[1], R. GOLDBACH[4], M. PETERSCHMITT[3] AND B. REYNAUD[1]

[1] *CIRAD-UMR PVBMT-Université de la Réunion, Pôle de Protection des Plantes Ligne Paradis, 97410 Saint Pierre, La Réunion, France*
[2] *CEFE-CNRS, UMR 5175, 1919 Route de Mende, 34293 Montpellier Cedex 05, France*
[3] *CIRAD, UMR BGPI, TA 41/K, 34398 Montpellier, Cedex 5, France*
[4] *Wageningen University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands*

(*Received 13 January 2006 and in revised form 5 February 2006*)

Summary

Human-mediated bioinvasions provide the opportunity to study the early stages of contact between formerly allopatric, divergent populations of a species. However, when invasive and resident populations are morphologically similar, it may be very difficult to assess their distribution in the field, as well as the extent of ecological overlap and genetic exchanges between invasive and resident populations. We here illustrate the use of data obtained from a set of eight microsatellite markers together with Bayesian clustering methods to document invasions in a group of major tropical pests, *Bemisia tabaci*, which comprises several morphologically indistinguishable biotypes with different agronomic impacts. We focus on the island of La Réunion, where an invasive biotype (B) has recently been introduced and now interacts with the resident biotype (Ms). The temporal and spatial distribution, host-plant range and genetic structure of both biotypes are investigated. We showed (i) that, without prior information, clustering methods separate two groups of individuals that can safely be identified as the B and Ms biotypes; (ii) that the B biotype has invaded all regions of the island, and showed no signs of genetic founder effect relative to the Ms biotype; (iii) that the B and Ms biotypes coexist in sympatry throughout most of their geographical ranges, although they tend to segregate into different host plants; and finally (iv) that asymmetrical and locus-specific introgression occurs between the two biotypes when they are in syntopy.

1. Introduction

The genetic and ecological consequences of a new contact established between divergent populations of a species have usually been investigated through the study of naturally occurring hybrid zones (Barton & Hewitt, 1985). Most of these zones behave as stationary systems at the timescale of the study, and usually do not provide the opportunity to observe the early stages of contacts between divergent populations. However, many important – if temporary – processes only occur immediately after the first contact: for example, one population may displace the other before a stable hybrid zone occurs; the two populations may or may not coexist in sympatry (or syntopy) before segregating into different geographical (respectively, ecological) settings; or favourable alleles from one population may rapidly invade the other. During the twentieth century, human trade has promoted long-range introduction events well beyond the natural dispersion range of species, bringing into contact species or genotypes that never used to coexist (Mooney & Cleland, 2001). This provides the opportunity to study the early stages of a contact between previously allopatric divergent populations and examine what kind of geographical, ecological and genetical relationships are established between the two in this context. We here provide an example of such a study in an insect taxon.

* Corresponding author. CIRAD-UMR PVBMT Pôle de Protection des Plantes Ligne Paradis, 97410 Saint Pierre, La Réunion, France. Tel: +262 (262) 499235. Fax: +262 (262) 499293. e-mail: delatte@cirad.fr

Bemisia tabaci (Gennadius) (Hemiptera: Aleyrodidae) is a polyphagous whitefly species distributed throughout tropical and subtropical regions worldwide (Oliveira *et al*., 2001). In the past two decades, *B. tabaci* has become a major problem both as a pest and as vector of plant viruses, especially begomoviruses (Markham *et al*., 1995; Jones, 2003; Varma & Malathi, 2003). However, it has been well known since the 1950s that this species covers several distinct entities named 'biotypes' (Bird & Maramorosch, 1978). Biotypes cannot be distinguished morphologically, but they may exhibit measurably different life-history traits, host ranges, levels of insecticide resistance and abilities to carry plant viruses (Brown *et al*., 1995*a*; Horowitz *et al*., 2005; Byrne *et al*., 2003), and correspond to different mitochondrial DNA (mtDNA) clades (Frohlich *et al*., 1999). More than 20 biotypes have been described (Perring *et al*., 2001), most of which are specific to one region; however, recent invasions have brought some pairs into sympatry. The most famous example is biotype B, which rapidly displaced the indigenous biotype A in the 1980s (Brown *et al*., 1995*a*) from most of its range in the United States (Perring *et al*., 1991; Gill, 1992; Bellows *et al*., 1994). Other introductions of the B biotype have had different outcomes: in Spain, the indigenous Q biotype coexists with B but has remained predominant in the field (Pascual & Callejas, 2004). However, whether the B biotype displaced or simply coexists with indigenous ones, new begomovirus epidemics have been reported (Polston & Anderson, 1997; Ribeiro *et al*., 2003). Many questions remain disputed with regard to the exact nature of biotypes, especially their degree of genetic isolation and whether interbreeding occurs when two of them are brought into sympatry. The lack of morphological differentiation also poses a major problem in detecting the introductions and analysing ecological interactions between biotypes.

Since 1997 a severe outbreak of tomato yellow leaf curl virus (TYLCV) (Peterschmitt *et al*., 1999) has been observed in La Réunion, an island in the Indian Ocean 700 km east of Madagascar. The viral epidemic was associated with a rise in its vector population *B. tabaci*. Prior to 1997 no *B. tabaci* infestations of similar importance had ever been reported in La Réunion; however, *B. tabaci* had been reported in La Réunion since 1938 (Bourriquet, 1938). A preliminary study on *B. tabaci* by RAPD-PCR and sequencing of a part of the cytochrome oxidase I gene (COI) has revealed the presence of the B biotype together with an indigenous biotype named Ms (Delatte *et al*., 2005). The introduction of the B biotype into the insular environment of La Réunion provides a great opportunity to analyse a colonization process from the beginning. However, because the biotypes cannot be distinguished morphologically, it has always been very difficult to follow the progress of biotype B invasions in the field. We have developed microsatellite markers to this end. Microsatellites were chosen as they are co-dominant and highly polymorphic markers, distributed throughout the nuclear genome and generally neutral unless linked to loci under strong selection. These characteristics make them good candidates for the study of population differentiation (Tautz, 1989), the colonization process and genetic changes in the populations.

We developed eight new polymorphic microsatellite markers in order (i) to test whether microsatellites allow the recognition of two distinct subpopulations, corresponding to biotypes Ms and B, consistent with COI sequences in La Réunion populations, (ii) to establish the current frequency of biotype B in La Réunion, and to ask to what extent the two biotypes overlap geographically and/or ecologically (host plant), using an extensive sample of 662 individuals, (iii) to examine whether the recent introduction of biotype B is associated with a reduction in genetic diversity, and (iv) to detect possible gene flow between the invasive and the resident biotypes.

2. Materials and methods

(i) *Whitefly sampling*

Three samplings of adult whiteflies were carried out (S1, February–March 2001; S2, September–October 2001; S3, February–March 2002), totalling 662 individuals on 10 different host plants (see below) at 19 different sites distributed over the whole island (Fig. 1). Host plants included annual poinsettia (*Euphorbia heterophylla*), poinsettia (*Euphorbia pulcherrima*), tomato (*Lycopersicon esculentum*), eggplant (*Solanum melongena*), cabbage (*Brassica oleracea*), ipomea (*Ipomea maxima*), watermelon (*Cucurbita* sp.), bean (*Vigna* sp.), lantana (*Lantana camara*), apple of Peru (*Nicandra physaloides*) and cassava (*Manihot esculenta*). The south and east regions are the wettest part of La Réunion, with the east mostly an area of sugar cane. The west and north are drier and contain most of the vegetable plantations.

(ii) *Microsatellite library*

An enriched library was built using a protocol modified from Billote *et al*. (1999). Total genomic DNA was extracted from 25 mg of pooled *B. tabaci* individuals from a laboratory strain of the B biotype using a DNA plant tissue kit (Qiagen, Courtaboeuf, France), and digested with *Sau* 3AI enzyme. The cohesive ends of DNA fragments were polished using Kleenow polymerase (Gibco BRL). Fragments were screened for AC repeats using a biotin-labelled $(TG)_8$ probe.

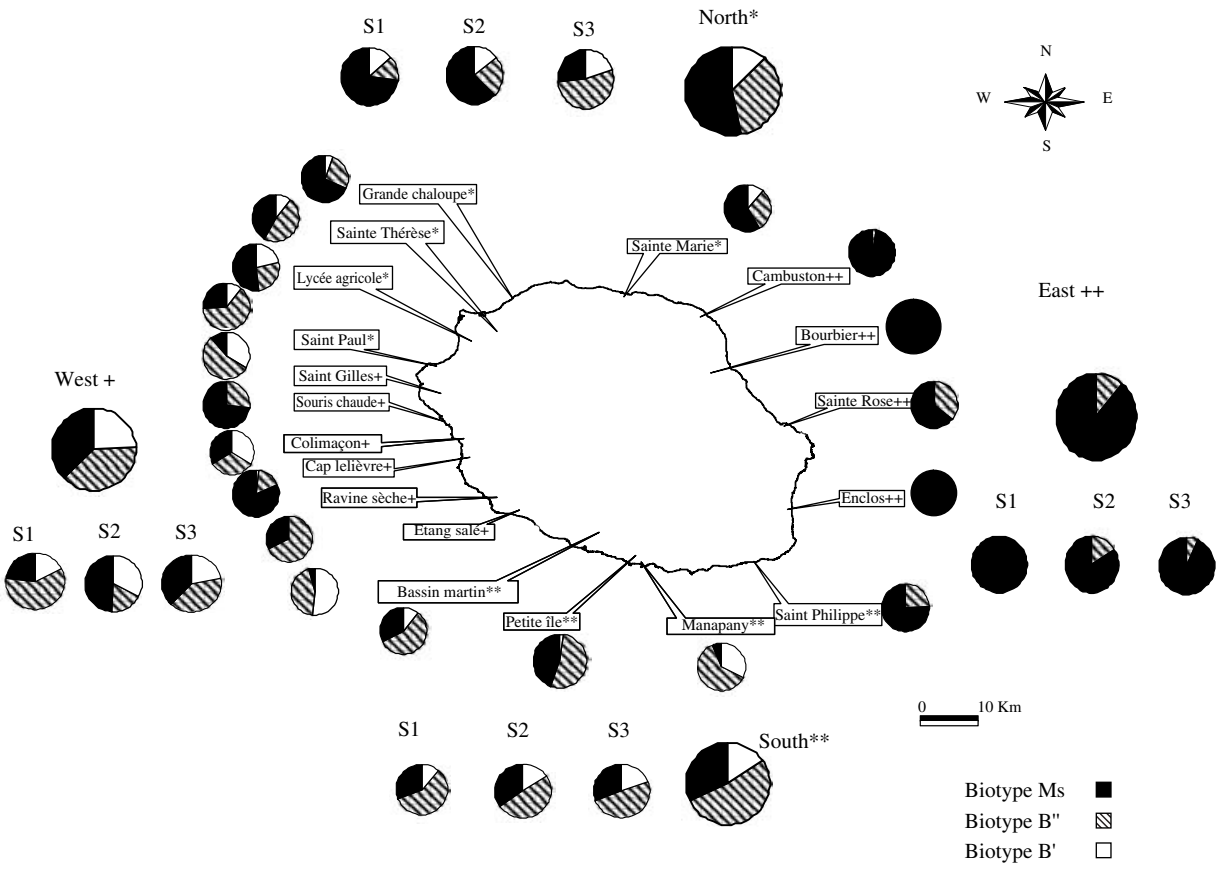

Fig. 1. Map of La Réunion with whitefly biotype B and Ms global repartition in absolute numbers per sector and sampling period (S1, S2, S3). The different sites are represented individually with years grouped, for biotype Ms and groups B′ and B″. The sampled sites are represented, and the symbols refer to the sectors they belong to.

Thirty-four clones containing microsatellite sequences were recovered, and 16 were tested with DNA from *B. tabaci* (biotypes B, Q and Ms) and *Trialeurodes vaporariorum*. The eight most polymorphic ones, amplifying for both biotypes B and Ms, were kept for this study (Table 1).

(iii) *DNA extraction, amplification and detection of microsatellites*

Whiteflies have haploid males and diploid females. Each field-captured whitefly was sexed before DNA extraction (Delatte *et al*., 2005). One primer of each locus was fluorescently end-labelled (Table 1), and a 'pigtail' was added at the 5′ end (Brownstein *et al*., 1996). PCR was conducted in a 20 μl volume with three different PCR mixes: (i) for primers P7,P11 and P53, 10× PCR buffer, 0·15 mM dNTP, 0·25 pmol per primer 1·2 mM $MgCl_2$, 0·2 unit *Taq* polymerase (Goldstar, Eurogentec, Seraing, Belgium), 2 μl DNA; (ii) for primer P59, the concentration of $MgCl_2$ was changed to 2 mM; (iii) for primers P5, P32, P41 and P62, the concentrations of $MgCl_2$ and dNTP were changed to 3 mM and 0·2 mM respectively. Two PCR programmes were used: for primers P5, P32, P41 and P62, a single soak at 94 °C for 5 min, followed by 5 cycles at 94 °C for 30 s, 65 °C for 1 min, 72 °C for 1 min, then 30 cycles at 94 °C for 30 s, 64·5 °C (−0·5 °C/cycle) for 1 min, 72 °C for 1 min, then 5 min at 72 °C; for other primers, 94 °C for 5 min followed by 40 cycles at 94 °C for 30 s, 50 °C for 30 s, 72 °C for 1 min, and then 10 min at 72 °C. Amplification products (2 μl) were mixed with 10 μl of ultra-pure formamide and 0·3 μl of size marker (GeneScan ROX, Foster City, CA, USA), and loaded onto an ABI Prism 700 automated sequencer. Allele sizes were calculated using GeneScan 3.1 and Genotyper 2.5, available from Plant Research International (PRI, Wageningen, The Netherlands).

The primary data from the study can be accessed at http://journals.cambridge.org/.

(iv) *Data analyses*

Only females were used for the genetic diversity analyses due to their diploid state. To differentiate the B from the Ms biotype, we used the software Structure v2.1 of Pritchard *et al*. (2000). This software differentiates mixed populations on the basis of allele frequency at each locus. The number of individuals

Table 1. *Microsatellite sequences, motives, predicted sizes and accession numbers registered in EMBL*

Name/Primer sequence (5′–3′) and fluorochrome	Repeat	Predicted size	Accession no.
Group 1			
53a/TTCGTAACGTCTTTAAATTTTTGC-FAM 53b/**TTG**TGGAGCATATAGCCTTTTTGG[a]	20(GT)	148 bp	AJ866707
7a/AGGGTGTCAGGTCAGGTAGC-HEX 7b/**TTT G** CGTAATGGTAACATGTTTAGAAAA[a]	8(GT)	176 bp	AJ866710
5a/ATTAGCCTTGCTTGGGTCCT-NED 5b/**TTT G** CAAAAACAAAAGCATGTGTCAAA[a]	8(GT)	207 bp	AJ866711
62a/CTTCCTTAGCACGGCAGAAT-FAM 62b/**TTT G** GCGCAATTTTTAGCGTCTGT[a]	8(GT)	209 bp	AJ866714
Group 2			
41a/AATCTTAGTGCTAAAGTTTTCTT-FAM 41b/**TT G** TGTTAGGATGATAGGCTTGGA[a]	7(AC)	165 bp	AJ866709
11a/CCAGAAAAGTGGACTTAAGA-HEX 11b/**TTT G** GATCTGGGTGTTTTCTTCTA[a]	9(GT)	180 bp	AJ866708
32a/GCTCAAAATTGTTGGCTCTGA-NED 32b/**TTTG**AGCTCCGCCCTTAAATTGTT[a]	9(CA)	199 bp	AJ866713
59a/CGGCGTTTCTCGTTTTCTT-FAM 59b/**TTTG**CCAACTGAAGCACATCAATCA[a]	44(T) 18(G)	208 bp	AJ866712

[a] Parts of the sequences in bold represent the 'pigtail' added to the original sequence.

and allele frequencies in each cluster, as well as individual assignments, are outputs of the analysis. The number of clusters, K, is decided by the user. Although our initial guess was $K=2$, we also tried higher numbers (up to 10) and used the method of Evanno *et al.* (2005) to determine the optimal K. This turned out to be $K=2$. The software was run twice, with and without the option admixture, which allows for some mixed ancestry within individuals. To use Structure, Hardy–Weinberg and linkage equilibrium are assumed within each group. Both hypotheses were tested *a posteriori* on each cluster using exact tests implemented in Genepop 3.3 (Raymond & Rousset, 1995). In addition, the identity of the clusters was checked by including in the analysis two types of controls with known biotype identity: laboratory cultures and COI-sequenced wild individuals. Laboratory controls were two Ms and three B insects from our laboratory strains. The mitochondrial cytochrome oxidase I (COI) gene, known as diagnostic for biotypes, was amplified and sequenced for 35 of the 662 wild individuals typed for microsatellites, using primers C1-J2195 and L2-N3014 (Frohlich *et al.*, 1999). A neighbour-joining tree of the sequences (370 nucleotides) was constructed from the maximum likelihood distance matrix using DNAMAN (Saitou & Nei, 1987). Sequences from the B, Q and A biotypes of *B. tabaci*, and *Trialeurodes vaporariorum*, downloaded from EMBL were included in the tree.

A correspondence analysis (COA) was performed using Genetix 4.01 (Belkhir *et al.*, 1996–2004) to visualize the major axes of genetic variation within the sample, and the position of the groups defined by the Structure software along these axes.

Within each of the two clusters inferred by Structure (i.e. putative biotypes), pairwise F_{st} among sites were computed using Weir & Cockerham's (1984) estimator θ, and isolation by distance, the correlation between genetic and geographical distance, was tested by the regression of $F_{st}/1-F_{st}$ on the logarithm of Euclidian geographical distances (Rousset, 1997).

Genetic diversity within each cluster was quantified by the number of alleles per locus, the observed heterozygosity (H_o) and expected heterozygosity (H_e). For the two clusters separately, Weir & Cockerham (1984) estimates of F_{is} within localities and of F_{st} among localities were calculated using Genepop 3.3. The null hypotheses of Hardy–Weinberg frequencies within populations, and lack of population structure, were tested using exact tests using Genepop 3.3. A hierarchical analysis of molecular variance (AMOVA) was obtained with Arlequin software (Excoffier *et al.*, 1992), partitioning the genetic variance into three components: (i) within-site within-cluster, (ii) among-sites within-cluster, and (iii) between-cluster.

Using raw data, large heterozygote deficiencies were present at all loci (except P62 for biotype Ms) and both biotypes (Table 2B). In order to obtain robust results, we repeated all analyses using pooled alleles that grouped together alleles of similar size (see Section 3 for details on the pooling procedure). Results of the analysis (with the exception of heterozygote deficiencies) were similar to those of the non-pooled dataset. Unless otherwise stated, we present only the more robust results obtained after pooling alleles (the raw dataset is available on the *Genetical Research* website).

Table 2. *Genetic diversity of the B and Ms biotypes (as delimited by Structure) within and among sites (all periods pooled)*

	After (A)				Before (B)			
Locus/Biotype	H_o	H_e	F_{is}	F_{st}	H_o	H_e	F_{is}	F_{st}
Ms								
P53	0·46	0·69	0·29***	0·05***	0·58	0·93	0·36***	0·04***
P62	0·50	0·50	−0·21*	0·02*	0·66	0·67	−0·02***	0·03***
P7	0·42	0·40	−0·06***	0·04***	0·41	0·72	0·37***	0·07***
P5	0·31	0·44	0·27***	0·06***	0·43	0·67	0·33***	0·05***
P41	0·04	0·06	0·25***	0·28***	0·07	0·43	0·83***	0·10***
P59	0·33	0·50	0·33***	0·10***	0·36	0·86	0·56***	0·05***
P11	0·12	0·13	0·05	0·01	0·16	0·62	0·72***	0·06***
P32	0·28	0·44	0·35***	0·10***	0·36	0·73	0·50***	0·05***
Average	*0·31*	*0·39*	*0·17****	*0·06****	*0·38*	*0·70*	*0·43****	*0·05****
B								
P53	0·66	0·70	0·04***	0·03***	0·71	0·91	0·19***	0·04***
P62	0·24	0·41	0·44***	0·16***	0·32	0·74	0·52***	0·11***
P7	0·19	0·20	−0·14*	0·04***	0·35	0·78	0·52***	0·11***
P5	0·10	0·36	0·57***	0·17***	0·23	0·63	0·57***	0·14***
P41	0·21	0·20	−0·07	0·03	0·28	0·58	0·45***	0·10***
P59	0·54	0·68	0·26***	0·05***	0·59	0·94	0·36***	0·04***
P11	0·01	0·03	0·24	0·01	0·11	0·75	0·85***	0·08***
P32	0·57	0·60	−0·02	0·01	0·62	0·86	0·26***	0·02***
Average	*0·32*	*0·40*	*0·18****	*0·07****	*0·40*	*0·77*	*0·45****	*0·08****

Results are presented after (A) and before (B) pooling alleles (see text for details on allele pooling). The observed heterozygosity (H_o), the expected heterozygosity (H_e) and the fixation indices (F_{is} and F_{st}) of Weir & Cockerham (1984) were given by the software Genepop. Hardy–Weinberg tests are indicated together with F_{is} values, and tests of homogeneity among populations together with F_{st} values.
*$P<0·05$; ** $P<0·01$; *** $P<0·001$. Non-significant P values are not indicated. The P values were combined over all loci using Fisher's method.

Once all individuals were classified as putative B or Ms, we were in a position to analyse the variation in relative frequency of the two biotypes in space and time. This was performed using an analysis of deviance model for binary data ($y=1$ for B, $y=0$ for Ms), incorporating the effects of (i) sampling period (S1, S2, S3), (ii) sampling location described as quadrant (east, west, north, south) and site within quadrant (four to six sites) and (iii) host plant described as type (cultivated vs weed) and species within type. The significance of particular effects was tested using model simplification (Crawley, 1993). Accordingly, the change in deviance due to the removal of a given term was compared with a chi-square distribution. The residual deviance consistently showed hints of under-dispersion, rather than over-dispersion, making the chi-square tests relatively conservative in this case.

3. Results

(i) *Microsatellite variability and allele size classes*

From the total 662 whiteflies collected and tested, 34 were unsuccessfully amplified and 61 which were amplified for fewer than six of the loci were discarded. The eight loci had 9 to 33 alleles each. Many alleles scored were of less than two nucleotides, possibly affected by rounding errors, and a first genetic analysis revealed often strong heterozygote deficiencies, especially for allele pairs with a small size difference (2 bp or less), suggesting scoring artefacts. To obtain conservative estimates of population structure and *F*-statistics, we decided to pool our alleles into allele-length intervals. To the extent that microsatellites can be considered neutral, population processes such as non-random mating, geographic structure, Wahlund effects and introgression must act on pooled alleles the same way as on raw alleles, the former being simply slightly more homoplasic than the latter. On the other hand, scoring artefacts may be attenuated by pooling alleles of similar size. To this end, allele frequencies were calculated for each locus and plots of frequency versus allele size were generated. Each plot typically displayed several peaks spanning several base pairs. Our pooling procedure was such that each peak corresponds to a synthetic allele, the limit classes between two peaks being conservatively pooled with the highest peak. After transformation the eight microsatellite loci

had four to 11 synthetic alleles per locus. The results obtained with synthetic alleles are consistent with those based on raw alleles (except of course for within-population heterozygote deficiencies and genetic diversities: see below), though more robust to scoring artefacts. Heterozygote deficiencies largely decreased upon pooling alleles, even in loci that remained very polymorphic, suggesting that f values in the original dataset were largely artefactual (Table 2, part B). In contrast to scoring errors, inbreeding, null alleles or Wahlund effects are expected to generate heterozygote deficiencies that remain unaffected by our pooling procedure, not being specific to pairs of alleles of similar size.

(ii) *Biotype differentiation and phylogenetic analyses*

Five hundred and sixty-seven genotypes (females and males), together with three laboratory whitefly controls of the B strain and two laboratory whitefly controls from the Ms strain, were given to Structure. The software identified two clusters (cluster 1 of 308 individuals; cluster 2 of 259 individuals). Individual assignation probabilities were all close to either 0 or 1, with no whitefly in the [0·20; 0·80] interval. Cluster 1 contained the B biotype laboratory controls, while cluster 2 contained the Ms laboratory controls (Fig. 2).

Of 35 insects tested with COI, 15 were classified by Structure as Ms. All of them had 100% nucleotide identity (only two were registered in EMBL: AJ877263 and AJ877264). The remaining 20 insects were classified as B by Structure and again had 100% identity (four were registered in EMBL: AJ877259, AJ877260, AJ877261, AJ877261). In the phylogenetic tree these two groups belonged to different well-supported clades (Fig. 2*A*). The first clade contains only sequences from the Indian Ocean region (including La Réunion), while the second contains sequences from known invasive populations of the B biotype from Israel or the United States. Thus, mitochondrial phylogeny is so far 100% consistent with the B and Ms classification obtained by Structure software. Hereafter, for simplicity, we use the terms 'biotype B' and 'biotype Ms' to refer to the genetic clusters provided by Structure.

(iii) *Genetic variation within and between the B and Ms biotypes*

Genetic diversity was similar between the two biotypes (Table 2), whether pooled alleles were used or not. Synthetic and raw allele frequencies are given in Appendix A. For five loci in biotype Ms (P53, P5, P41, P59, P32) and three in biotype B (P62, P5, P59), the observed heterozygosity (H_o) was lower than the expected heterozygosity (H_e). The average within-population heterozygote deficiencies were similar for the B and Ms biotypes (Table 2A). Both biotypes also show genetic substructure among sites, as attested by significant F_{st} (Table 2A). The F_{st} values are comparable between biotypes. Cases of linkage disequilibrium were detected between some loci (see Appendix B).

No significant correlation was observed between geographic and genetic ($F_{st}/(1-F_{st})$) distance within either biotype (Mantel tests; biotype B, $r=-0{\cdot}019$; biotype Ms, $r=-0{\cdot}111$; both NS).

The AMOVA showed that most of the variation was distributed within population (66%) and between biotypes (31%) (Table 3). The genetic differences between B and Ms account for much more genetic variance (31%; $F_{ct}=0{\cdot}32$, $P<0{\cdot}001$) than those among populations within biotypes (3%; $F_{ct}=0{\cdot}04$, $P<0{\cdot}001$).

The genetic COA allows a graphical representation of the major axes of genetic differentiation within the sample. Only diploid females were included in the COA. Although this analysis did not make use of the groupings identified by the Structure software, there is a striking correspondence between Structure clusters and the position of individuals along the first two axes of the COA, which together represent 12·03% of total genetic variation (see Fig. 3*A*). Despite a large dispersion within each biotype, individuals from the two biotypes identified by Structure form well-separated clusters in the first factorial plane, especially along axis 1. However, in the factorial plane the B group seems to contain two subgroups: one close to (B″), the other far from (B′) the Ms group (Fig. 3). Indeed, the distribution of the first principal coordinates (F1) within the B biotype is strongly bimodal, with one mode centred on 0·2 (subgroup B′, 91 individuals) and the other on 0·8 (subgroup B″, 217 individuals) (see Fig. 3). This separation was also found by Structure, using the option admixture. Indeed, Structure classified all B′ individuals as genotypes with mixed ancestry (i.e. approximately half of the genes from one biotype, half from the other) and B″ as 'pure' genotypes (Fig. 3*C*). In order to investigate in greater detail the genetic makeup of the B′ and B″ groups, we first defined as 'Ms alleles' all alleles with higher frequency in the Ms than in the B biotype (including both B′ and B″), and as 'B alleles' all other alleles, then evaluated for each individual the relative proportions of each category over all scored loci. The results are presented in Fig. 3*B*, where the histograms make it apparent that the subgroups B′ and B″ differ with respect to the distribution of this proportion (non-parametric Kruskal–Wallis ANOVA, $\chi^2=114{\cdot}0$, $P<0{\cdot}001$). This confirms that the B″ group is genetically intermediate between B′ and Ms, as suggested by its position on the factorial plane. However, not all loci participate

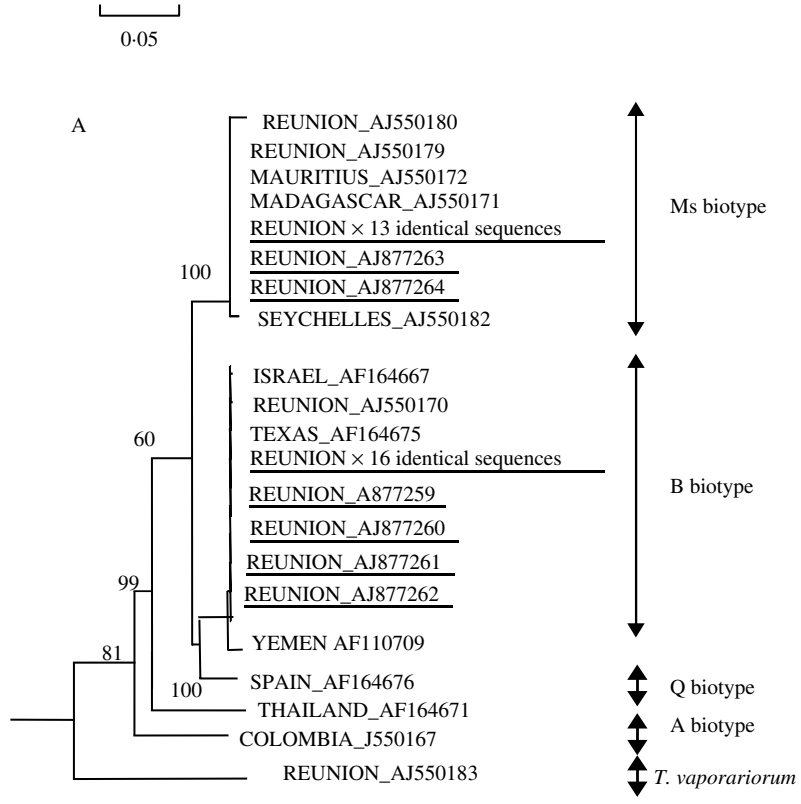

Fig. 2. Identification of biotypes in the sample by molecular data. (*A*) Rooted neighbour-joining tree showing the genetic distance among 370 nucleotide cytochrome oxidase I fragments of *Bemisia tabaci*, either sequenced in this study or selected from GenBank. Sequences are identified by their accession number and the geographic origin of the tested whitefly. The sequences generated in this study are underlined and followed by their code number. One *Trialeurodes vaporariorum* sequence was added to this comparison, as an outgroup. The scale measures the maximum likelihood distance between sequences. Numbers associated with nodes represent the percentage of 2000 bootstrap iterations supporting the nodes; only percentages $\geqslant 60\,\%$ are indicated. (*B*) Summary plot of estimates of Q (the estimated membership coefficient for each individual in each cluster) given by the software Structure v2.1. Each of the 567 individuals is represented by a single vertical line broken into K populations ($K=2$ in our case), with length proportional to each of the K inferred. The numbers 1 and 2 correspond to B and Ms biotypes, respectively. The software was run with the option no admixture.

equally in this effect. At most loci the frequency of the Ms synthetic allele in the B″ group is near (though slightly above) that in the B′ group (Appendix A). There are two exceptions: the first is locus P62, at which the frequency of the Ms allele in B″ (0·898) is very high, equal to that in the Ms biotype (0·900), and very different from the B′ group (0·069). The second is locus P5, which shows a similar pattern (frequencies 0·953, 0·815 and 0·034 in the Ms, B″ and B′ groups respectively) (Appendix A).

Other subgroups can be defined within the B cluster, on the basis of coexistence or not with the Ms biotype in the field. Within the B cluster, some individuals (hereafter 'syntopic individuals') belong to samples in which Ms individuals were also captured (i.e. at the same site, on the same plant and at

Table 3. *Analysis of molecular variance* (*V*)*, and F-statistics of genetic differentiation between B and Ms biotypes and among sampling sites of* Bemisia tabaci (*populations*) *computed by the method of Excoffier* et al. (*1992*)

Source of variation								
Among biotypes			Among sites within biotypes			Within sites within biotypes		
V	%	F_{ct}	V	%	F_{sc}	V	%	F_{st}
0·60	31·37	0·32***	0·05	2·54	0·04***	1·25	65·74	0·34***

The *P* value is of $P<10^{-5}$ for the three *F*-statistics.

the same time), while others ('non-syntopic individuals') belong to samples in which no Ms individuals were caught (Fig. 3*A*). The syntopic B and non-syntopic B subsamples were genetically differentiated ($F_{st}=0{\cdot}07$, $P<0{\cdot}001$ over all loci). Specifically, although syntopic and non-syntopic individuals were present in the two subclusters B′ and B″, non-syntopic individuals were more concentrated in the B′ cluster (61 B′ and 48 B″) and syntopic individuals in the B″ cluster (29 B′ and 123 B″; Kruskal–Wallis test on the first principal coordinate, syntopic versus non-syntopic: $\chi^2=33{\cdot}05$, $P<0{\cdot}001$, see Fig. 3*A*). On the other hand, no symmetrical difference was observed between syntopic Ms and non-syntopic Ms individuals (Fig. 3*A*, Kruskal–Wallis ANOVA, $\chi^2=0{\cdot}006$, NS).

(iv) *Frequency of the B biotype as a function of spatial, temporal and host-plant variables*

Once non-significant interactions were simplified out, the final statistical model included the effects of host plant type (cultivated or weed) and species within type; sampling season; geographical sector and population within the sector; as well as a significant interaction between sector and sampling season (Table 4). Biotype B tended to be concentrated in crops while biotype Ms was predominant on weeds (Fig. 4, Table 4). Over the total amount of *B. tabaci* biotype B, 68 % of individuals identified as biotype B, but only 26 % identified as biotype Ms, were collected on vegetable crops (respectively 206 insects for B and 67 for Ms). The B population was found in all the sampled host plants whereas the Ms population was absent from cotton and bean (Fig. 4). With regard to host-plant species, eggplant and cabbage favoured the B biotype, while annual poinsettia and lantana favoured the Ms biotype (Table 4, Fig. 4).

Biotype B was present at almost every sampling site and at every sampling period except in the east part during the last samplings. Biotype B was dominant throughout the island except in the eastern region (Fig. 1). Among the two subgroups B″ and B′ within biotype B, the former is more widespread; it is present at all sites occupied by biotype B, and usually predominant. Only two populations of the west coast (Etang Salé and Colimaçon) contain as many or more B′ than B″ individuals.

The deviance analysis showed that there were variations in the proportion of biotype B relative to Ms at the temporal, spatial and ecological levels (Table 4). Globally the east and north sectors had more Ms biotypes than the west and south. An increase in the proportion of Ms was observed in the east over time, whereas in the three other sectors a decrease in proportion over time was observed. This decrease is clearly seen in the northern sector, with a decrease of Ms in favour of the B″ group.

4. Discussion

(i) *Posterior biotype discrimination on the basis of microsatellite data*

Identification of *B. tabaci* biotypes and assessment of their relative proportions in the field is of primary importance in agronomic terms, because biotypes differ greatly in their ecological requirements as well as agronomic impact (reviewed in: Perring, 2001; Oliveira *et al*., 2001). Many invasions by biotypes of *B. tabaci* may remain cryptic because established biotypes, morphologically indistinguishable from invasive ones, are already present (Bird & Maramorosch, 1978). In this respect, *B. tabaci* is not an isolated example (see e.g. Genner *et al*., 2004) and there is a growing need for reliable genetic screens to evaluate the impact of cryptic invasions in many different groups. Our results show that microsatellite analysis, combined with Bayesian statistics (Structure software), can efficiently be adapted to large-scale assessment of biotype frequency at the regional scale (especially if previously undescribed biotypes are involved and, therefore, external references are lacking). Some of the basic hypotheses required by Structure are not met by our dataset, i.e. there are significant linkage disequilibria and heterozygote

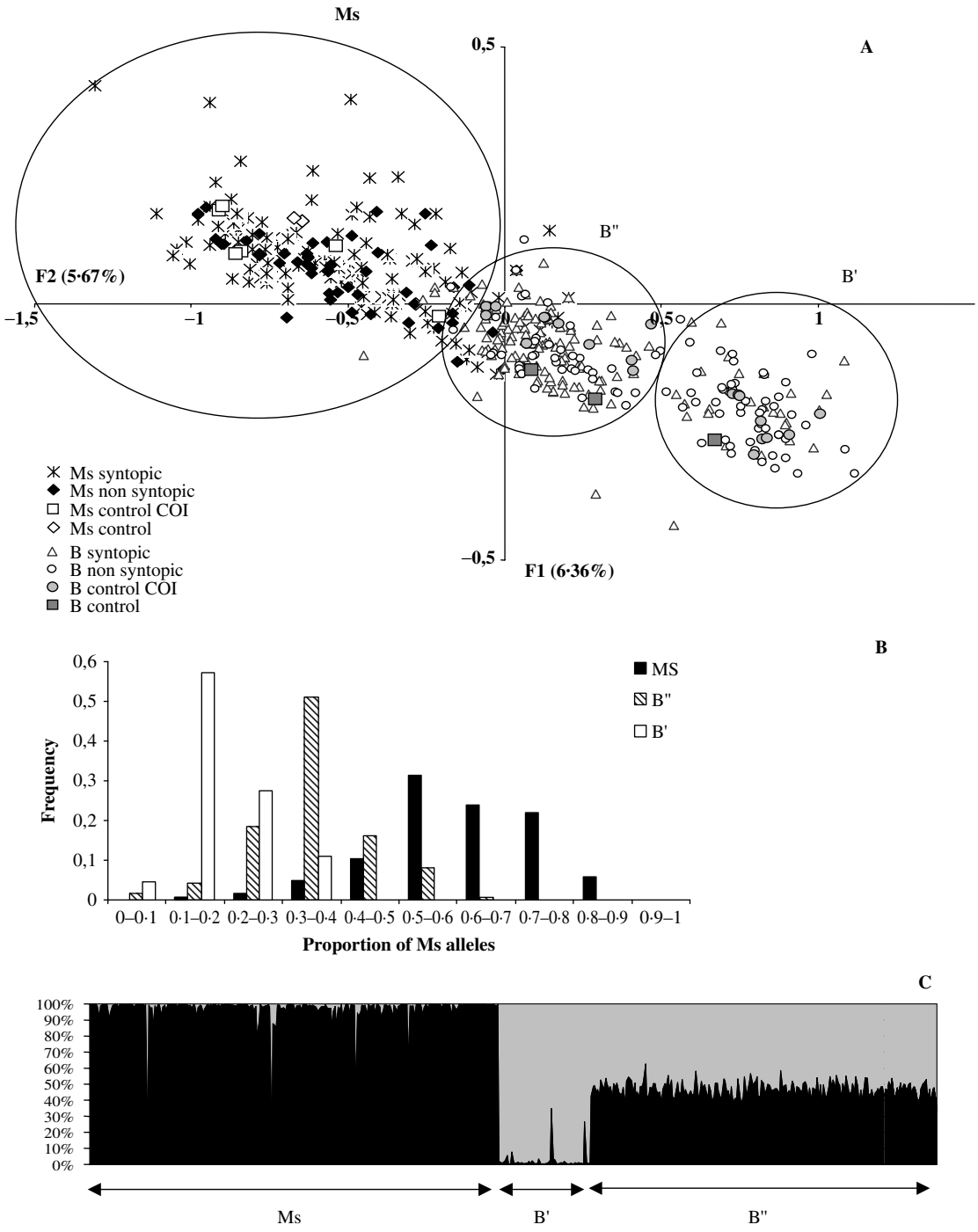

Fig. 3. Genetic structure of *Bemisia tabaci* sampled over La Réunion and proportion of biotype Ms alleles per group of genotype. (*A*) First factorial plane of the genetic correspondence analysis. Biotypes B and Ms as identified by Structure are represented by different symbols; we also distinguish two subgroups within the B biotype: B′ and B″ (see text). (*B*) Distribution of the proportion of alleles of the Ms type within individuals of each of the three groups Ms, B′ and B″. (*C*) Summary plot of estimates of Q (the estimated membership coefficient for each individual in each cluster) given by the software Structure v2.1 with the admixture option. Each of the 567 individuals is represented by a single vertical line broken into *K* populations ($K=2$ in our case), with length proportional to the inferred proportion of B ancestry; individuals from the two subgroups B′ and B″ have been represented in different sections of the graph to emphasize their genetic differences.

deficiencies within the inferred groups. This is partly due to a spatial structuring of the populations (significant F_{st} among populations within biotypes) and potentially to the presence of null alleles or scoring artefacts. However, several lines of evidence support the validity of the groups given by Structure,

Table 4. *Deviance analysis on the effect of space, time and host plant on relative proportions of the B and Ms biotypes*

Factors	d.f.	Changes in deviance relative to reference model	P (χ^2)	Reference model	Deviance of reference model
Crop or weed (c)	1	18·67	<0·001	y + p(s) + y.s	464·57
Host plant (h(c))	9	24·92	0·003	y + p(s) + y.s + c	464·53
Year (y)	2	6·49	0·039	h(c) + p(s)	455·19
Sector (s)	3	20·02	<0·001	h(c) + y	619·49
Population (sector) (p(s))	15	150·8	<0·001	h(c) + y + s + y.s	546·41
Year. sector (y.s)	6	27·72	<0·001	h(c) + p(s) + y	448·7
Residual	525	420·98			
Global deviance	561				774·99

Interactions between factors are denoted by dots (x.y), while nested factors are denoted by parentheses (x(y)). Only significant factors have been retained in the final model. Each factor or interaction is tested using changes in deviance between two models: with and without the factor. The model without the factor is denoted 'reference model'.

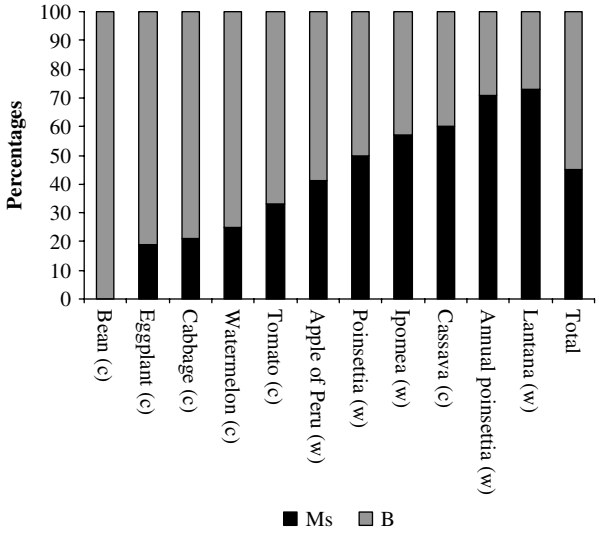

Fig. 4. Percentages of biotypes B (in grey) and Ms (in black) on different host plants, summed over all sites and times. Host plants were separated into two categories: weed (w) or cultivated (c).

despite the violation of assumptions. First, the assignation probabilities given by Structure are always very close to 0 or 1, suggesting that it has efficiently converged towards well-defined group limits. Second, the differentiation (F_{ct}) between the two groups assumed to be biotypes B and Ms is very high (0·32) relative to geographic variation within biotypes ($F_{sc}=0·04$). Third, the two groups are well separated in the first factorial plane of the COA, an independent type of analysis that makes no assumptions regarding Hardy–Weinberg and linkage equilibrium. Fourth, all the controls (5 laboratory controls plus 35 COI-sequenced individuals) have been correctly classified. This consistency suggests that, as the two groups show pronounced differentiation, the Structure classification is relatively robust. Two other microsatellite banks recently published (De Barro *et al.*, 2003; Tsagkarakou & Roditakis, 2003) will provide more microsatellite markers which might be very interesting for generalizing our kind of study to large-scale sampling of different biotypes.

(ii) *Genetic variation within biotypes*

Both biotypes displayed considerable genetic variation at the microsatellite loci analysed (up to $H_e=0·9$). Biotype Ms has already been found in other islands of the region (Delatte *et al.*, 2005), and its area of origin could be on another island, especially given that La Réunion is one of the youngest islands in the area (2·2 million years). Genetic diversities are large and of similar magnitude within biotypes Ms and B, one of which is ancient and indigenous, the other a recently arrived invader. This suggests that the introduction of biotype B was not associated with a severe genetic bottleneck. More than one introduction might be responsible for its invasion. Interestingly, in a few years the introduced biotype has reached the same levels of geographic structure (F_{st}) across the island as the local biotype.

(iii) *Ecological, spatial and temporal determinants of biotype distributions*

Genetic screening provides the opportunity to analyse the details of biotype distribution. Our first main result is to document a rapid spread of the invasive B biotype towards all parts of the island, although it has not wiped out the resident biotype

in any sector. In almost all sampled sites individuals of the two biotypes are found in sympatry. However, the proportions of the two biotypes are far from constant and vary according to geographic or ecological factors. The B biotype is found predominantly in the north, west and south parts of the island, while Ms predominates in the east, although it is present at all the sampled sites. This is consistent with field observations. Indeed the first *B. tabaci* upsurge (1991) was observed in the south and so were the first TYLCV symptoms (at Saint Pierre in 1997). According to epidemiological surveys in 1997–1999, TYLCV spread from Saint Joseph (south, close to Manapany) to La Possession (north-west, close to Saint Paul) in 2 years, and no symptoms of TYLCV have been reported so far from the eastern part of La Réunion (Reynaud *et al*., 2003). A general scheme of the invasion process of the B biotype seems to emerge, with a starting point somewhere in the south-west and an invasion flow down south and up north, the east being the last part of the island to be reached by the invasion.

Ecological differences are visible between the two biotypes, especially with regard to host plants (Figs 1, 4). The low frequency of the B biotype in the very humid eastern regions (Table 4) could reflect climatic preferences and/or the distribution of host plants. Indeed B is dominant on vegetable crops, whereas Ms is found on weeds, which are abundant in the east where there are relatively few vegetable crops (Fig. 4).

(iv) *Interactions between biotypes*

In their review Reitz & Trumble (2002) provide compelling evidence supporting the importance of interspecific competition as a mechanism structuring insect species communities, despite the common belief that food is not limited for these species. De Bach (1966) noticed that established species are more often displaced than completely excluded, and this might eventually happen to the Ms biotype. However, we cannot exclude the possibility that the Ms biotype is in the process of being almost completely excluded by B, as happened to the A biotype in the United States (Brown *et al*., 1995*a*).

Another potential outcome to consider is inter-biotype mating and emergence of a hybrid swarm. The COA results suggests two distinct subgroups within biotype B, one of which (B′) is genetically distant from the Ms biotype while the other (B″) has a higher frequency of alleles typical of the Ms group. Two hypotheses may account for these two groups: (i) either there are two biotypes, not one, within the cluster identified by Structure and identified as B, or (ii) there is some local introgression of Ms alleles within B genotypes, and B′ and B″ represent, respectively, pure and introgressed forms of the B biotype. The second hypothesis is more likely for several reasons. First, the genetic differences between B′ and B″ bear mainly on the frequency of alleles typical of the Ms biotype. Indeed, B″ is exactly intermediate between B′ and Ms in the first factorial plane of the COA analysis, and Structure classifies B″ individuals as individuals with mixed ancestry. This is not expected if B′ and B″ represent two independent biotypes. Second, COI sequences are available in both B′ (10 individuals) and B″ (10 individuals), and they are all identical to reference B sequences. Third, we observed more B″ individuals in syntopic populations, i.e. when Ms and B individuals are found together, and more B′ individuals in non-syntopic populations. This is consistent with the idea that introgressed individuals should be found more frequently when the B and Ms biotypes are in close contact.

This is still enough to observe two striking characteristics, yet common to many other studies of introgression between species. First, introgression in the B″ group is unequal across loci, as it is very moderate at all loci but two (P62 and P5), for which allele frequencies nearly equal those of the Ms biotype. This is a quite regular phenomenon in the hybridization literature (Barton & Hewitt, 1985; Bierne *et al*., 2003), where it has been attributed to differences in the selection regime among different parts of the genome. Such differences may depend, for example, on the physical proximity to a gene involved in hybrid breakdown, or on the contrary to a gene where a favourable allele exists in one species and spreads into the other. Second, introgression appears to be asymmetrical, i.e. alleles typical of Ms introgress into a B background but the reverse has not been found. This is confirmed by the fact that maternally transmitted mitochondrial sequences in the putative introgressed group (B″) are of the B type, in agreement with the majority of microsatellite alleles. This unequal introgression, in our case, might occur if two conditions are fulfilled: (i) Mating is one-way (Ms males mate with B females), so that hybrids always inherit the B cytoplasm. In insects, such situations have often been associated with cytoplasmic incompatibility due to bacterial endosymbionts such as *Wolbachia* (Nirgianaki *et al*., 2003; O'Neil *et al*., 1992). More research is needed to evaluate the role of *Wolbachia* in the different biotypes of *B. tabaci*. (ii) Male hybrids have very low fitness or can only mate with B females. This would prevent male recombined genotypes crossing with Ms females. It has often been observed that genetic incompatibilities between different incipient species or differentiated populations tended to be recessive (Dobzhansky–Muller incompatibilities; Gravilets, 2003). In this context, it would not be surprising that the mixing of the two genomes through

hybridization and recombination affects haploid males more than diploid females.

The assumption that *B. tabaci* biotypes are genetically isolated due to mating incompatibilities and host-related specialization prevails in the literature, though actual tests are rare (De Barro & Hart, 2000, De Barro *et al*., 2005; Maruthi *et al*., 2004). Biotypes are ecologically separated by their distinctive host ranges (see above for the case of B and Ms). However, this is only a partial isolation, as attested by the many individuals found in syntopy in the present study. Introgressed individuals (the B″ group) are now widespread in the island and dominant at almost all sites over non-introgressed (B′) individuals. However, there is no clinal geographical structure typical of classical hybrid zones. More data are needed to definitively document the ability of the two biotypes to interbreed when present in the same plant, and more evolutionary time is needed to confirm the extent, and establish the long-term outcome, of introgression in the field.

We would like to thank Isabelle Litrico for her advice and Martial Grondin for his technical assistance in the collection of whiteflies. This study was funded by CIRAD-3P and the Conseil Régional of la Réunion. H.D. is a recipient of a sandwich PhD from Wageningen University.

Appendix A. Allele frequencies at each of the eight loci of groups Ms, B″, B′ and B (total B′ + B″), respectively

Alleles classified as typical of Ms (see text) are in italics. Alleles were conservatively grouped into size classes (see text), so each allele represents a range of sizes.

P53					P62					P7					P5				
Allele size	Ms	B″	B′	B	Allele size	Ms	B″	B′	B	Allele size	Ms	B″	B′	B	Allele size	Ms	B″	B′	B
130–135	0·266	0·085	0·022	0·064	*180*	0·01	0	0	0·000	131	0	0·003	0	0·002	181	0	0·011	0	0·008
136–140	0·128	0·418	0·467	0·434	201	0	0	0·017	0·006	154	0	0·018	0·034	0·023	202–204	0·008	0·034	0·183	0·085
141–145	0·007	0·06	0·093	0·071	203–205	0·033	0·025	0·075	0·041	*160*	0·035	0·016	0	0·011	207–210	0·038	0·14	0·783	0·358
148–150	0·021	0·016	0·011	0·015	*206–208*	0·247	0·105	0·063	0·091	173–176	0·022	0·068	0·18	0·103	*211*	0·102	0	0·006	0·002
151–157	0·312	0·197	0·165	0·186	209–210	0·062	0·066	0·822	0·312	178–183	0·75	0·878	0·787	0·849	*212–214*	0·565	0·815	0·028	0·548
158–159	0·05	0·063	0·093	0·073	*211–214*	0·643	0·793	0·006	0·537	*184–188*	0·193	0·013	0	0·009	*215–219*	0·281	0	0	0·000
160–164	0·209	0·123	0·143	0·130	215–220	0·005	0·006	0·011	0·007	207	0	0·003	0	0·002	*227*	0·005	0	0	0·000
165–169	0·007	0·038	0·005	0·027	226	0	0·006	0·006	0·006	214	0	0·003	0	0·002					
Synthetic Ms	0·808	0·421	0·341	0·395	Synthetic Ms	0·900	0·898	0·069	0·628	Synthetic Ms	0·228	0·029	0	0·020	Synthetic Ms	0·953	0·815	0·034	0·550

P41					P59					P11					P32				
Allele size	Ms	B″	B′	B	Allele size	Ms	B″	B′	B	Allele size	Ms	B″	B′	B	Allele size	Ms	B″	B′	B
132	0·007	0	0	0·000	*118*	0·005	0	0	0·000	*169*	0·007	0·008	0	0·005	192–196	0·109	0·457	0·542	0·486
152	0·017	0·011	0	0·008	*171*	0·017	0	0·022	0·008	*174–176*	0·069	0·005	0·006	0·005	*200–203*	0·635	0·387	0·333	0·369
158–162	0·03	0·122	0·207	0·149	175–178	0·072	0·214	0·124	0·183	178–182	0·92	0·982	0·994	0·986	*204–206*	0·121	0·034	0·042	0·036
163–167	0·947	0·866	0·793	0·843	*179–182*	0·636	0·074	0·022	0·056	*186*	0·002	0	0	0·000	*207–209*	0·112	0·015	0·024	0·018
					183–186	0·069	0·335	0·455	0·377	194	0	0·005	0	0·004	210–214	0·023	0·107	0·06	0·091
					193–198	0·17	0·03	0·017	0·025	*208*	0·002	0	0	0·000					
					200–204	0	0·033	0·073	0·047										
					205–209	0·002	0·163	0·202	0·177										
					210–212	0	0·003	0·051	0·019										
					213–215	0·024	0·128	0·022	0·091										
					216–219	0·005	0·021	0·011	0·017										
Synthetic Ms	0·971	0·877	0·793	0·851	Synthetic Ms	0·828	0·104	0·061	0·089	Synthetic Ms	0·078	0·013	0·006	0·010	Synthetic Ms	0·868	0·436	0·399	0·423

Appendix B. *P* values of the linkage disequilibrium given by Genepop 3·3 for biotypes B and Ms

Biotype B	P53	P62	P7	P5	P41	P59	P11	
P53	----------							
P62	NS	----------						
P7	***	***	----------					
P5	***	***	*	----------				
P41	NS	*	***	***	----------			
P59	***	NS	NS	***	***	----------		
P11	NS	NS	NS	NS	NS	***	----------	
P32	NS	***	NS	NS	NS	NS	NS	----------

Biotype Ms	P53	P62	P7	P5	P41	P59	P11	
P53	----------							
P62	NS	----------						
P7	***	***	----------					
P5	**	***	***	----------				
P41	NS	*	NS	**	----------			
P59	NS	*	**	NS	NS	----------		
P11	NS	NS	NS	NS	NS	NS	----------	
P32	***	NS	NS	NS	NS	*	NS	----------

* $P<0{\cdot}05$; ** $P<0{\cdot}01$; *** $P<0{\cdot}001$; NS, not significant.

Appendix C. Raw allele frequencies at each of the eight loci of B and Ms, respectively, given as /allele size B frequency Ms frequency/

Locus P53: 130 0 0.003/131 0.038 0.19/132 0 0.037/133 0.008 0.008/134 0.012 0/135 0.002 0/136 0.004 0/137 0.172 0.021/138 0.182 0.047/139 0.08 0.063/140 0.012 0.005/141 0.026 0.003/142 0.02 0.005/143 0.014 0/144 0.004 0.003/145 0.002 0/148 0.002 0.003/149 0.004 0/150 0.006 0.018/151 0.002 0.003/152 0.006 0.055/153 0.002 0.103/154 0.07 0.024/155 0.082 0.04/156 0.02 0.074/157 0.008 0.026/158 0.036 0.04/159 0.03 0.016/160 0.008 0.068/161 0.008 0.074/162 0.05 0.029/163 0.058 0.032/164 0.002 0.005/165 0 0.003/166 0.03 0/168 0 0.003/169 0 0.003/

Locus P62: 180 0 0.012/201 0.006 0/203 0 0.015/204 0.039 0.018/205 0 0.003/206 0.004 0.003/207 0.071 0.241/208 0.016 0.015/209 0.083 0.062/210 0.227 0.009/211 0.006 0/212 0.014 0.112/213 0.075 0.012/214 0.445 0.5/218 0.008 0/226 0.006 0/

Locus P7: 131 0.002 0/154 0.018 0/155 0.008 0/160 0.012 0.035/161 0 0.003/170 0.004 0/171 0.002 0/172 0.002 0/173 0.006 0.003/174 0.029 0/175 0.045 0.02/176 0.022 0.003/178 0.053 0.01/179 0.166 0.151/180 0.356 0.482/181 0.246 0.093/182 0.018 0/183 0.002 0.005/184 0.004 0.02/185 0.002 0.055/186 0.002 0.101/187 0 0.013/188 0 0.008/207 0.002 0/214 0.002 0/

Locus P5: 180 0.004 0/181 0.004 0/202 0.002 0/203 0.086 0.009/204 0.004 0/206 0.014 0/207 0.012 0/208 0.01 0/209 0.318 0.05/211 0.002 0.107/212 0 0.006/213 0.521 0.503/214 0.023 0.018/215 0 0.009/217 0 0.228/218 0 0.03/219 0 0.036/227 0 0.006/

Locus P41: 132 0 0.008/152 0 0.019/153 0.004 0/154 0.004 0/158 0.002 0.008/159 0.009 0/160 0.002 0.012/161 0.04 0/162 0.062 0.008/163 0.036 0.004/165 0.004 0.015/166 0.598 0.754/167 0.234 0.173/168 0.004 0/

Locus P11: 169 0 0.008/170 0.006 0/174 0.004 0.025/175 0.002 0.031/176 0 0.005/177 0 0.008/178 0.004 0/179 0.221 0.162/180 0.39 0.571/181 0.2 0.175/182 0.107 0.01/183 0.064 0/186 0 0.003/194 0.004 0/208 0 0.003/

Locus P32: 191 0.011 0/192 0.022 0.01/193 0.048 0.029/194 0.233 0.049/195 0.174 0.026/196 0.002 0.003/199 0.002 0/200 0.046 0.003/201 0.054 0.026/202 0.054 0.059/203 0.194 0.507/204 0.015 0.026/205 0.02 0.069/206 0.004 0.043/207 0.013 0.007/208 0.002 0.023/209 0.007 0.101/210 0.007 0/211 0.002 0/212 0.002 0/213 0.08 0.02/214 0.004 0/215 0.004 0/

Locus P59: 118 0 0.006/171 0 0.019/172 0.009 0/174 0.019 0/175 0.110 0.036/176 0.030 0.038/177 0.019 0/178 0.002 0.006/179 0.002 0.046/180 0.025 0.254/181 0.023 0.246/182 0.006 0.063/183 0.049 0.006/184 0.125 0.036/185 0.100 0.008/186 0.051 0.014/187 0.030 0/188 0.025 0/190 0.002 0/193 0 0.011/194 0.011 0.033/195 0.004 0.107/196 0.006 0.027/197 0.004 0.006/198 0 0.006/199 0.002 0/202 0.013 0/203 0.034 0/204 0.002 0.003/205 0.042 0/206 0.036 0/207 0.042 0/208 0.047 0.003/209 0.004 0/210 0.002 0/211 0.013 0/212 0.006 0/213 0.004 0.003/214 0.081 0.022/217 0.011 0/218 0.006 0.006/219 0.002 0/

References

Barton, N. H. & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**, 113–148.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. (1996–2004). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

Bellows, T. S., Perring, T. M., Gill, R. J. & Headrick, D. H. (1994). Description of a species of *Bemisia* (Homoptera, Aleyrodidae) infesting North American agriculture. *Annals of the Entomological Society of America* **87**, 195–206.

Bierne, N., Daguin, C., Bonhomme, F., David, P. & Borsa, P. (2003). Direct selection on allozymes is not required to explain heterogeneity among markers across a *Mytilus* hybrid zone. *Molecular Ecology* **12**, 2505–2510.

Billote, N., Lagoda, P. J. L., Risterucci, A. & Baurens, F. (1999). Microsatellite enriched libraries, applied methodology for the development of ISSR markers in tropical crops. *Fruits* **54**, 277–288.

Bird, J. & Maramorosch, K. (1978). Virus and virus diseases associated with whiteflies. *Advances in Virus Research* **22**, 55–110.

Bourriquet, G. (1938). Note concernant les maladies des plantes cultivées à la Réunion. *Revue Agricole Réunion* **43**, 33–38.

Brown, J. K. (2000). Molecular markers for the identification and global tracking of whitefly vector-Begomovirus complexes. *Virus Research* **71**, 233–260.

Brown, J. K., Frohlich, D. R. & Rosell, R. C. (1995*a*). The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or a species complex? *Annual Review of Entomology* **40**, 511–534.

Brown, J. K., Coats, S., Bedford, I. D., Markham, P. G., Bird, J. & Frohlich, D. (1995*b*). Characterization and distribution of esterase electromorphs in the whitefly, *Bemisia tabaci* (Genn.) (Homoptera, Aleyrodidae). *Biochemical Genetics* **33**, 205–214.

Brownstein, M. J., Carpten, J. D. & Smith, J. R. (1996). Modulation of non-templated nucleotide addition by Taq DNA polymerase, primer modification that facilitates genotyping. *Biotechniques* **20**, 1004–1010.

Byrne, F. J., Castle, S., Prabhaker, N. & Toscano, N. C. (2003). Biochemical study of resistance to imidacloprid in B biotype *Bemisia tabaci* from Guatemala. *Pest Management Science* **59**, 347–352.

Chakraborty, R., De Andrade, M., Daiger, S. P. & Budowle, B. (1992). Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic application. *Annals of Human Genetics* **56**, 45–57.

Colvin, J., Omongo, C. A., Maruthi, M. N., Otim-Nape, G. W. & Thresh, J. M. (2004). Dual begomovirus infections and high *Bemisia tabaci* populations: two factors driving the spread of a cassava mosaic disease pandemic. *Plant Pathology* **53**, 577–584.

Costa, H. S. & Brown, J. K. (1991). Variation in biological characteristics and in esterase patterns among populations of *Bemisia tabaci* Genn. and the association of one population with silverleaf symptom development. *Entomologia Experimentalis and Applicata* **61**, 211–219.

Costa, H. S., Brown, J. K., Sivasupramaniam, S. & Bird, J. (1993). Regional distribution, insecticide resistance, and reciprocal crosses between the A- and B-biotypes of *Bemisia tabaci*. *Insect Sciences and its Application* **14**, 255–266.

Crawley, M. J. (1993). *GLIM for Ecologists*. Oxford: Blackwell Science.

De Bach, P. (1966). The competitive displacement and coexistence principles. *Annual Review of Entomology* **11**, 183–212.

De Barro, P. J. & Hart, P. J. (2000). Mating interactions between two biotypes of the whitefly, *Bemisia tabaci* (Hemiptera, Aleyrodidae) in Australia. *Bulletin of Entomological Research* **90**, 103–112.

De Barro, P. J., Driver, F., Trueman, J. W. H. & Curran, J. (2000). Phylogenetic relationships of world populations of *Bemisia tabaci* (Gennadius) using ribosomal ITS1. *Molecular Phylogenetics and Evolution* **16**, 29–36.

De Barro, P. J., Scott, K. D., Graham, G. C., Lange, C. L. & Schutze, M. K. (2003). Isolation and characterization of microsatellite loci in *Bemisia tabaci*. *Molecular Ecology Notes* **3**, 40–43.

De Barro, P. J., Trueman, J. W. H. & Frohlich, D. R. (2005). *Bemisia argentifolii* is a race of *B. tabaci* (Hemiptera, Aleyrodidae), the molecular genetic differentiation of *B. tabaci* populations around the world. *Bulletin of Entomological Research* **95**, 193–203.

Delatte, H., Reynaud, B., Granier, M., Thornary, L., Lett, J. M., Goldbach, R. & Peterschmitt, M. (2005). A new silverleaf inducing biotype of *Bemisia tabaci* (Hemiptera, Aleyrodidae) Ms, indigenous to the islands of the South West Indian Ocean. *Bulletin of Entomological Research* **95**, 29–35.

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

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

Frohlich, D. R., Brown, J. K., Bedford, I. & Markham, P. (1996). Mitochondrial 16S ribosomal subunit as a molecular marker in *Bemisia*, and implications for population variability. In *Bemisia 1995: Taxonomy, Biology, Damage, Control and Management* (ed. D. Gerling & R. T. Mayer), pp. 143–145. Andover, UK: Intercept.

Frohlich, D. R., TorresJerez, I., Bedford, I. D., Markham, P. G. & Brown, J. K. (1999). A phylogeographical analysis of the *Bemisia tabaci* species complex based on mitochondrial DNA markers. *Molecular Ecology* **8**, 1683–1691.

Gavrilets, S. (2003). Perspective. Models of speciation: what have we learned in 40 years? *Evolution* **57**, 2197–2215.

Genner, M. J., Michel, E., Erpenbeck, D., De Voogd, N., Witte, F. & Pointier, J. P. (2004). Camouflaged invasion of Lake Malawi by an Oriental gastropod. *Molecular Ecology* **13**, 2135–2141.

Gill, R. (1992). A review of the sweetpotato whitefly in Southern California. *Pan-Pacific Entomologist* **68**, 144–152.

Horowitz, A. R., Kontsedalov, S., Khasdan, V. & Ishaaya, I. (2005). Biotypes B and Q of *Bemisia tabaci* and their relevance to neonicotinoid and pyriproxyfen resistance. *Archives of Insect Biochemistry and Physiology* **58**, 216–225.

Jones, D. R. (2003). Plant viruses transmitted by whiteflies. *European Journal of Plant Pathology* **109**, 195–219.

Markham, P. G., Bedford, I. D., Liu, S., Frolich, D. F., Rosell, R. & Brown, J. K. (1995). The transmission of

geminiviruses by biotypes of *Bemisia tabaci* (Gennadius). In: *Bemisia 1995: Taxonomy, Biology, Damage, Control and Management* (ed. D. Gerling & R. T. Mayer), pp. 69–75. Andover, UK: Intercept.

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

Mooney, H. A. & Cleland, E. E. (2001). The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences of the USA* **98**, 5446–5451.

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

Nirgianaki, A., Banks, G., Frohlich, D. R., Veneti, Z., Braig, H. R., Miller, T. A., Bedford, I. D., Markham, P. G., Savakis, C. & Bourtzis, K. (2003). Wolbachia infections of the whitefly *Bemisia tabaci*. *Current Microbiology* **47**, 93–101.

Oliveira, M. R. V., Henneberry, T. J. & Anderson, P. (2001). History, current status, and collaborative research projects for *Bemisia tabaci*. *Crop Protection* **20**, 709–723.

O'Neil, S. L., Giordano, R., Colbert, A. M. E., Karr, T. & Robertson, H. M. (1992). 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. *Proceedings of the National Academy of Sciences of the USA* **89**, 2699–2702.

Pascual, S. & Callejas, C. (2004). Intra- and interspecific competition between biotypes B and Q of *Bemisia tabaci* (Hemiptera, Aleyrodidae) from Spain. *Bulletin of Entomological Research* **94**, 369–375.

Pascual, M., Aquardo, C. F., Soto, V. & Serra, L. (2001). Microsatellite variation in colonizing and palearctic populations of *Drosophila subobscura*. *Molecular Biology and Evolution* **18**, 731–740.

Perring, T. M. (2001). The *Bemisia tabaci* species complex. *Crop Protection* **20**, 725–737.

Perring, T. M., Cooper, A., Kazmer, D. J., Shields, C. & Shields, J. (1991). New strain of sweetpotato whitefly invades California vegetables. *California Agriculture* **45**, 10–12.

Peterschmitt, M., Granier, M. & Mekdoud, R. (1999). First report of tomato yellow leaf curl virus in Réunion Island. *Plant Disease* **83**, 303.

Polston, J. E. & Anderson, P. K. (1997). The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. *Plant Disease* **81**, 1358–1368.

Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Raymond, M. & Rousset, F. (1995). GENEPOP (version 3.3), population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248–249.

Reitz, S. R. & Trumble, J. T. (2002). Competitive displacement among insects and arachnids. *Annual Review of Entomology* **47**, 435–465.

Reynaud, B., Wuster, G., Delatte, H., Soustrade, I., Lett, J. M., Gambin, O. & Peterschmitt, M. (2003). Les maladies à bégomovirus chez la tomate dans les départements français d'Outre-Mer. *Phytoma* **562**, 13–17.

Ribeiro, S. G., Ambrozevicius, L. P., Avila, A. C., Bezerra, I. C., Calegario, R. F., Fernandes, J. J., Lima, M. F., Mello, R. N. d., Rocha, H. & Zerbini, F. M. (2003). Distribution and genetic diversity of tomato-infecting begomoviruses in Brazil. *Archives of Virology* **148**, 281–295.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145**, 1219–1228.

Saitou, N. & Nei, M. (1987). The neighbour joining method: A new method for reconstructing phylogenetic trees. *Molecular and Biological Evolution* **4**, 406–425.

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**, 6463–6471.

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

Varma, A. & Malathi, V. G. (2003). Emerging geminivirus problems: a serious threat to crop production. *Annals of Applied Biology* **142**, 145–164.

Weir, B. S. & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.