

Reliable molecular identification of nine tropical whitefly species

Tatiana M. Ovalle¹, Soroush Parsa^{1,2}, Maria P. Hernández¹ & Luis A. Becerra Lopez-Lavalle^{1,2}

¹Centro Internacional de Agricultura Tropical (CIAT), Km 17, Recta Cali-Palmira, Cali, Colombia

²CGIAR Research Program for Root Tubers and Bananas, Lima, Peru

Keywords

COI, RFLP-PCR, Tropical whiteflies, Molecular identification.

Correspondence

Luis A. Becerra Lopez-Lavalle, Centro Internacional de Agricultura Tropical (CIAT) Km 17, Recta Cali-Palmira, Cali, Colombia. Tel: +57 2445 0000; Fax: +57 2445 0073; E-mail: l.a.becerra@cgiar.org

Funding Information

This should state that the CGIAR Research Program for Root Tubers and Bananas provided the resources to do this work.

Received: 4 April 2014; Revised: 20 June 2014; Accepted: 2 July 2014

Ecology and Evolution 2014; 4(19): 3778–3787

doi: 10.1002/ece3.1204

Introduction

Whiteflies (Hemiptera: Aleyrodidae) are globally considered as one of the most important pests in agriculture (Bellotti and Arias 2001; Morales and Anderson 2001). They can attack a wide range of crop hosts often reducing their yields by more than 50% due to the combined effects of their feeding from phloem and their vectoring of plant viruses (Byrne and Bellows 1991). Most species of whiteflies are oligophagous or polyphagous; however, a few species are monophagous (feed on only one host plant; Resh and Cardé 2009). The most notorious whitefly species, *Bemisia tabaci*, can feed from over 500 different plant hosts worldwide (Greathead 1986).

Today, 1556 valid species are known to compose the Aleyrodidae family, the only included family in the Aleyrodidae subgroup of Hemiptera (Forero 2008). The Aleyrodidae is further subdivided into three subfamilies: Aleurodicinae, Aleyrodinae, and Udamoselinae (Forero 2008). The systematics of these subfamilies is based almost entirely on morphological characters of the forth-instar nymph (i.e., pupae), the most conspicuous developmental stage (Martin et al.

Abstract

The identification of whitefly species in adult stage is problematic. Morphological differentiation of pupae is one of the better methods for determining identity of species, but it may vary depending on the host plant on which they develop which can lead to misidentifications and erroneous naming of new species. Polymerase chain reaction (PCR) fragment amplified from the mitochondrial cytochrome oxidase I (*COI*) gene is often used for mitochondrial haplotype identification that can be associated with specific species. Our objective was to compare morphometric traits against DNA barcode sequences to develop and implement a diagnostic molecular kit based on a RFLP-PCR method using the *COI* gene for the rapid identification of whiteflies. This study will allow for the rapid diagnosis of the diverse community of whiteflies attacking plants of economic interest in Colombia. It also provides access to the *COI* sequence that can be used to develop predator conservation techniques by establishing which predators have a trophic linkage with the focal whitefly pest species.

2000). Whitefly pupae, however, can exhibit significant phenotypic plasticity in response to differences in leaf architecture and to environmental or physical factors (Guershon and Gerling 2001), potentially confounding research outcomes. On the other hand, adult characters have been used successfully in the subfamily Aleurodicinae, but much work is still needed before they can be more broadly used in whitefly systematics (Ghahari et al. 2009). Modern molecular techniques can assist whitefly systematics research also facilitating species identifications (Oliveira et al. 2000; Calvert et al. 2001, 2005; Shatters et al. 2009).

Mitochondrial DNA (mtDNA) has been extensively used in phylogenetic studies of animals because it evolves much more rapidly than nuclear DNA, favoring the accumulation of nucleotide differences (i.e., polymorphisms) between closely related species (Brown et al. 1979; Lunt et al. 1996). The accumulation of these polymorphisms is primarily caused by the loss or gain of restriction sites without a detectable change in genome size (Hebert et al. 2003a).

The nature of the sequence changes within the mitochondrial cytochrome oxidase I (*COI*) gene make it an

ideal candidate to be used as a DNA barcode system (Hebert *et al.* 2003b; Kress and Erickson 2008). Generally, sequencing of the polymerase chain reaction (PCR) fragment amplified from the *COI* gene is used for mitochondrial haplotype identification that can be then associated with specific species (Shatters *et al.* 2009). This approach can be particularly useful when rapid whitefly detection and identification is required for regulatory purposes especially when specimens are physically damaged. To be most valuable, the test needs to be highly accurate while maintaining high rates of throughput.

A basic approach to develop such type test has been to employ universal PCR primers to amplify the *COI* gene and to subject the products to restriction fragment length polymorphism (RFPL) analysis (Vidigal *et al.* 2002; Caldeira *et al.* 2003; Thyssen *et al.* 2005). In addition, a large number of species-specific *COI* sequences (nearly 150,000 species) have now been submitted to public sequence databases such as GenBank or The Barcode of Life Data System (BOLD) for use in taxonomic and phylogenetic studies of insects (Kwong *et al.* 2012; Ptaszyńska *et al.* 2012; Smith *et al.* 2012).

Our objective was to compare morphometric traits against DNA barcode sequences to develop and implement a diagnostic molecular kit based on a RFLP-PCR method using the *COI* gene for the rapid identification of whiteflies. This study will allow for the rapid diagnosis of the diverse community of whiteflies attacking plants of economic interest in Colombia. It will also provide access to the *COI* sequence that can be used to development predator conservation techniques. For instance, Lundgren *et al.* (2013) were able to establish which predators have a trophic linkage with *Aleurotrachelus socialis* in Colombia.

Materials and Methods

Morphological study

Over 40 years, CIAT's entomology research group has collected or received hundreds of whitefly voucher specimens from tropical and subtropical regions of the world, where beans, cassava, rice, tropical forages and fruits crops are grown by small holder farmers. These specimens deposited in the Arthropod Reference Collection at CIAT (CIAT-ARC) headquarters in Cali, Colombia constitute an invaluable resource for the identification and characterization of current, as well as potentially new tropical crop pest. Initially, pupae of whitefly were collected from multiple crops in the area of Palmira – Valle del Cauca (Table 1). Pupa at the 4th instar nymph stage were morphologically examined under a light microscope and identified using the keys of Martin *et al.* (2000), Martin (2004), Hodges and Evans (2005), Martin (2005), and Dooley (Dooley 2006).

Observations were carried out from January of 2011 to August of 2012 on CIAT's fields and laboratories located in Palmira, Valle del Cauca, Colombia. Pupae were collected from infested leaves of seven plant species and brought to the laboratory. Nine whitefly species were identified and 10–20 nymphs of each species (Table 1).

The pupae samples were then placed vials and allowed to develop to adults. Once emerged, 10–20 adults per species were introduced in 1.5-mL microcentrifuge tube and placed in liquid nitrogen for DNA isolation.

DNA isolation and PCR amplification

Total genomic DNA was obtained from nine whiteflies species using a modified cetyltrimethylammonium bromide method (CTAB; Doyle and Doyle 1990) by adding 2.5 M potassium acetate (pH 5.5) to precipitate protein. The DNA barcode region was amplified from all whitefly species using 10 μ M of universal DNA primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO 2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') for the amplification of mitochondrial cytochrome c oxidase subunit I – *COI* (Folmer *et al.* 1994). PCR was conducted in a Mastercycler[®] pro (Eppendorf, Hamburg, Germany) for five cycles (40 sec at 94°C, 40 sec at 45°C, and 60 sec at 72°C) followed by 35 cycles (40 sec at 94°C, 40 sec at 51°C, and 60 sec at 72°C) with a final extension period of 72°C for 10 min. All PCRs were performed in a 12 μ L reaction volume using 0.25 U of Platinum[®] Taq (Invitrogen[®] Carlsbad, CA), 1 \times PCR Buffer (buffer composition), 2.5 mM of each dNTP (Promega Corp. Madison, WI), 2.5 mM MgCl₂, and 20 ng of DNA.

Sequencing of the *COI* region

Amplified *COI* region from each white fly species was cloned into the PGEM[®]-T Easy vector (Promega Corp.). Plasmid DNA from *Escherichia coli* used for sequence analysis was purified using a SV minipreps preparation (Promega Corp.). DNA inserts in this vector were sequenced in both directions using the Big Dye[™] Terminator Cycle Sequencing kit with an Applied Biosystems 377 (Foster City, CA) DNA fragment analyzer by the Cornell University Life Sciences Core Laboratories Center (Ithaca, NY), using vector primers SP6 and T7. The resulting DNA sequences were edited and analyzed using Sequencher[®] 4.5 (Genes Codes Corp. Ann Arbor, MI).

DNA sequence analysis

DNA Sequences were edited to the whitefly species-specific *COI* sequence by eliminating vector and universal primer sequence and analyzed for identity with sequences

Table 1. List of whitefly species found associated with seven plant species in Palmira/Valle del Cauca (Colombia) based on morphological and anatomical features of the puparia.

Host plant ¹	WF species identification ²	Key used
<i>Citrus sinensis</i>	<i>Aleurothrix floccosus</i>	Puparia dark brown to black; margin with coarse teeth, a gland is present at the base of each tooth, which gives margin the appearance of having a double row of teeth. Vasiform orifice elevated, small lingula obscured by operculum. Caudal furrow absent. Puparia often in dense groups which are covered by secreted flocculent wax (Martin 1987)
<i>Ipomoea indica</i>	<i>Aleurotrachelus trachoides</i>	Puparia black, marginal row of teeth much paler than the rest of pupal case, each tooth with an even paler glandular spot at its base. Inner submargin, mesad of marginal teeth and glands, with regular rows of coarse black. This dotted zone is divided into blocks by narrow breaks perpendicular to the margin; lingula large setose, expanded apically, protruding beyond vasiform orifice (Martin 1987)
<i>Manihot esculenta</i>	<i>Aleurotrachelus socialis</i>	Puparia dark usually black, elliptical, margin dentate, not differentiated at caudal or thoracic marginal openings. Glands present at the base of the marginal teeth giving the appearance of a double row of teeth. Small dorsal patches of tiny spinules, and rather stout eight abdominal setae that are longer than the caudal setae. Vasiform orifice subcircular to subchordate, longer than wide with an exerted lingula or it is obscured by operculum (Martin 2005)
<i>Manihot esculenta</i>	<i>Trialeurodes variabilis</i>	Body oval; submarginal papillae in an irregularly spaced single row; dorsal disc pores distad or interspersed among submarginal papillae; subdorsal tubercles absent; lingula head distinctly lobulated; first abdominal and eighth abdominal setae short (KeyLucid Central for Whitefly) ³
<i>Matisia cordata</i>	<i>Lecanoideus floccissimus</i>	Submarginal zone without a ring of double-rimmed pores, only with a band of crowded wide-rimmed pores; dorsal disc mesad of compound pores only sparsely punctuated by septate pores (Martin et al. 2000)
<i>Passiflora edulis</i>	<i>Aleuronudus melzeri</i>	Cephalic and the three anterior abdominal compound pores subequal in size; 3 posterior abdominal compound pores reduced in size with the 5th posterior abdominal pore offset. Cephalic submedian setae absent. Lingula exerted (Cortez-Madrigal et al. 2008)
<i>Phaseolus vulgaris</i>	<i>Bemisia tabaci</i>	Puparia ovoid. Disc dorsal with or without setae. Caudal setae always stout, usually as long as the vasiform orifice whose sides are almost straight. Vasiform orifice triangular, inset from puparial margin by less than its own length. <i>Lingula spatulate</i> (Martin et al. 2000)
<i>Phaseolus vulgaris</i>	<i>Trialeurodes vaporariorum</i>	Body oval usually pale yellow; margin crenulations coarser, usually less than 13 occupying 0.1 mm; with a tiny tongue-like structure usually visible protruding beyond apical notch of vasiform orifice. Eight abdominal setae placed anterior to widest part of operculum. Legs lacking spines, only with setae (Martin et al. 2000)
<i>Psidium guajava</i>	<i>Aleurodicus dispersus</i>	Puparia relatively large (≥ 1.00 mm). Compound pores with cephalic pair similar in size to anterior most two abdominal pairs; each pore usually with a central process visible outer submarginal zone with a distinct ring of double-rimmed pores; dorsal disc mesad of compound pores densely punctuated by septate pores (Martin 2004; Martin et al. 2000)

¹WF infested plant materials were sampled at CIAT headquarters, Palmira, Colombia.

²Determined by Maria del Pilar Hernandez (CIAT).

³<http://keys.lucidcentral.org/keys/v3/whitefly/Old/Homepage.htm>.

deposit in the GenBank, EMBL or BOLD public databases. *COI* sequence contigs were assembled from five clones per specie to identify unique haplotypes with a discrimination requirement of 100% identity. Haplotype nucleotide sequences of the whitefly samples are deposited in GenBank (Table 2). Genetic distances were calculated with the alignment sequences for the nine species. Further, we used sequences reported in the GenBank for species listed in Table 1 as control group, when available (*Trialeurodes vaporariorum* and *B. tabaci*), and *Dialeurodes citri* (JQ340193), *Dialeurodes hongkongensis* (JQ340195) and *Dialeuropora* (JQ340197) as outgroup. Two phylogenetic reconstruction analyses were performed using the MEGA V6 and CLC main workbench V6.9 software. The first phylogeny analysis was performed using the statistical method neighbor-joining under

Jukes-Cantor (J-C) Poisson-model for base substitution implemented in the CLC main workbench V 6.9 software (Jukes and Cantor 1969). The second analysis was also performed using the standard neighbor-joining (NJ) using the JTT matrix-based method (Jones et al. 1992).

Restriction mapping for rapid whitefly identification

Each of the 10 *COI* sequences obtained in this study was digested in silico with 30 commercially available restriction enzymes: eight base cutter *NotI*; six base cutters *ApaI*, *AseI*, *DraI*, *EcoRI*, *EcoRV*, *HindIII*, *KpnI*, *NdeI*, *PstI*, *PvuII*, *SacI*, *Sall*, *ScaI*, *SmaI*, *SpeI*, *SphI*, *XbaI*, and *XhoI*; five base cutters *DdeI*, *HinfI*, and the four base cutters *AluI*, *HaeIII*, *HhaI*, *HpaII*, *MboI*, *MseI*, *MspI*, *RsaI*, and

Table 2. "In silico" characterization of the *COI* restriction sites using single enzyme reactions with *AluI* and *MboI* or a double digest reaction with *AseI/TaqI* in nine species of whiteflies.

WF ¹ species	<i>AluI</i>		<i>MboI</i>		<i>AseI + TaqI</i>	
	No. frag.(s)	Length(s)	No. frag.(s)	Length(s)	No. frag.(s)	Length(s)
<i>Aleurodicus dispersus</i>	5	336, 174, 118, 69 & 12	2–4	615 ² , 611, 98, & 94	5–7	207, 172, 151, 95, 84 & 12 ²
<i>Aleuronudus melzeri</i>	4	330, 243, 124 & 12	7	328, 128, 78, 77, 53, 27 & 18	6	202, 172, 119, 95, 89 & 32
<i>Aleurothrix floccosus</i>	3–5	Undigested (709) ² , 591 536 ² , 118 & 55 ²	6	294, 128, 112, 104, 53 & 18	2	649 & 60
<i>Aleurotrachelus socialis</i>	3–6	611 ² , 454, 330 ² , 157 & 98	3–5	615 ² , 611, 53, 49 ² & 45	3	554, 95 & 60
<i>Aleurotrachelus trachoides</i>	2–3	Undigested (709) ² , 575 & 134	4	534, 77, 53 & 45	3	391, 263 & 55
<i>Bemisia tabaci</i> (A) ³	3–7	591 ² , 454 ² , 336, 255, 118, 82 ² , & 36 ²	5	483, 128, 71, 16 & 11	5	483, 95, 60, 50 & 21
<i>Bemisia tabaci</i> (B) ³	4	367, 157, 98 & 87	6	534, 77, 53, 18, 16 & 11	5	354, 129, 95, 71 & 60
<i>Lecanoides floccissimus</i>	5–7	522 ² , 454 ² , 336, 243 ² , 174, 118, 69 & 12	6–11	410 ² , 406, 132 ² , 128, 81 ² , 77, 53, 31 ² , 27 22 ² & 18	8	241, 133, 122, 95, 84, 21, 8 & 5
<i>Trialeurodes vaporariorum</i>	1	Undigested (709)	5	456, 78, 77, 53 & 45	4	303, 174, 143 & 89
<i>Trialeurodes variabilis</i>	5–7	322 ² , 255 ² , 191, 134, 132, 131, 121 & 47	6	212, 199, 134, 71, 66 & 27	5	255, 169, 130, 95 & 60

¹Whitefly.²Intraspecific variation explaining WF population differences.³Samples associated with *Phaseolus vulgaris*.

TaqI, using the "Restriction site analysis tool" of CLC main workbench 6.8.3 (CLC Bio, Aarhus Denmark), to predict the product sizes of each *COI*-amplicon restriction reaction to implement a cost-effective *COI*-RFLP based species identification assay.

COI-RFLP assay

Restriction endonuclease digestion of amplified *COI* fragments was performed with one unit of restriction endonucleases *AluI*, *MboI*, and *TaqI* (NEB, Beverly, MA), 1× supplied restriction buffer and sterile Milli-Q H₂O in a final volume of 20 μL. Reactions were incubated at 37°C for 2 h. Restriction digestion products were resolved in 2% agarose gels in boric acid (BA) buffer at 80 V for 3 h, stained with GelRed™ (Biotium, Hayward, CA). Gel images were capture with a GelDoc™ BioRad documentation system and analyzed with the Image LAB™ software (BioRad, Hercules, CA).

Results

Molecular species identification (DNA barcoding)

Morphological characterization of the adult whitefly is very difficult and limits the capacity of identifying new

species. Almost universally, whitefly taxonomical characterization is performed at the nymphal stage, which is labor intensive and time consuming. Because of the limitations for a rapid identification of whitefly species by conventional means, a molecular approach was adopted to assess its use for rapid identification of whitefly adult samples. Initially, the mtDNA *COI* universal primers LCO1490 and HCO2198 were employed to amplify the *COI* region of nine whitefly species. The *COI* gene (COX1) is one of 13 coding protein genes in the mitochondrial genome. It is ordinarily composed by 512 amino acids flanked on its 5' terminus by tRNA-Trp-Tyr-Cys (15,121–15,315 bp) and at its 3' terminus by tRNA-Leu (1538–1603 bp). The LCO1490 primer sequence is located 18 bp downstream near the 3'-end of the tRNA-Cys, and the HCO2198 primer sequence is located on the opposite DNA strand 812 bp upstream at the 5'-end of the tRNA_{Leu}. Used together, these primers amplify a partial COX1 fragments commonly referred as the *COI* sequence or the Barcode of Life (Hajibabaei et al. 2007).

As identified morphologically, all nine whitefly species successfully amplified the *COI* sequence. The PCR amplification generated a fragment of approximately 709 bp that was subsequently cloned and sequence as described above. One hundred and fifty clones were recovered, but 54 were carrying the right size insert (approximately 5 per sample). DNA sequences obtained comprised 219 of

the 512 amino acids that define the COX1 gene of the mitochondrial genome for all ten whitefly samples that represented the nine species under investigation (Table 1). The sequences were employed to search the GenBank, EMBL, and BOLD databases to find novel sequences with the greatest identity. The information gathered from these three databases allowed the assignment of each whitefly sample to a single species (Table 2).

DNA sequences of the partial COI region obtained in this study were subjected to phylogenetic analysis, along with seven related whitefly COI sequences retrieved from

the GenBank/EMBL/BOLD databases, whereas three sequences *D. citri* (JQ340193), *D. hongkongensis* (JQ340195), and *Dialeuropora* spp. (JQ340197) were employed as out-group. All 20 COI DNA sequences were translated to amino acid sequences and a matrix with pairwise distance values was used to obtain a phylogenetic tree with neighbor-joining (Fig. 1), and the agreed tree provided strong support for the morphologically determined whitefly classification (Table 1). Trees obtained from MEGA and CLC main workbench gave essentially the same results. The COI amino acid sequence and

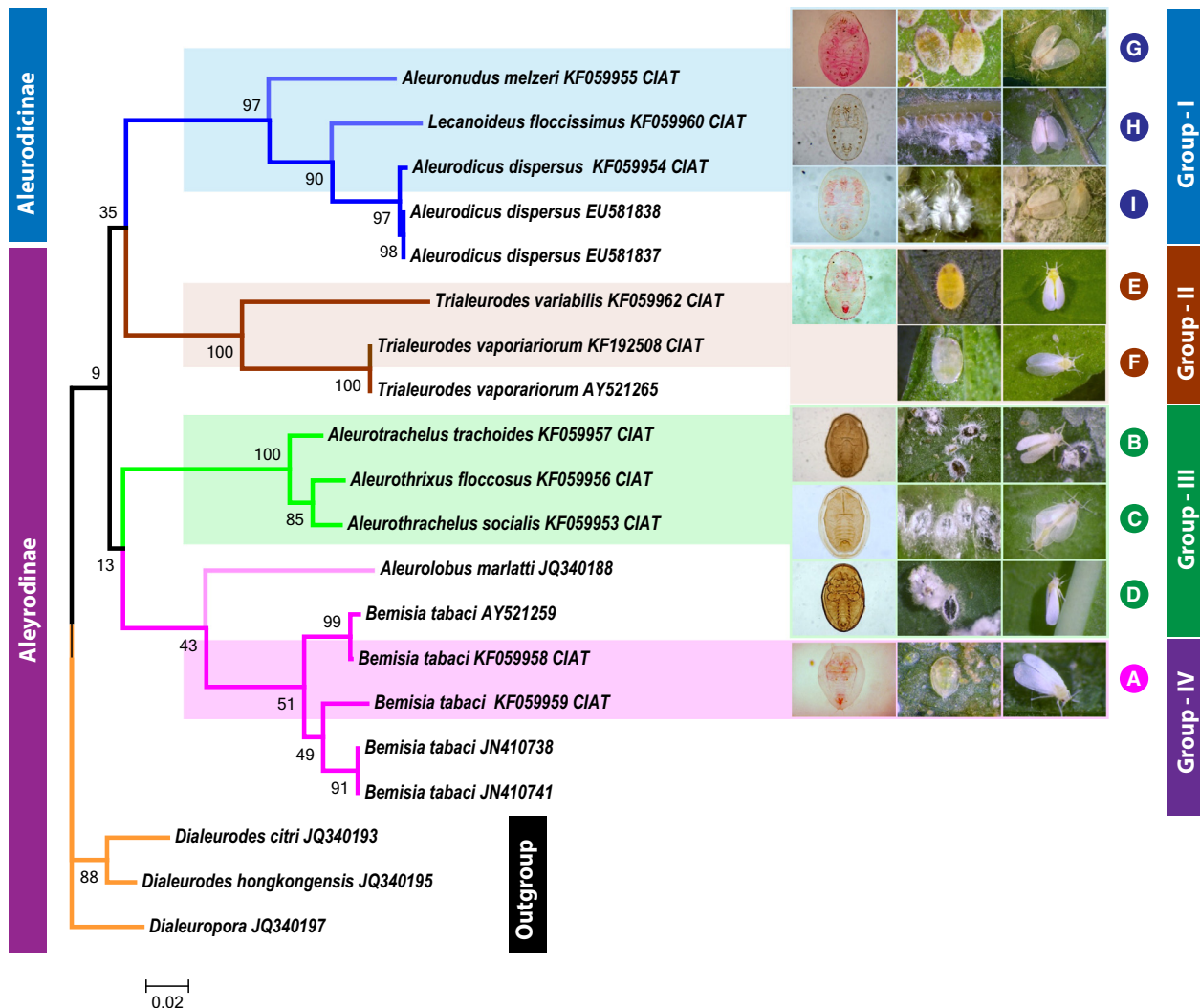


Figure 1. Phylogenetic relationships inferred using the neighbor-joining method in MEGA V6 and CLC main workbench V6.9 for tropical whitefly species, in relation to well-characterized whitefly taxa for which sequences were available in the GeneBank (Saitou and Nei 1987; Tamura et al. 2011). The optimal tree with the sum of branch length = 1.67299211 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10,000 replicates) are shown next to the branches (Felsenstein 1985). The analysis involved 29 amino acid sequences. There were a total of 219 positions in the final dataset. Four well-defined whitefly phylogenetic groups (G-I to G-IV) were observed. The nine species identified at CIAT were named from A to I. A = *Aleoronudus melzeri*, B = *Lecanoideus floccissimus*, C = *Aleurodicus dispersus*, D = *Trialeurodes variabilis*, E = *Trialeurodes vaporariorum*, F = *Aleurotrachelus trachoides*, G = *Aleurothrixus floccosus*, H = *A. trachoides*, and I = *Bemisia tabaci*.

phylogenetic analyses clustered all species in four groups. Group-I *Aleuronudus melzeri*, *Lecanoideus floccissimus*, and *Aleuronudus disperses*, group-II *Trialeurodes variabilis* and *T. vaporariorum*, group-III clustered *Aleurotrachelus trachoides*, *A. socialis*, *Aleurotrachelus floccosus*, and in group-IV *B. tabaci* together with *Aleurotrachelus marlattii*. Of the four groups, group-I belonged to the subfamily Aleurodicinae, while groups II–IV including the outgroup belonged to the subfamily Aleyrodinae.

COI-RFLP analysis

For the COI-RFLP analysis, COI fragments from the nine whitefly species were digested in silico with 30 restriction

enzymes. Digestions with *AluI* and *MboI* proved to be the most informative, whereas *AseI* and *DdeI* are fully informative only if used together or in association with *TaqI* (Table 2). The polymorphic restriction patterns revealed by *AluI* or *MboI* in silico showed that for the 10 whitefly samples used in this study one can easily and inexpensively develop a diagnostic test for species identification.

The COI-PCR amplification yield both on genomic as well as cloned materials yielded fragments of 709 bp. After restriction enzyme digestion, all samples displayed reproducible restriction enzyme pattern. As predicted in silico, the profile produced by *AluI* or *MboI* allowed us to distinguish among the nine whitefly species. *AluI* showed the best profile for differentiating all ten samples

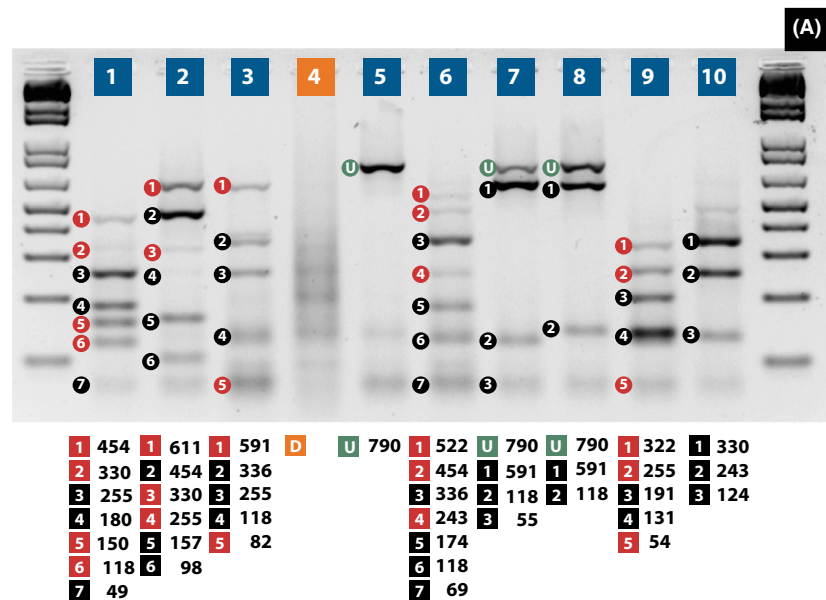
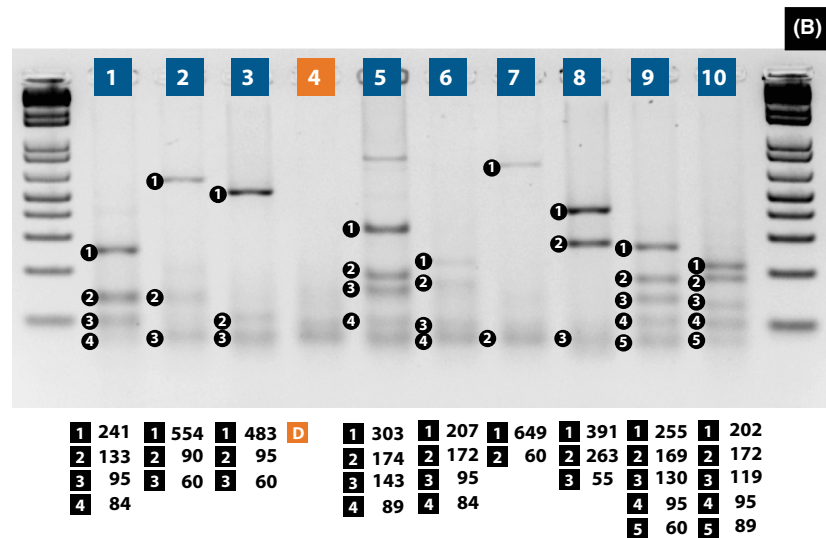


Figure 2. Displays two 2% agarose gels showing the restriction fragment length polymorphism profiles obtained by digesting the COI with *AluI* (A) and *AseI* + *TaqI* (B) enzyme from *Lecanoideus floccissimus* (lane 1), *Aleurotrachelus socialis* (lane 2), *Bemisia tabaci* (Sample 1, lane 3), *B. tabaci* (Sample 2, lane 4), *Trialeurodes vaporariorum* (lane 5), *Aleurodicus dispersus* (lane 6), *Aleurothrixus floccosus* (lane 7), *Aleurotrachelus trachoides* (lane 8), *Trialeurodes variabilis* (lane 9), *Aleuronudus melzeri* (lane 10). Molecular size markers are shown on the right and left side of the figure. Number filled in black corresponds to fragments size expected based on available sequence information and those in red to fragments expected by sequence variation due to intraspecies variation detected by RFLP-PCR of the COI amplicon.



(Fig. 2A). *MboI* enzyme also allowed the identification of all samples; however, *A. trachoides* and *B. tabaci* DNA restriction profile were very similar (data not shown). The double enzyme identification system (*AseI/TaqI*) also allowed distinguishing among all ten samples used here. The restriction patterns for some species were complex, and the sum of sizes of the bands produced upon digestion was not equal to the size of the undigested PCR products. For instance, in the patterns produced by *AluI* for *L. floccissimus* (Fig. 2A) the sum of the sizes of seven bands (454 + 330 + 255 + 180 + 150 + 118 + 49 bp) was not equal to the estimated 709 bp. It might be due to sequence variability of the *COI* gene caused by genetic differentiation among individual whiteflies originated from different populations. This is a very plausible explanation as DNA extractions were performed by randomly pooling the 10–20 adult whiteflies. However, the sensitivity for identifying genetic differences within the *COI* sequence disappeared when a double enzyme digestions were performed using *AseI* + *TaqI*. In the patterns produced by these two enzymes for *B. tabaci* the sum of three bands (483 + 95 + 60) was not equal to the estimated 709. It might be due to overlapping of similar size fragments produced by the enzyme.

Discussion

In the present study, we conducted classical morphological identification of the whitefly's 4th instar larvae and molecular identification of the adult form using the part of the COX1 gene region of the mtDNA. Based on the morphological traits, we identified nine species of whiteflies associated with seven crop species at CIAT's headquarters (Table 1). This approach is limited because it demands specialized taxonomic expertise and it relies on a life-stage that may display significant environmentally induced morphological variation, confounding identifications (Neal and Bentz 1999).

To circumvent these limitations, we amplified Folmer's COX1 region (i.e., *COI*) to match DNA samples with morphologically identified whitefly species.

Before undertaken a phylogenetic inspection of the whitefly species sampled, *COI* sequences were further examined to characterize for the presence of orthologues or *COI*-like sequences by looking at the translated amino acid sequence. Moulton et al. (2010) suggested incorporating to the barcode methodology the analysis of the derived amino acid sequences particularly when a protein-coding gene region like *COI* is used. The translation of each *COI* sequence examined here consistently produced a fragment of 219 amino acids. Our aim to investigate the applicability of DNA barcoding for confirming species identifications of whiteflies was only limited when

insufficient numbers of worldwide whitefly entries were available on GenBank and BOLD databases. This was predominantly the case for the South American region. Nevertheless, where whitefly *COI* sequence information was available, species identification was correctly and consistently achieved using GenBank or BOLD. Thus, DNA barcode identification was achieved for *B. tabaci* in cluster G-IV, *T. vaporariorum* in cluster G-II, and *Aleurodicus dispersus* in G-I.

Our results suggest that *COI* sequence analysis is a very accurate and effective tool for species identification in whiteflies. Thus, by comparing molecular markers with diagnostic morphological traits, we provide six new *COI* sequences that will contribute to build a species-specific sequence library. A *COI*-based identification system will undoubtedly provide high resolution tools to nontaxonomist for conducting field surveys to collect information about the numbers and distribution of whitefly species among crops and its associated weeds.

COI sequence analysis has proven to be a good molecular marker for intra- and interspecific variation in whitefly species not only in this study but others (Thao and Baumann 2004; Boykin et al. 2007; Shoorcheh et al. 2008; Chu et al. 2012; Henri et al. 2013). However, a simple, rapid, and cost-effective identification system for whitefly species surveillance cannot rely exclusively on sequencing. Invariably in this study, the *COI* universal primers (Folmer et al. 1994) amplified a 709 bp fragment suggesting that a DNA diagnostic approach without a sequencing step will require either restriction enzyme digestion of the PCR product (RFLP-PCR) or develop molecular markers that adequately distinguish intra- and interspecific variation among different whitefly species. PCR amplification of the *COI* region followed by digestion with two single restriction endonucleases (*AluI* or *MboI*) or a double enzyme system (*AseI* + *TaqI*) could allow a rapid identification of whitefly species at the adult stage (Table 2). This molecular method of identification can be easily used to monitor the spread of native whiteflies or for potential outbreaks of invasive exotic types in South America and elsewhere. As shown in Figure 2(A and B), the *COI*-RFLP assay successfully differentiated all adult forms of the nine whitefly species identified morphologically at CIAT. We performed digestions with *AluI* (Fig. 2A) and *AseI* + *TaqI* (Fig. 2B) of the *COI* sequences obtained using total genomic DNA of bulked individuals for each whitefly species. The restriction profiles for all whitefly were screened using 2% agarose gels as they provide good resolution and size discrimination. These results demonstrate that *COI*-RFLP, using *AluI* or *AseI* + *TaqI* enzymes, is an important tool for differentiation of whitefly species and its biotypes. Furthermore, restriction digestions with the *AluI* enzyme also allowed

the identification of intraspecific changes indicating that this approach can also be used to conduct survey or monitor whitefly population distribution.

In addition to facilitating their authoritative identification, molecular protocols can also help characterize predator communities feeding on whiteflies (Lundgren *et al.* 2013). Molecular gut content analyses can establish which predators have a trophic linkage with a focal whitefly species, assisting the development of biological control programs for its management. Using this method, Lundgren *et al.* (2013) characterized the predator community of cassava whitefly *A. socialis* in Colombia. The *COI* sequence information generated in this study has the potential of facilitating similar efforts for a broader list of pestiferous whiteflies globally.

Acknowledgments

The authors acknowledge the CGIAR Root, Tubers and Bananas Research program for the financial support to conduct this study. We also would like to thank Ms. Katherine Castillo for her invaluable advised in DNA cloning and sequence analysis.

Data accessibility

The *COI* sequence data used for this study can be found in the following links:

<https://www.ncbi.nlm.nih.gov/nuccore/KF059953>
<https://www.ncbi.nlm.nih.gov/nuccore/KF059954>
<https://www.ncbi.nlm.nih.gov/nuccore/KF059955>
<https://www.ncbi.nlm.nih.gov/nuccore/KF059956>
<https://www.ncbi.nlm.nih.gov/nuccore/KF059957>
<https://www.ncbi.nlm.nih.gov/nuccore/KF059958>
<https://www.ncbi.nlm.nih.gov/nuccore/KF059959>
<https://www.ncbi.nlm.nih.gov/nuccore/KF059960>
<https://www.ncbi.nlm.nih.gov/nuccore/KF059962>
<https://www.ncbi.nlm.nih.gov/nuccore/KF192508>
<https://www.ncbi.nlm.nih.gov/nuccore/EU581837>
<https://www.ncbi.nlm.nih.gov/nuccore/EU581838>
<https://www.ncbi.nlm.nih.gov/nuccore/JQ340188>
<https://www.ncbi.nlm.nih.gov/nuccore/JQ340193>
<https://www.ncbi.nlm.nih.gov/nuccore/JQ340195>
<https://www.ncbi.nlm.nih.gov/nuccore/JQ340197>
<https://www.ncbi.nlm.nih.gov/nuccore/JN410738>
<https://www.ncbi.nlm.nih.gov/nuccore/JN410741>
<https://www.ncbi.nlm.nih.gov/nuccore/AY521259>

Conflict of Interest

None declared.

References

- Bellotti, A. C., and B. Arias. 2001. Host plant resistance to whiteflies with emphasis on cassava as a case study. *Crop Prot.* 20:813–823.
- Boykin, L. M., R. G. Shatters Jr, R. C. Rosell, Cindy L. McKenzie, Ruth Ann. Bagnall, Paul. 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.
- Brown, W. M., M. George, and A. C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl Acad. Sci. USA* 76:1967–1971.
- Byrne, D. N., and T. S. Bellows. 1991. Whitefly biology. *Annu. Rev. Entomol.* 36:431–457.
- Caldeira, R. L., O. S. Carvalho, C. L. Mendonça, *et al.* 2003. Molecular differentiation of *Angiostrongylus costaricensis*, *A. cantonensis*, and *A. vasorum* by polymerase chain reaction-restriction fragment length polymorphism. *Mem. Inst. Oswaldo Cruz* 98:1039–1043.
- Calvert, L. A., M. Cuervo, J. A. Arroyave, *et al.* 2001. Morphological and mitochondrial DNA marker analyses of whiteflies (Homoptera: Aleyrodidae) colonizing cassava and beans in Colombia. *Ann. Entomol. Soc. Am.* 94:512–519.
- Calvert, L., N. Villarreal, and D. Frohlich. 2005. Using molecular techniques to analyse whitefly species and biotypes in Latin America. Pp. 251 *in* P. K. Anderson, F. J. Morales, eds. *Building a knowledge*. No. 341. CIAT, Cali, Colombia.
- Chu, D., X. Hu, C. Gao, *et al.* 2012. Use of mitochondrial cytochrome oxidase I polymerase chain reaction-restriction fragment length polymorphism for identifying subclades of *Bemisia tabaci* Mediterranean Group. *J. Econ. Entomol.* 105:242–251.
- Cortez-Madriral, H., M. Cruz Martinez-Lopez, F. Reyes-Izquierdo, and L. D. Ortega-Arenas. 2008. First record of *Lecanoideus floccissimus* (Hemiptera: Aleyrodidae) in cocoa from Tabasco, Mexico. *Rev. Colomb. Entomol.* 34:33–40.
- Dooley, J.W. 2006. Whitefly pupa of the World: compendium and key to the genera of the *Aleurodicinae* & the *Aleyrodinae*. USDA, APHIS, PPQ. South San Francisco, USA (<http://keys.lucidcentral.org/keys/v3/whitefly/Default.htm>).
- Doyle, J., and J. Doyle. 1990. A rapid total DNA preparation procedure for fresh plant tissue. *Focus* 12:13–15.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791.
- Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotech.* 3:294–299.
- Forero, D. 2008. The systematics of the Hemiptera. *Rev. Colomb. Entomol.* 34:1–21.

- Ghahari, H., S. Abd-Rabou, J. Zahradnik, and H. Ostovan. 2009. Annotated catalogue of whiteflies (Hemiptera: Sternorrhyncha: Aleyrodidae) from Arasbaran, Northwestern Iran. *J. Entomol. Nematol.* 1:007–018.
- Greathead, A. H. 1986. Host plants. Pp. 17–26 *in* M. J. W. Cock, ed. *Bemisia tabaci* – a literature survey: on the cotton whitefly with an annotated bibliography. CAB International Institute of Biological Control, Ascot, U.K.
- Guershon, M., and D. Gerling. 2001. Effect of foliar tomentosity on phenotypic plasticity in *Bemisia tabaci* (Hom., Aleyrodidae). *J. Appl. Entomol.* 125:449–453.
- Hajibabaei, M., G. A. C. Singer, P. D. N. Hebert, and D. A. Hickey. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends Genet.* 23:167–172.
- Hebert, P. D., S. Ratnasingham, and J. R. de Waard. 2003a. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B Biol. Sci.* 270:S96–S99.
- Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. de Waard. 2003b. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B Biol. Sci.* 270:313–321.
- Henri, H., G. Terraz, O. Gnankiné, F. Fleury, and L. Mouton. 2013. Molecular characterization of genetic diversity within the Africa/Middle East/Asia Minor and Sub-Saharan African groups of the *Bemisia tabaci* species complex. *Int. J. Pest Manage.* 59:329–338.
- Hodges, G. S., and G. A. Evans. 2005. An identification guide to the whiteflies (Hemiptera: Aleyrodidae) of southeastern united states. *Fla Entomol.* 88:518–534.
- Jones, D. T., W. R. Taylor, and J. M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8:275–282.
- Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pp. 21–132 *in* M. N. Munro, ed. *Mammalian protein metabolism*. Academic Press, New York, USA.
- Kress, W. J., and D. L. Erickson. 2008. DNA barcodes: genes, genomics, and bioinformatics. *Proc. Natl Acad. Sci. USA* 105:2761–2762.
- Kwong, S., A. Srivathsan, and R. Meier. 2012. An update on DNA barcoding: low species coverage and numerous unidentified sequences. *Cladistics* 28:639–644.
- Lundgren, J. G., L. A. Becerra López Lavalle, S. Parsa, and K. A. G. Wyckhuys. 2013. Molecular determination of the predator community of a cassava whitefly in Colombia: pest-specific primer development and field validation. *J. Pest. Sci.* 87:125–131.
- Lunt, D. H., D. X. Zhang, J. M. Szymura, and O. M. Hewlitt. 1996. The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. *Insect Mol. Biol.* 5:153–165.
- Martin, J. H. 1987. An identification guide to common whitefly pest species of the world (Homopt Aleyrodidae). *Trop. Pest Manage.* 33:298–322.
- Martin, J. 2004. Whiteflies of Belize (Hemiptera: Aleyrodidae). Part 1-introduction and account of the subfamily Aleurodicinae Quaintance & Baker. *Zootaxa* 1: 1–119.
- Martin, J. 2005. Whiteflies of Belize (Hemiptera: Aleyrodidae) Part 2-a review of the subfamily Aleyrodinae Westwood. *Zootaxa* 1:1–116.
- Martin, J., D. Mifsud, and C. Rapisarda. 2000. The whiteflies (Hemiptera: Aleyrodidae) of Europe and the Mediterranean basin. *Bull. Entomol. Res.* 90:407–448.
- Morales, F. J., and P. K. Anderson. 2001. The emergence and dissemination of whitefly-transmitted geminiviruses in Latin America. *Arch. Virol.* 146:415–441.
- Moulton, M. J., H. Song, and M. F. Whiting. 2010. Assessing the effects of primer specificity on eliminating numt coamplification in DNA barcoding: a case study from Orthoptera (Arthropoda: Insecta). *Mol. Ecol. Resour.* 10:615–627.
- Neal, J. W., and J.-A. Bentz. 1999. Evidence for the stage inducing phenotypic plasticity in Pupae of the Polyphagous Whiteflies *Trialeurodes vaporariorum* and *Bemisia argentifolii* (Homoptera: Aleyrodidae) and the raison d'être. *Ann. Entomol. Soc. Am.* 92:774–787.
- Oliveira, M. R. Vd., M. S. Tigano, and S. Aljanabi. 2000. Molecular characterization of whitefly (*Bemisia* spp.) in Brazil. *Pesqui. Agropecu. Bras.* 35:1261–1268.
- Ptaszyńska, A. A., J. Łętowski, S. Gnat, and W. Małek. 2012. Application of COI sequences in studies of phylogenetic relationships among 40 Apionidae species. *J. Insect Sci.* 12: 1–14.
- Resh, V. H., and R. T. Cardé. 2009. *Encyclopedia of insects*. Academic Press, New York, USA.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406–425.
- Shatters, R. G., C. A. Powell, L. M. Boykin, H. Liansheng, and C. L. McKenzie. 2009. Improved DNA barcoding method for *Bemisia tabaci* and related Aleyrodidae: development of universal and *Bemisia tabaci* biotype-specific mitochondrial cytochrome oxidase I polymerase chain reaction primers. *J. Econ. Entomol.* 102:750–758.
- Shoorcheh, H. R., B. Kazemi, S. Manzari, J. Brown, and A. Sarafrazi. 2008. Genetic variation and mtCOI phylogeny for *Bemisia tabaci* (Hemiptera, Aleyrodidae) indicate that the 'B' biotype predominates in Iran. *J. Pest. Sci.* 81:199–206.
- Smith, M. A., C. Bertrand, K. Crosby, et al. 2012. *Wolbachia* and DNA barcoding insects: patterns, potential, and problems. *PLoS One* 7:e36514.
- Tamura, K., D. Peterson, N. Peterson, et al. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28:2731–2739.
- Thao, M. L., and P. Baumann. 2004. Evolutionary relationships of primary prokaryotic endosymbionts of

- whiteflies and their hosts. *Appl. Environ. Microbiol.* 70:3401–3406.
- Thyssen, P. J., A. C. Lessinger, A. M. Azeredo-Espin, and A. X. Linhares. 2005. The value of PCR-RFLP molecular markers for the differentiation of immature stages of two necrophagous flies (Diptera: Calliphoridae) of potential forensic importance. *Neotrop. Entomol.* 34:777–783.
- Vidigal, T. H., L. C. Montresor, A. J. Simpson, and O. S. Carvalho. 2002. Polymerase chain reaction and restriction fragment length polymorphism of cytochrome oxidase subunit I used for differentiation of Brazilian *Biomphalaria* species intermediate host of *Schistosoma mansoni*. *Mem. Inst. Oswaldo Cruz* 97:47–52.