MtDNA variability in whitefly (*Bemisia tabaci*) populations in Brazil

G.E. Valle[1], A.L. Lourenção[1], M.I. Zucchi[2], J.B. Pinheiro[3] and A.G. Abreu[2]

[1]Centro de Fitossanidade, Instituto Agrônomico de Campinas, Campinas, SP, Brasil
[2]Agência Paulista de Tecnologia dos Agronegócios, Piracicaba, SP, Brasil
[3]Departamento de Genética, Escola Superior de Agronomia "Luiz de Queiroz",
Universidade de São Paulo, Piracicaba, SP, Brasil

Corresponding author: G.E. Valle
E-mail: gevalle@yahoo.com.br

ABSTRACT. *Bemisia tabaci* (Hemiptera: Aleyrodidae) consists of a complex of morphologically indistinct biotypes that vary mainly in their capacity to transmit plant viruses and to induce physiological disorders in plants of economic importance. The adaptability of *B. tabaci* to many regions of the world has fostered the appearance of various biotypes and has resulted in a broad spectrum of host plants. Our goal was to identify which biotypes were present in four *B. tabaci* populations in Brazil. We quantified genetic variability between and within populations. Three individuals were collected from three host plant species: two populations on soybean (Campinas and Rondonópolis), one on pumpkin (Barreiras) and one on tomato (Cruz das Almas) in three States of Brazil (São Paulo, Mato Grosso, and Bahia). We chose one sequence of the B biotype, obtained from GenBank; the Campinas population, which had been previously characterized as biotype B, was used as a control for this biotype. We also included one sequence of the Q biotype, obtained from GenBank, as an outgroup. The COI region of the mtDNA gene was partially amplified with the CI-J-2195 and L2-N-3014 pair of primers, and the reaction products were sequenced. Based on distance-

based algorithm analyses, we found that all haplotypes belong to biotype B, which was confirmed by the haplotype network. Genetic structure analyses showed that the host plant species does not influence population structuring of this pest; only the geographic location mattered.

Key words: Insecta; Aleyrodidae; Whitefly; Genetic diversity; Molecular characterization

INTRODUCTION

Whiteflies are insects that cause direct damages to plants and transmit plant viruses (Berlinger, 1986; Norman et al., 2011). Worldwide, they are considered to be an extremely important group, transmitting more than 110 different plant viruses; in addition, they are the only transmitters of geminivirus (Jones, 2003). One of the most important species is *Bemisia tabaci* (Genn.), which is currently considered to be a complex (Brown et al., 1995a; Perring, 2001) with a wide host range. These whiteflies develop on over 900 plant species (Perring, 2001; Berry et al., 2004) and are considered to be pests of agroeconomic importance (Inbar and Gerling, 2008), where the so-called biotype B is ranked as one of the world's worst invasive organisms (Global Invasive Species Database, http://www.issg.org/database).

The existence of *B. tabaci* biotypes was proposed around 1950 to describe distinct populations of this insect based on their adaptation to the host plant and their capacity to transmit viruses (Mound, 1963). Biotype B is considered to be a serious agricultural pest and was once described as a new species (Bellows et al., 1994). In Brazil, the biotype B was introduced in the early 1990s (Lourenção and Nagai, 1994). Currently, more than 20 distinct biotypes have been characterized, designated A to T (Bedford et al., 1994; Brown et al., 1995b; Banks et al., 1999; Perring, 2001).

Molecular technology has provided valuable molecular marker investigational tools. It is believed that knowledge of the genetic differentiation of *B. tabaci* populations may lead to important implications for the control of this pest in Brazil and worldwide, since it includes populations of different host plants, from different geographic regions, and genetic variations within these populations. The objective of this study was to characterize *B. tabaci* populations collected on three host plants from four regions of Brazil being one of these populations previously characterized as biotype B.

MATERIAL AND METHODS

Adults of *B. tabaci* whitefly populations were collected in the field, identified, and kept in 70% ethanol at -4°C for later molecular analysis. Mitochondrial sequences were used in the genetic diversity study of *B. tabaci* populations; the sequences were called P1 (P101 and P102), P2 (P204, P205, and P206), P3 (P307, P308, and P309), and P4 (P410 and P412), associated with populations from Barreiras (BA) - pumpkin (*Cucurbita moschata* D.), Rondonópolis (MT) - soybean (*Glycine max* L.), Cruz das Almas (BA) - tomato (*Lycopersicon esculentum* L.), and Campinas (SP) - soybean, respectively. The GenBank accession numbers are: HM584841 (P102), HM584842 (P101), HM584840 (P204), HM584839 (P205), HM584838 (P206), HM584837 (P307), HM584836 (P308), HM584835 (P309), HM584834

(P410), and HM584833 (P412). The Campinas population was collected from a reared stock maintained on soybean plants at Centro de Fitossanidade do Instituto Agronômico (IAC), and was considered to be a control. The rearing was implemented by collecting whitefly adults from a tomato crop in the city of Paulínia, São Paulo State, and transferred to soybean plants maintained under greenhouse conditions. After a few generations, adults were sent to Dr. Judith K. Brown, University of Arizona, USA, who identified them as *B. tabaci* biotype B. The identification is periodically confirmed by placing pumpkin plants at the rearing site to express leaf silvering, a physiological disorder resulting from feeding by immature insects of this biotype (Brown et al., 1995a). These populations were compared with two biotypes in GenBank (biotype B - AM696208 and biotype Q - FJ025793; called here biB and biQ, respectively).

The DNA extraction procedure adopted in the characterization and study of the genetic structure of the *B. tabaci* whitefly population was similar to that used by Lima et al. (2000). DNA was extracted from a whitefly female, which was macerated in a 1.5-mL Eppendorf tube in 60 µL extraction buffer (10 mM Tris-HCl, pH 8, 1 mM EDTA, 0.3% Triton X-100, 60 µg/mL proteinase K). The macerate was incubated for 15 min at 65°C, followed by heating for 10 min at 95°C. The samples were then stored at -20°C.

With respect to the mitochondrial regions, two partial sequences of the mtCOI gene (about 860 bp) were evaluated using primers CI-J-2195 and L2-N-3014 in three individuals of each *B. tabaci* population.

The polymerase chain reaction (PCR) mixture contained 20 ng DNA, 0.6 µM forward and reverse primer, 0.15 mM dNTPs, 1X buffer solution (50 mM KCl, 100 mM Tris-HCl, pH 8.5), 2.5 mM $MgCl_2$, 0.5 U Taq DNA polymerase, and autoclaved Milli-Q water added to a final volume of 25 µL.

The amplifications were carried out in a thermocycler (Biorad MyCycler), containing an initial denaturation stage of 2 min at 95°C, followed by 45 cycles with denaturation for 1 min at 95°C, annealing for 1 min at 52°C, and extension for 1 min at 72°C, finishing with a final extension of 5 min at 72°C.

Later, the fragments were purified with the Wizard® SV Gel and PCR Clean-Up System kit; the samples were then sequenced, using an Applied Biosystem Model 3730 DNA sequencer.

Sequences were obtained with the 3730/3730xl Data Collection Software v3.0 (Applied Biosystems) and processed with the BioEdit software (Hall, 1999). Poor-quality sequence data and sequences whose bases could not be firmly identified were excluded.

In phylogenetic analysis, it is important to determine the distance between a pair of sequences. These values are important for the application of both clustering methods, such as UPGMA (unweighted pair group method with arithmetic mean; Michener and Sokal, 1957) and NJ (neighbor-joining; Saitou and Nei, 1987), and probabilistic methods such as maximum parsimony (MP; Fitch, 1981) and maximum likelihood (ML; Felsenstein, 1981, 2004; Bryant et al., 2005). Analyses were performed using distances and characters to differentiate the sequences of mitochondrial loci.

The genetic structure of the populations was investigated using analysis of molecular variance (AMOVA), based on the variation of haplotype frequencies, taking into consideration the number of mutations between haplotypes. A hierarchical analysis was performed by separating the covariance components due to differences within populations, between populations in the same intrapopulation cluster versus the null hypothesis of non-association by means of 1000 permutations, using the Arlequin 2000 software program (Schneider et al., 2000).

RESULTS AND DISCUSSION

From the analyses of character-based optimization methods (Figures 1 and 2), the closest sequences were P410, P309, and P412, with 65.3% node consistency, P101, P102, P204, P307, and P308, with 58.0% by the ML clustering method. The following values were found by the MP method: 100% of node consistency between P412 P410, P309, P205, P206, biB, and biQ and 65% for P101, P102, P204, P307, and P308. It can be seen that the values observed in the sequences analyzed showed different patterns.
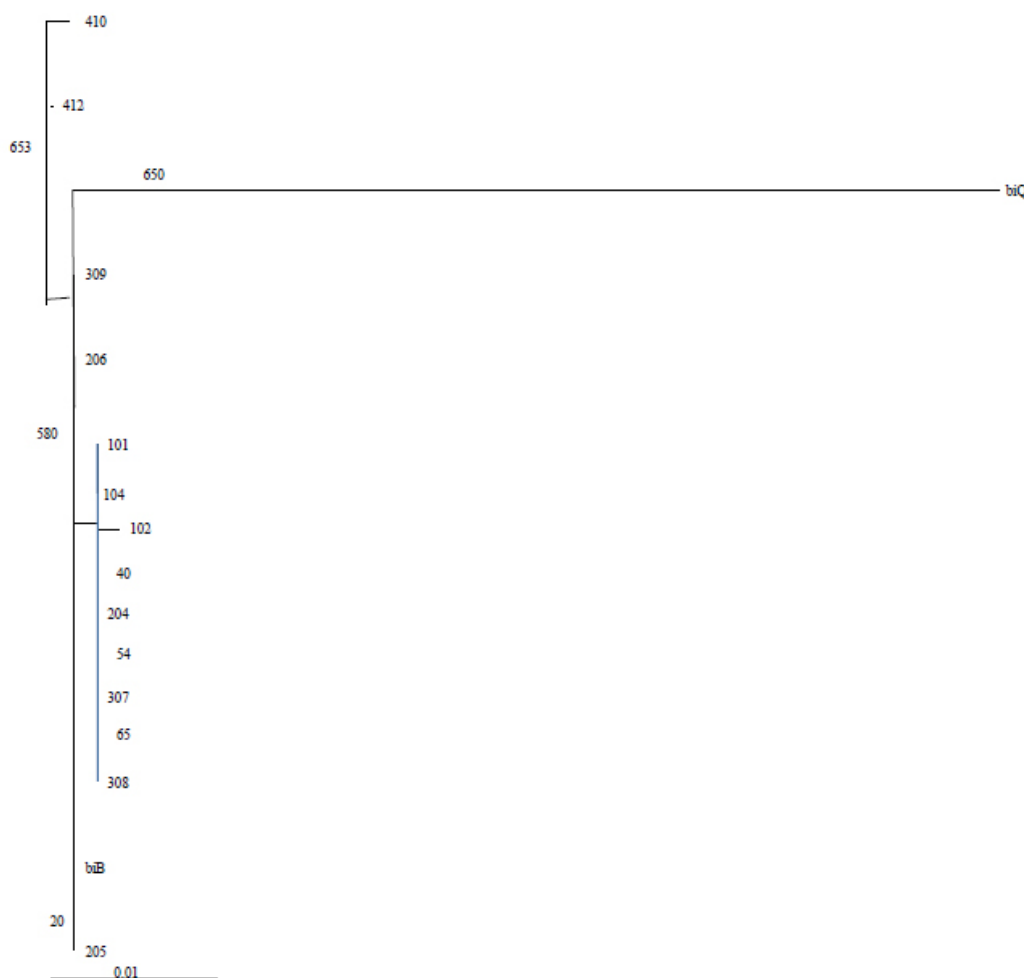

Figure 1. Genetic divergence pattern between four *Bemisia tabaci* populations, defined by the maximum likelihood method, based on genetic identity determined from the character for sequences of mitochondrial loci.
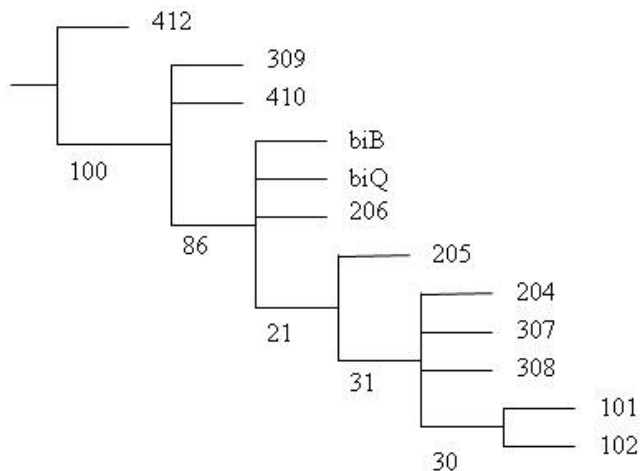
Figure 2. Genetic divergence pattern between four *Bemisia tabaci* populations, defined by the maximum parsimony method, based on the genetic identity determined from the character for sequences of mitochondrial loci.

According to the distance-based methods, NJ and UPGMA (Figures 3 and 4), the closest sequences were P410 and P412, with 64% node consistency, P205, P206, P309, and biB with 64%, and P204, P308, P101, and P102 with 65% by NJ clustering. The following values were found by UPGMA clustering: 44% for P206, biB, P309, and P205, and 45% for P101, P307, P308, and P204. Therefore, it can be concluded that, regardless of the method used, the distances observed in the sequences analyzed showed that the Q biotype was different in relation to the other haplotypes analyzed, which are believed to belong to the B biotype.
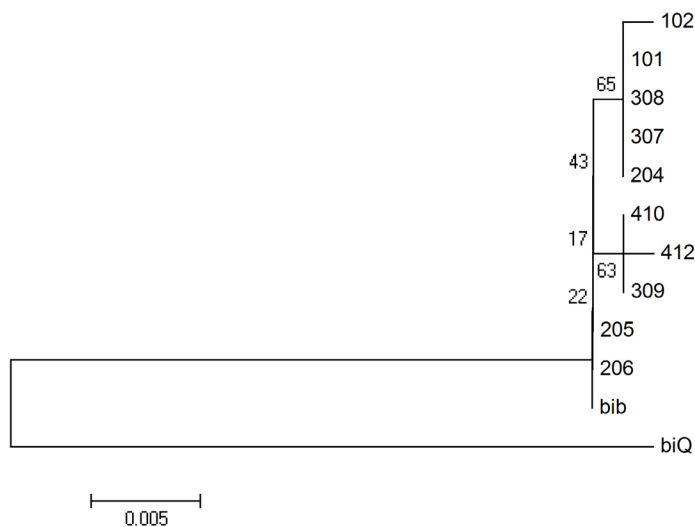

Figure 3. Genetic divergence pattern between four *Bemisia tabaci* populations, defined by the neighbor-joining cluster method, based on the genetic identity determined from K2P distances for sequences of mitochondrial loci.
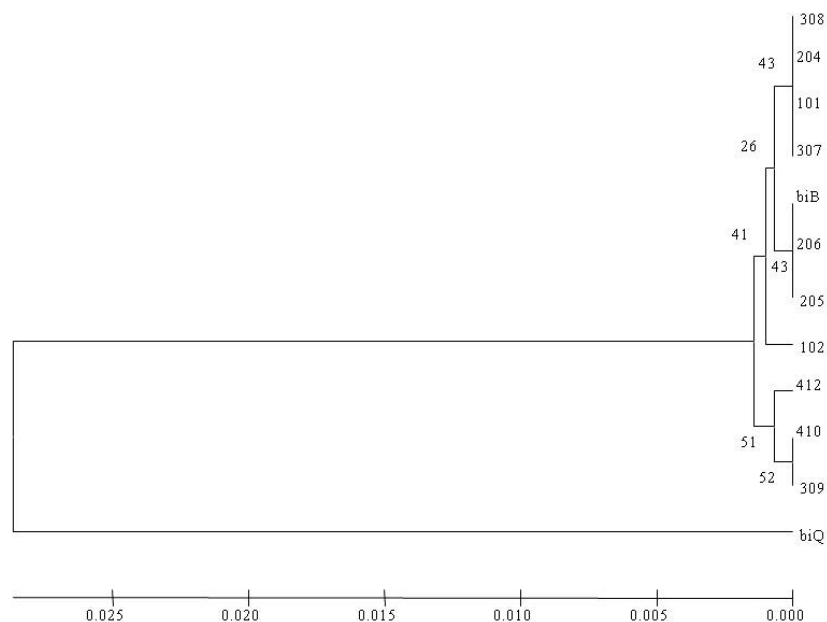
Figure 4. Genetic divergence pattern from an unrooted tree between four *Bemisia tabaci* populations, defined by UPGMA clustering, based on the genetic identity determined from K2P distances for sequences of mitochondrial loci.

Distance-based analyses (NJ and UPGMA; Figures 3 and 4) and optimization analyses (ML and MP; Figures 1 and 2) indicated different patterns, showing that only the distance-based analyses were adequate in this case.

Mitochondrial COI sequencing for *B. tabaci* has been used to identify relationships for the *B. tabaci* complex (Frohlich et al., 1999; Boykin et al., 2007). The taxonomic status of *B. tabaci* has been problematic; however, genetic differentiation and behavioral differences among populations have led to a proliferation of alternative designations and biotypes (Perring, 2001; De Barro et al., 2000, 2005; Boykin et al., 2007; Dinsdale et al., 2010), partly because authors have been constrained to treating the group as a single taxonomic species.

The *B. tabaci* populations seem to be more specific to the regions than the hosts and do not seem to be much different from one another: it is believed that they belong to the same biotype. Therefore, based on the data from the populations examined at the mitochondrial sequence level, it can be presumed that only biotype B is present in the Brazilian regions where collections were made.

AMOVA was conducted to detect the population structuring among the various haplotypes. In all forms of clustering used (Tables 1, 2, 3, and 4), there was a significant structuring between populations, indicated by significant F_{ST} values. Initially, AMOVA was conducted by clustering hosts (soybean, pumpkin, and tomato); it was observed that the variation between groups was negative, which indicates that intrapopulation variability is greater than interpopulation variability.

Next, we tested tomato and pumpkin populations versus soybean populations. The difference was positive, which shows that there is genetic structuring between different host groups.

Table 1. Analysis of molecular variance (AMOVA) for the whitefly *Bemisia tabaci* on different hosts.

Source of variation	Sum of squares	Variation components	Variation percentage
Between groups	1.93	-0.2318	-34.3059
Between populations within groups	1.63	0.5185	76.7462
Within populations	2.33	0.3889	57.5597
Total	5.90	0.6756	
F_{ST}		0.4244	

Table 2. Analysis of molecular variance (AMOVA) for the whitefly *Bemisia tabaci* on different host groups.

Source of variation	Sum of squares	Variation components	Variation percentage
Between groups	1.50	0.0826	11.1612
Between populations within groups	2.07	0.2685	36.2863
Within populations	2.33	0.3889	52.5525
Total	5.90	0.7400	
F_{ST}		0.4745	

Table 3. Analysis of molecular variance (AMOVA) for the whitefly *Bemisia tabaci* in different states.

Source of variation	Sum of squares	Variation components	Variation percentage
Between groups	3.13	0.3650	47.2544
Between populations within groups	0.43	0.0185	2.3975
Within populations	2.33	0.3889	50.3480
Total	5.90	0.7724	
F_{ST}		0.4965	

Table 4. Analysis of molecular variance (AMOVA) for the whitefly *Bemisia tabaci* in different populations.

Source of variation	Sum of squares	Variation components	Variation percentage
Between groups	2.65	0.6888	62.3839
Between populations within groups	0.92	0.0265	2.3959
Within populations	2.33	0.3889	35.2201
Total	5.90	1.1042	
F_{ST}		0.6478	

Another hypothesis tested was population clustering by state (BA, SP, and MT). A 47% difference between groups was observed, that is, the geographic origin of the populations probably influences their differentiation.

Since the Campinas population always occupied the external branches in the dendrograms, an AMOVA was run that compared the Campinas population with all other populations. The variation percentage between groups was 62%, which indicates that there really is a great difference between Campinas and the other localities. Based on the results obtained in AMOVA, notable structuring can be observed between the populations analyzed (Tables 1, 2, 3, and 4). It can be seen that the greatest part of the variation was observed between individuals from populations collected in different regions.

The haplotype network defined the cluster comparisons that were investigated in search for population structuring and gene flow (Templeton et al., 1995). Gene flow depends on the migration rate between populations. Migration rate, in turn, depends on the number of

populations that are within a region, their geographic distribution, spatial pattern, distance between them, and dispersal ability of the species (Frankhan et al., 2002).

Gene flow between groups homogenizes the genetic variation between them (Genovart et al., 2003). When gene flow rate is low or null, drift effects can become more easily visible in the differentiation of populations (Slatkin, 1987).

The inferences used to make comparisons between haplotypes consider that identical haplotypes have the same origin and that differences between haplotypes are due to random mutation events.

The genetic structuring, i.e., the non-random association of alleles or haplotypes, could be related to a restricted gene flow, historical events (fragmentation, geographic expansion, colonization) or to a combination of these factors (Templeton et al., 1995).

According to the gene flow model restricted via isolation by distance, it is expected that the most recent haplotype tends to be the least frequent and least dispersed (Templeton et al., 1995; De Brito et al., 2002). However, since each haplotype had only one copy, the frequency criterion cannot be used; only the number of differences observed between them can be used as a criterion. In general, the haplotypes show some mutations between one another (Figure 5). Just like in the dendrograms, the most different haplotypes, which belonged to the Campinas population (P410 and P412), occupied an external position in the tree. The most distant haplotype so far was biQ, which indicates that it was a different form between them, with possible differentiation by geographic isolation.
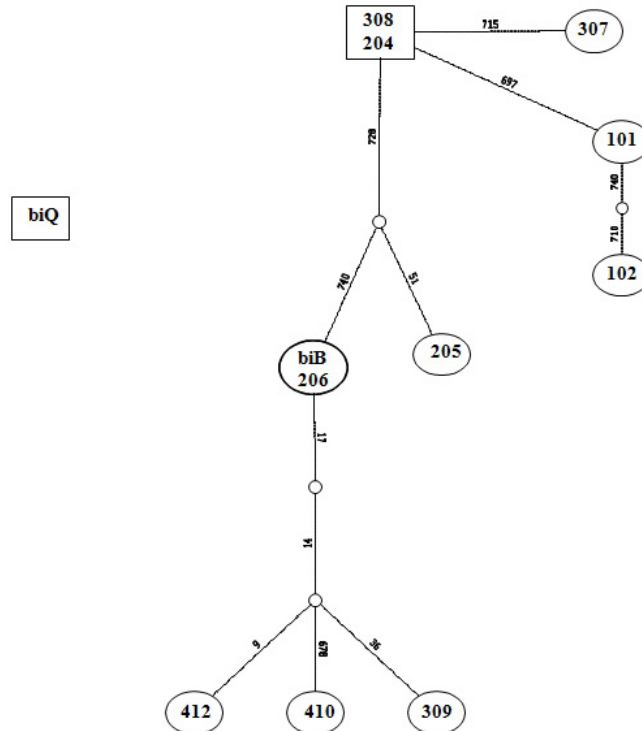

Figure 5. Haplotype network of *Bemisia tabaci* whitefly populations.

In spite of the differences found between some pairs of haplotypes, the low level of divergence observed between sequences does not corroborate the hypothesis that more than one *B. tabaci* biotype exist among the populations studied in the present article, which represent important agricultural areas of Brazil. Apparently, the host plant species colonized by the populations do not influence their structuring. Thus, differentiation exists between the populations sampled in relation to their geographic location. Frohlich et al. (1999) analyzed the mitochondrial DNA of *B. tabaci* individuals from several countries and also found a separation in *B. tabaci* due to its geographic origin. Therefore, the genetic differentiation between *B. tabaci* populations seems to occur by gene flow restriction between them, not by adaptation to a given host plant species.

REFERENCES

Banks GK, Bedford ID, Beitia FJ, Rodriguez-Cerezo E, et al. (1999). A novel geminivirus of *Ipomoea indica* (Convolvulaceae) from Southern Spain. *Plant Dis.* 83: 486.

Bedford ID, Briddon RW, Brown JK, Rosell RC, et al. (1994). Geminivirus transmission and biological characterization of *Bemisia tabaci* (Gennadius) biotypes from different geographic regions. *Ann. Appl. Biol.* 125: 311-325.

Bellows TS, Perring TM, Gill RJ and Headrick DH (1994). Description of a species of *Bemisia* (Homoptera: Aleyrodidae). *Ann. Entomol. Soc. Am.* 87: 195-206.

Berlinger MJ (1986). Host plant resistance to *Bemisia tabaci*. *Agric. Ecosyst. Environ.* 17: 69-82.

Berry SD, Fondong VN, Rey C, Rogan D, et al. (2004). Molecular evidence for five distinct *Bemisia tabaci* (Homoptera: Aleyrodidae) geographic haplotypes associated with cassava plants in Sub-Saharan Africa. *Ann. Entomol. Soc. Am.* 97: 852-859.

Boykin LM, Shatters RG Jr, Rosell RC, McKenzie CL, 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.

Brown JK, Frohlich DR and Rosell RC (1995a). The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or a species complex? *Ann. Rev. Entomol.* 40: 511-534.

Brown JK, Coats SA, Bedford ID, Markham PG, et al. (1995b). Characterization and distribution of esterase electromorphs in the whitefly, *Bemisia tabaci* (Genn.) (Homoptera: Aleyrodidae). *Biochem. Genet.* 33: 205-214.

Bryant D, Galtier N and Poursat M (2005). Likelihood Calculation in Molecular Phylogenetics. In: Mathematics of Evolution and Phylogeny (Gascuel O, ed.). Oxford University Press, Oxford, 33-62.

De Barro PJ, Driver F, Trueman JWH and Curran J (2000). Phylogenic relationship of word populations of *Bemisia tabaci* (Gennadius) using ribosomal ITS1. *Mol. Phylog. Evol.* 16: 29-36.

De Barro PJ, Trueman JW and Frohlich DR (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 Brito RA, Manfrin MH and Sene FM (2002). Mitochondrial DNA phylogeography of Brazilian populations of *Drosophila buzzatii*. *Gen. Mol. Biol.* 25: 161-171.

Dinsdale A, Cook L, Riginos C, Buckley YM, et al. (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.

Felsenstein J (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* 17: 368-376.

Felsenstein J (2004). Inferring Phylogenies. Sinauer, Sunderland.

Fitch WM (1981). A non-sequential method for constructing trees and hierarchical classifications. *J. Mol. Evol.* 18: 30-37.

Frankhan R, Ballou JD and Briscoe DA (2002). Introduction to Conservation Genetics. Cambridge University Press, Cambridge.

Frohlich DR, Torres-Jerez I, Bedford ID, Markham PG, et al. (1999). A phylogeographical analysis of the *Bemisia tabaci* species complex based on mitochondrial DNA markers. *Mol. Ecol.* 8: 1683-1691.

Genovart M, Oro D and Bonhomme F (2003). Genetic and morphological differentiation between the two largest breeding colonies of Audouin's Gull *Larus audouinii*. *Ibis* 145: 448-456.

Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis. Available at [http://www.mbio.ncsu.edu/BioEdit/bioedit.html.]. Accessed May 25, 2005.

Inbar M and Gerling D (2008). Plant-mediated interactions between whiteflies, herbivores, and natural enemies. *Annu.*

Rev. Entomol. 53: 431-448.

Jones DR (2003). Plant viruses transmitted by whiteflies. *Eur. J. Plant Pathol.* 109: -195.

Lima LHC, Návia D, Inglis PW and de Oliveira MRV (2000). Survey of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) biotypes in Brazil using RAPD markers. *Gen. Mol. Biol.* 23: 781-785.

Lourenção AL and Nagai H (1994). Surtos populacionais de *Bemisia tabaci* no Estado de São Paulo. *Bragantia* 53: -53.

Michener CD and Sokal RR (1957). A quantitative approach to a problem in classification. *Evolution* 11: 130-162.

Mound LA (1963). Host-correlated variation in *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae). *Proc. Royal Entomol. Soc. London (A).* 38: 171-180.

Norman JW, Riley DG, Stansly PA, Ellsworth PC, et al (2011). Management of Silverleaf Whitefly: A Comprehensive Manual on the Biology, Economic Impact and Control Tactics. USDA, Washington.

Perring TM (2001). The *Bemisia tabaci* species complex. *Crop Prot.* 20: 725-737.

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

Schneider S, Roessli D and Excoffier L (2000). Arlequin ver. 2000. A Software for Population Genetic Data Analysis. Genetics and Biometry Laboratory. University of Geneva, Geneva.

Slatkin M (1987). Gene flow and the geographic structure of natural populations. *Science* 236: 787-792.

Templeton AR, Routman E and Phillips CA (1995). Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140: 767-782.