DNA Barcode Development for Three Recent Exotic Whitefly (Hemiptera: Aleyrodidae) Invaders in Florida

Authors: Aaron M. Dickey, Ian C. Stocks, Trevor Smith, Lance Osborne, and Cindy L. McKenzie

Source: Florida Entomologist, 98(2) : 473-478
Published By: Florida Entomological Society
URL: https://doi.org/10.1653/024.098.0213
DNA barcode development for three recent exotic whitefly (Hemiptera: Aleyrodidae) invaders in Florida

Aaron M. Dickey1, Ian C. Stocks2, Trevor Smith3, Lance Osborne1, and Cindy L. McKenzie3*

Abstract
Several new whitefly (Hemiptera: Aleyrodidae) species have become established in Florida in the past decade. Three of these, fig whitefly (FW), rugose spiraling whitefly (RSW), and Bondar’s nesting whitefly (BNW), have caused noticeable damage to residential plants in the landscape including ficus hedges, palms, and bird of paradise. Whiteflies are difficult to identify and 4th instar nymphs are needed for morphological identification making whiteflies good candidates for identification via DNA barcoding. A DNA barcoding cocktail to amplify the 5′ end of the cox1 mitochondrial gene from these species was developed. Subsequently, primers were developed for each species, validated with multiple populations collected throughout Florida, and a phylogenetic tree was constructed for placement of the 3 species in the whitefly tree of life. Besides FW, RSW, and BNW, 2 additional species of whiteflies were detected in collections, namely Paraleyrodes pseudonaranjae Martin (Hemiptera: Aleyrodidae) and a species provisionally designated Aleurodicinae sp1. RSW and BNW clustered with congeners within the phylogeny, and FW was resolved as a possible sister taxa to the genus Bemisia. The barcoding cocktail should allow sequencing of 5′ cox1 from multiple genera and both sub-families of whiteflies, and the primers developed for each species will facilitate rapid identification of these 3 invasive whiteflies.

Key Words: cox1 barcoding; invasive species; Singhiella simplex; Aleurodicus rugioperculatus; Paraleyrodes bondari; Paraleyrodes pseudonaranjae

Resumen
Varias especies de mosca blanca (Hemiptera: Aleyrodidae) se han establecido recientemente en la Florida durante la década pasada. Tres, en particular, han causado daños notables a las plantas residenciales en el campo como los setos de ficus, las palmeras y plantas de aves del paraíso. Estas plagas son la mosca blanca de higo (MBH), la mosca blanca rugosa de espiral (MBRS) y la mosca blanca de anidación de Bondar (MBAB). Las moscas blancas son difíciles de identificar y se necesitan las ninfas de cuarto estadio para la identificación morfológica lo que las convierte en buenos candidatos para la identificación a través de los códigos de barras de ADN. Un cóctel de códigos de barras de ADN para amplificar el extremo 5′ del gen mitocondrial cox1 fue desarrollado utilizando la inosina nucleósido junto con oligonucleótidos tradicionales degenerados y no degenerados. Posteriormente, especies primers específicos fueron desarrollados para cada especie, validados con múltiples poblaciones recolectadas en toda la Florida, y un árbol filogenético fue construido para la colocación de las tres especies en el árbol de la vida de la mosca blanca. Además de MBH, MBRS y MBAB; Se detectaron dos especies adicionales de moscas blancas en las colecciones, Paraleyrodes pseudonaranjae Martin (Hemiptera: Aleyrodidae), y una especie designados provisionalmente como Aleurodicinae sp1. Las MDRS y MBAB fueron agrupadas con congéneres dentro de la filogénea y MBH se resolvió como un posible taxón hermano al género, Bemisia. El cóctel de código de barras debe permitir la secuenciación de 5′ cox1 para una amplia diversidad de moscas blancas y los específicos primers de especies desarrollados facilitará la identificación rápida de estas tres especies invasivas de mosca blanca.

Palabras Clave: cox1; especies invasoras; Singhiella simplex; Aleurodicus rugioperculatus; Paraleyrodes bondari; Paraleyrodes pseudonaranjae

The state of Florida hosts a large number of exotic species with many new “invasives” arriving annually (Simberloff et al. 1997). Invasive species cause billions of dollars worth of damage (Pimentel et al. 2005) and major ecological disruptions to native ecosystems (Simberloff et al. 2013). Among invasive insects establishing in Florida over the past decade are 3 whitefly species (Hemiptera: Aleyrodidae) that cause highly visible wax and sooty mold buildup in urban plantings as well as defoliation: the fig whitefly (FW) Singhiella simplex Singh (subfamily Aleurodinae), the rugose spiraling whitefly (RSW) Aleurodicus rugioperculatus Martin, and Bondar’s nesting whitefly (BNW) Paraleyrodes bondari Perachi (both subfamily Aleurodicae). These 3 species have been established in Florida since 2007, 2009, and 2011 respectively (Hodges 2007; Stocks 2012; Stocks & Hodges 2012). Although FW feeds primarily on Ficus benjamina in Florida, it is recorded from 4 species of Ficus (Hodges 2007). RSW and BNW have broader host ranges including plants in the genera Ficus, Psidium, Annona, and Pouteria (BNW) and 32 genera in 12 plant families (RSW) (Stocks 2012; Stocks and Hodges 2012). The apparent preference of BNW for F. benjamina in Florida, a
host not in the literature (Stocks 2012), suggests BNW may represent a cryptic species complex and underscores the importance of accurate pest identification when developing treatment plans.

DNA barcoding is a method of identifying an organism based on its DNA sequence at a specific genetic locus (Hebert et al. 2003). This method has great advantages for identifying taxa morphologically difficult to distinguish (Hebert et al. 2004), including many paraneopteran insects (Foottit et al. 2008; Dinsdale et al. 2010). For example, adult whiteflies have very few characters available for morphological identification (Martin 1987), show high intraspecific variability (Rosell et al. 1997; Calvert et al. 2001), and require a specific life stage, namely 4th instar nymphs, for accurate identification (Hodges & Evans 2005). This makes them excellent candidates for identification using DNA barcoding. The genetic locus used as a DNA barcode is generally taxon specific, with portions of the 16S, rbcL, matK, and cox1 genes being used for bacteria, plants, and animals respectively (Hebert et al. 2003; Oline 2006; Hollingsworth et al. 2009). Although the 5' end of the mitochondrial gene cox1 is considered the standard for DNA barcoding in animals, this ~650 bp segment is not always easily sequenced due to mismatches between universal primers and the sequences of target taxa (Ivanova et al. 2007). In whiteflies, the cox1 5' end has been largely abandoned as a barcode, particularly in the pest complex Bemisia tabaci Gennadius, in favor of the 3' end of the same gene for which primers more faithfully amplify DNA (Shatters et al. 2009), but see Ashfaq et al. (2014). For difficult or unknown targets, a new strategy for barcoding primer design is to couple multiple degenerate primers into a single primer cocktail with strategic placement of the nucleoside inosine, which pairs with A, C, and T. This strategy was used successfully in a large-scale barcoding effort in fish (Ivanova et al. 2007). Herein, we report the successful use of this strategy to develop 1) a barcode primer cocktail and 2) DNA barcodes for FW, BNW, and RSW. We subsequently designed species-specific primers to amplify an abbreviated portion of the barcode from each species, validated these primers with at least 2 populations each, and performed a phylogenetic analysis to predict evolutionary placement of these 3 invaders in the cox1 whitefly phylogeny.

Materials and Methods

The primer cocktail was designed by aligning mitochondrial cox1 sequences from the NCBI nr database that contained one or both regions upon which universal invertebrate primers have been developed (Folmer et al. 1994) for species in the insect superorder Paraneoptera, containing Thysanoptera, Psocodoea, and Hemiptera. Cocktail 3 (Table 1) consisted of 6 primers: a forward and a reverse primer of each of the 3 types a) non-degenerate, b) traditional degenerate (IUPAC) coded, and c) inosine degenerate, with inosine replacing each degenerate position. The non-degenerate forward and reverse primers were modified from Folmer et al. (1994) to be the consensus sequence from the Paraneoptera alignment. The locations of traditional and inosine degeneracy followed Ivanova et al. (2007), and M13 tails were added to the 5' end of each primer (Ivanova et al. 2007).

Samples of whiteflies from which individuals had been identified by I. Stocks, G. Hodges (RSW and BW), and P. Avery (FW) were used in barcode development. Genomic DNA was extracted from individual whiteflies by boiling in lysis buffer (Dickey et al. 2012) or using the DNeasy kit (Qiagen Inc., Valencia, California, USA). Polymerase chain reactions (PCR; 25 μL) were run using the GoTaq kit (Promega, Madison, Wisconsin, USA) with 1 μL DNA template. The total primer concentration in each reaction was 1 μM corresponding to a concentration of each primer in the cocktail of 30, 10, and 10 nM for inosine, degenerate, and non-degenerate primers, respectively (Ivanova et al. 2007) (Table 1). PCR products were visualized using agarose gel electrophoresis, cleaned with a Nucleospin (Machery-Nagel, Bethlehem, Pennsylvania, USA) kit, and directly sequenced bi-directionally using a BigDye Terminator cycle sequencing kit and an 3730XL DNA sequencer (both Thermo Fisher, Waltham, Massachusetts, USA). M13, cocktail primers, and species-specific primers were used for sequencing reactions. If needed to ensure full-length barcodes, a second PCR was conducted using 1 μL of cleaned-up PCR product as template. This was done for FW, and no PCR errors were detected between the initial (349 nt) and final (658 nt) barcode sequences.

Following the amplification of barcodes, specific primers were designed for each of the 3 species using Primer 3 (Rozen & Skaletsky 2000). These primers were used to validate multiple populations of each species with single individuals from collections identified by I. Stocks or G. Hodges or samples identified in the field by L. Osborne, V. Kumar, or A. Francis (Table 2). PCR primer concentration was 0.8 μM, and PCR and sequencing reactions were run using the same kits used in barcode development.

The alignment used for phylogeny inference consisted of 104 Aleyrodidae and 3 Aphididae (outgroup) barcodes and accession numbers are provided (Fig. 1). Barcodes not sequenced during this study were downloaded from GenBank and BOLD. All genetic groups within 3% nearest neighbor distance available in GenBank and BOLD as of 15 Jan 2015 are represented in the phylogeny. Barcodes were translated and the resulting amino acid sequences were used to construct the phylogeny. Outgroup taxa were the aphid species Acyrthosiphon pismum Harris, Cervaphis quercus Takahashi, and Pterocomma pilosum Buckton (Hemiptera). The alignment was completed in Mesquite 2.75 (Maddison & Maddison 2001) using Clustal X (Thompson et al. 1997). The best-fitting model of sequence evolution was selected using ProtTest 3.4

### Table 1. Cocktail 3: a whitefly barcoding primer cocktail.

<table>
<thead>
<tr>
<th>Name</th>
<th>Ratio</th>
<th>Primer sequence 5’ to 3’</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cocktail 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1-J1514f</td>
<td>1</td>
<td>TTGAAAAACGACGTTCCTTTCCAAAATCACATAGATATTT</td>
<td>Current</td>
</tr>
<tr>
<td>C1-J1514g</td>
<td>1</td>
<td>TTGAAAAACGACGTTCCTTTCCAAAATCACATAGATATTT</td>
<td>Current</td>
</tr>
<tr>
<td>C1-J1514h</td>
<td>10</td>
<td>TGTTAAAAACGACGTTCCTTTCCAAAATCACATAGATATTT</td>
<td>Current</td>
</tr>
<tr>
<td>C1-N2173j</td>
<td>0.5</td>
<td>CAGGAAAAACGACGTTCCTTTCCAAAATCACATAGATATTT</td>
<td>Current</td>
</tr>
<tr>
<td>C1-N2173k</td>
<td>0.5</td>
<td>CAGGAAAAACGACGTTCCTTTCCAAAATCACATAGATATTT</td>
<td>Current</td>
</tr>
<tr>
<td>C1-N2173l</td>
<td>0.5</td>
<td>CAGGAAAAACGACGTTCCTTTCCAAAATCACATAGATATTT</td>
<td>Current</td>
</tr>
<tr>
<td>M13F</td>
<td>1</td>
<td>TTGAAAAACGACGTTCCTTTCCAAAATCACATAGATATTT</td>
<td>Messing 1983</td>
</tr>
<tr>
<td>M13R</td>
<td>1</td>
<td>TTGAAAAACGACGTTCCTTTCCAAAATCACATAGATATTT</td>
<td>Messing 1983</td>
</tr>
</tbody>
</table>

The primer naming convention follows Simon et al. (2006); Gene (C1-Cytochrome Oxidase I)-Majority(J) or Minority(N) strand and location in the ancestral arthropod genome. Lower-case letters denote unique primers at the same location. The M13 tail of each primer is shown in italics. PCR conditions: 94 °C for 1 min; 36 cycles of 94 °C for 30 s, 50 °C for 1 min, 72 °C for 1 min; 72 °C for 10 min. Primers are modified from Folmer et al. (1994).
Dickey et al.: Exotic whitefly barcoding

A complete 655 nt DNA barcode was obtained each from *A. rugioperculatus* (RSW), *P. bondari* (BNW), and *S. simplex* (FW) using cocktail 3, and the 3 barcodes are deposited in GenBank (Table 3). Pairwise nucleotide differences among these barcodes ranged from 19.6% to 27.8% p-distance. No SNPs were detected in RSW, BNW, or FW. 4th species, *Paraleyrodes pseudonaranjae* Martin, was identified based on its barcode sequence and found to be present in 2 mixed species populations collected from pond apple *Annona glabra* L. (Magnoliales: Annonaceae) and mango *Mangifera* sp. (Sapindales: Anacardiaceae) (Table 2). Barcodes from this species have also been deposited in GenBank (Table 3). The 2 *P. pseudonaranjae* haplotypes were separated by 3 SNPs (1.2%), 2 synonymous and 1 non-synonymous (I or V) mutation. These haplotypes were originally detected using the BNW primers (Table 2) and have been deposited as abbreviated barcode sequences. A longer nucleotide sequence was recovered from a cocktail 3 PCR product from *P. pseudonaranjae* haplotype 2 but not the full-length barcode. Haplotype 1 of this species could not be amplified with cocktail 3. 5th species, designated here as *Aleurodicinae sp1* (Table 3), was amplified from all specimens collected from *F. benjamina* in Orange County, Florida, in 2012 (Table 2). This species groups with the genus *Paraleyrodes* in a phylogenetic tree with weak to moderate support (67% bootstrap, 89 SH) (Fig. 1) and its nucleic acid sequence differs from BNW by 17.3%.

Two populations from which 4th instar nymphs had been identified as BNW were discovered to be mixed populations. From 8 whiteflies barcoded from pond apple, 5 were BNW and 3 were *P. pseudonaranjae*. From 9 whiteflies barcoded from mango, 3 were BNW, 2 were RSW, and 4 were *P. pseudonaranjae*. The FW population in Saint Lucie County, Florida, contained Aleurodicinae sp1 at a very low frequency (2 of 20 samples). The remainder of the populations contained only the species of whitefly morphologically identified. DNA from all individuals in the populations of whiteflies used for validation could be amplified using at least 1 of the primers designed for each species.

The phylogenetic tree resolved Aleurodidae as monophyletic with high support (100% bootstrap, 100 SH) (Fig. 1). The subfamilies Aleurodicinae and Aleyrodinae were not well supported within the tree. Several genera, including *Aleyrodus*, were resolved as paraphyletic. *Aleyrodus rugioperculatus* (RSW) is placed in the tree within the clade containing *A. dugesii* Cockerell, *A. dispersus* Russell and Lecanoi-
Fig. 1. Maximum Likelihood phylogeny of whiteflies using the mitochondrial cox1 barcode. Log-likelihood: −9515.93005, substitution model: MtART+I+G, support values: bootstrap % / SH approximate likelihood ratio test. Highlighted nodes 1: Aleyrodidae, 2: Paraleyrodes, 3: Aleurodicus dispersus–Lecanoideus floccissimus complex, 4: Singhiella + Massileurodes + Aleurolobus + Bemisia. Accessions KF059961 and HQ446157 are possible specimen misidentifications in GenBank given the phylogeny.
deus floccissimus Martin with moderate to high support (93% bootstrap, 97 SH). Paraleyrodes bondari (BNW) groups with its congener, *P. pseudonaranjae*, with weak to moderate support (67% bootstrap, 89 SH). *Singhiella simplex* (FW) or a clade containing FW is recovered as a possible sister to *Bemisia* with moderate to high support (84% bootstrap, 95 SH).

**Discussion**

Whiteflies are an important insect group because of their role as plant pests and invasive species, and correct species identification is critical for accurate detection of new and existing invasives and their range expansions and for consistent communication regarding management outcomes. DNA barcoding adds to the repertoire of identification methods available to experts and non-experts and allows for rapid taxonomic identification for life stages other than 4th instar nymphs. The primers developed for each species are useful for identifying the 3 invasive species in Florida. They should be particularly beneficial for identifying RSW and documenting its expanding distribution within the southern US and its increasing host breadth (Stocks 2013). The primers designed for BNW and RSW often, but not always, amplified other members of the Aleurodicinae for which they were not specifically designed (Table 2). For the purposes of identification by sequencing, this was a useful feature and allowed initial detection of *P. pseudonaranjae* in mixed species samples. *Paraleyrodes pseudonaranjae* co-occurred with BNW on pond apple and with both BNW and RSW on mango. *Paraleyrodes pseudonaranjae* has been established in Florida since the 1940s (Martin 2001), and it is noteworthy that it can co-occur with BNW and RSW on at least 2 hosts.

The taxon identified here as *Aleurodicinae sp1* was identified as BNW in the field and without slide vouchers. The clustering of this species with other members of the genus *Paraleyrodes* in a phylogenetic tree suggests that it may represent another species of *Paraleyrodes* (Fig. 1). No 4th instar nymphs are available from this species for morphological identification, and it has not been detected in our BNW collections in Florida since 2012.

The phylogeny presented supports the possible inclusion of RSW into the *Aleurodicus dispersus–Lecanoideus floccissimus* complex (Callejas et al. 2005), which agrees with results based on morphology. *Aleurodicus* is a large genus and many of the species are difficult to identify based on morphology. Furthermore, the apparent non-monophyly of highly pestiferous genera such as *Bemisia* and *Aleurodicus* in the phylogeny highlights the need to continue work delimiting and resolving generic boundaries in Aleurodidae. It should be noted that this is a preliminary phylogenetic analysis based on a single gene and clustered taxon sampling representing mostly pests. Single gene phylogenies do not account for sources of false phylogenetic signal such as incomplete lineage sorting (Carstens & Knowles 2007) and introgression (Fontaine et al. 2015), and phylogenies can be strongly affected by taxon sampling (Pick et al. 2010). Future molecular phylogenetics of whiteflies should include many more genera, all 3 subfamilies and multiple loci.

The methods of Ivanova et al. (2007) allowed for the successful DNA barcoding of 3 newly invasive whiteflies in Florida for which there was no prior genetic information. Using the cocktail, we also obtained and deposited to GenBank barcodes from the following species established in Florida: *Metaleurodicus cardini* Back, *Dialeurodes scheffleri* Hodges & Dooley, *Aleurotrachelus trachoides* Quaintance & Baker, and *Trialeurodes variabilis* Quaintance (Table 3). Successful sequencing of these species further validates the utility and broad application of both cocktail 3 and the approach of Ivanova et al. (2007) in whitefly barcoding. The barcodes developed in this study are at the 5′ end of the coxi gene and thus are compatible with the recommendations of the Consortium for the Barcode of Life initiative (www.barcodeoflife.org) and align with the majority of animal barcodes sequenced to date. Barcodes for the species described in this paper are unique additions to GenBank and were not in either GenBank or the Barcode of Life Database prior to this study with the exception of *P. pseudonaranjae* KP032221, a 100% sequence identity match to KF595126 already in GenBank.

Recently, there has been another barcoding effort in whiteflies focused on the 5′ end of coxl (Ashfaq et al. 2014) finding 90% of *Bemisia* tabaci representing 7 cryptic species in the complex that were successfully barcoded using a cocktail of 3 primers including trnW forward primers (Park et al. 2010). Primers that cross gene boundaries may be problematic for groups of arthropods with highly rearranged mitochondrial genomes such as whiteflies, though this does not appear to be the case for trnW (Thao et al. 2004). Park et al.'s (2010) primers should be further tested with additional genera of whiteflies to demonstrate their broad utility. In contrast, cocktail 3 has amplified barcodes from 7 genera and 2 subfamilies of whiteflies. Both Ashfaq et al. (2014) and the current study represent important early steps encouraging the whitefly community to adopt the standard 5′ animal barcode region. This may be a challenging transition given the widespread adoption of the 3′ alternative region and its associated cryptic species identification framework for the major pest *B. tabaci* (Dinsdale et al. 2010). Although we anticipate that the 3′ barcode (Folmer et al. 1994; Shatters et al. 2009) will continue to be widely used with whiteflies, we encourage the whitefly community to use trnW (Park et al. 2010), cocktail 3 (current work), and future applications of these primer design approaches to obtain standard 5′ DNA barcodes from whiteflies that will be comparable across metazoa.

**Table 3.** Whitefly DNA barcodes submitted to GenBank.

<table>
<thead>
<tr>
<th>Species</th>
<th>GenBank accession</th>
<th>Size (bp)</th>
<th>County (year collected)</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aleurodicinae sp1</em></td>
<td>KP032214</td>
<td>1474</td>
<td>Saint Lucie</td>
<td><em>Ficus benjamina</em></td>
</tr>
<tr>
<td><em>Paraleyrodes bondari</em></td>
<td>KP032215</td>
<td>658</td>
<td>Orange (2014)</td>
<td><em>Ficus benjamina</em></td>
</tr>
<tr>
<td><em>Singhiella simplex</em></td>
<td>KP032216</td>
<td>658</td>
<td>Orange</td>
<td><em>Ficus benjamina</em></td>
</tr>
<tr>
<td><em>Trialeurodes variabilis</em></td>
<td>KP032217</td>
<td>658</td>
<td>Orange</td>
<td><em>Carica papaya</em></td>
</tr>
<tr>
<td><em>Aleurotrachelus trachoides</em></td>
<td>KP032218</td>
<td>658</td>
<td>Orange</td>
<td><em>Capsicum annuum</em></td>
</tr>
<tr>
<td><em>Aleurodicus rugiperfolius</em></td>
<td>KP032219</td>
<td>658</td>
<td>Saint Lucie (2014)</td>
<td><em>Streitzaia nicolai</em></td>
</tr>
<tr>
<td><em>Dialeurodes scheffleri</em></td>
<td>KP032220</td>
<td>636</td>
<td>Orange</td>
<td><em>Scheffleri sp.</em></td>
</tr>
<tr>
<td><em>Paraleyrodes pseudonaranjae</em> haplotype 2</td>
<td>KP032221</td>
<td>596</td>
<td>Dade</td>
<td><em>Annona glabra</em></td>
</tr>
<tr>
<td><em>Metaleurodicus cardini</em></td>
<td>KP032222</td>
<td>566</td>
<td>Glades</td>
<td><em>Psidium guajava</em></td>
</tr>
<tr>
<td><em>Paraleyrodes pseudonaranjae</em> haplotype 1</td>
<td>KP032223</td>
<td>249</td>
<td>Dade</td>
<td><em>Annona glabra</em></td>
</tr>
</tbody>
</table>

*bp: base pairs*
Acknowledgments

We thank John Prokop and Florian Grant for technical support. This research was made possible, in part, by a Cooperative Agreement from the United States Department of Agriculture’s Animal and Plant Health Inspection Service (APHIS). It may not necessarily express APHIS’ views.

References Cited


