IDENTIFICATION OF *Encarsia* FÖRSTER (HYMENOPTERA: APHELINIDAE) SPECIES FROM NORTHEASTERN MÉXICO BY CYTOCHROME C OXIDASE SUBUNIT 1 (*CO1*) SEQUENCING †

[IDENTIFICACIÓN DE ESPECIES DE *Encarsia* FÖRSTER (HYMENOPTERA: APHELINIDAE) DEL NORESTE DE MÉXICO MEDIANTE SECUENCIACIÓN DE LA SUBUNIDAD 1 DE CITOCROMO C OXIDASA (*CO1*)]

D. A. García-Guerrero[1], O. García-Martínez[2], S. Nikolaevna-Myartseva[3]; L. Palomo-Ligas[4], L. A. Aguirre-Uribe[2], M. Flores-Dávila[2], A. C. Flores-Gallegos[4], R. Rodríguez-Herrera[4*] and H. Hernández-Torres[5].

[1]Centro de Investigación para los Recursos Naturales (CIReNa), antigua Escuela Normal Rural s/n, CP 33941, Salaices, Municipio de López, Chihuahua, Mexico. E-mail. ggda28@gmail.com

[2]Departamento de Parasitología Agrícola, Universidad Autónoma Agraria Antonio Narro, Calz Antonio Narro 1923, Buenavista, 25315 Saltillo, Coah. México. E-mail. drogarcia@yahoo.com.mx, auribe@yahoo.com, cisef9@hotmail.com

[3]Facultad de Ingeniería y Ciencias, Universidad Autónoma de Tamaulipas, Centro Universitario, Campus, 87149. Ciudad Victoria 87149, Tamaulipas, México. E-mail. smyartse@docentes.uat.edu.mx

[4]Departamento de Alimentos, Facultad de Ciencias Químicas, Universidad Autónoma de Coahuila, Blvd. V. Carranza s/n. Col. República Oriente C.P.25280 Saltillo, Coahuila, México. E-mail. lissethe_palomo@uadec.edu.mx, carolinaflores@uadec.edu.mx, raul.rodriguez@uadec.edu.mx

[5]Unidad Académica Multidisciplinaria Mante, Universidad Autónoma de Tamaulipas, Boulevard Enrique Cárdenas González 1201 Pte., Col. Jardín, C. P. 89840, Ciudad Mante, Tamaulipas, México. E-mail. hermelindo.torres@uat.edu.mx

**Corresponding author*

SUMMARY

Background. Aphelinidae, comprising more than 1120 species in 40 genera, are economically significant parasitoids known for preying on Aleyrodidae, Coccidae, Pseudococcidae, and Aphididae. However, there is scarcity of registered sequences for Aphelinidae and the *Encarsia* genus in Mexico's databases. With a widespread distribution, *Encarsia* stands as the largest genus within this family, boasting 96 species. The application of sequencing techniques, particularly targeting the cytochrome c oxidase subunit 1 (*CO1*) mitochondrial gene, has considerably eased species identification, providing valuable information at the species level. **Objective.** To identify *Encarsia* species through *CO1* gene sequencing. By doing so, the study sought to augment the information available for this group in the GenBank and globally, facilitating further taxonomic comprehension and unveiling potential cryptic species. **Methodology.** During 2015 and 2016, we collected Aleyrodidae and Diaspididae nymphs from various locations, namely Tampico Alto and Ciudad Cuauhtémoc in Veracruz, and Altamira, Tamaulipas, and Saltillo, Coahuila. Taxonomic identification of the species *Encarsia citrina*, *Encarsia perplexa*, and *Encarsia tamaulipeca* was conducted alongside *CO1* gene sequencing. The obtained sequences were subsequently deposited in the GenBank under the accession keys MF444685, MF444686, and MF444687, respectively. **Results.** Through the application of *CO1* gene sequencing, we successfully identified three *Encarsia* species in the regions under investigation. Notably, the registration of the latter two species marked their first-ever presence in the GenBank, further augmenting the knowledge base for this genus on a global scale. **Implications.** The inclusion of these new sequences in the GenBank represents a significant step forward for Aphelinidae, specifically *Encarsia*, in Mexico. This expansion of data will

† *Submitted August 7, 2023 – Accepted December 17, 2024. http://doi.org/10.56369/tsaes.5083*

ISSN: 1870-0462.

ORCID = D. A. García-Guerrero: http://orcid.org/0000-0002-6526-3827; **O. García-Martínez:** http://orcid.org/0000-0002-3115-418X; **S. Nikolaevna-Myartseva:** http://orcid.org/0000-0003-3173-6208; **L. Palomo-Ligas:** http://orcid.org/0000-0002-6399-5856; **L. A. Aguirre-Uribe:** http://orcid.org/0000-0001-8462-1012; **M. Flores-Dávila:** http://orcid.org/0000-0002-8196-6119; **A. C. Flores-Gallegos:** http://orcid.org/0000-0001-5092-1404; **R. Rodríguez-Herrera:** http://orcid.org/0000-0002-6428-4925; **H. Hernández-Torres:** http://orcid.org/0000-0002-6170-0137

serve as a valuable tool for validating traditional identification methods, exploring intra- and interspecies variations, and shedding light on previously unknown cryptic species. **Conclusions.** By utilizing *CO1*sequencing, this study successfully identified and registered three *Encarsia* species in the GenBank, including two species previously unrecorded worldwide. The newfound genetic data will be instrumental in enhancing our understanding of this ecologically important group, contributing to more precise taxonomic assessments and encouraging further investigations into the diversity and distribution of Aphelinidae in Mexico and beyond.
Key words: Aphelinidae; Encarsia; *CO1* gene sequencing; Species identification.

RESUMEN

Antecedentes: Los Aphelinidae, compuestos por más de 1120 especies en 40 géneros, son parasitoides económicamente importantes conocidos por depredar Aleyrodidae, Coccidae, Pseudococcidae y Aphididae. *Encarsia*, es el género más extenso en esta familia, con 96 especies. Sin embargo hay escasez de secuencias registradas para Aphelinidae y el género *Encarsia* en las bases de datos de México. La aplicación de técnicas de secuenciación, particularmente dirigidas al gen mitocondrial citocromo c oxidasa subunidad 1 (*CO1*), ha facilitado considerablemente la identificación de especies, proporcionando información valiosa a nivel de especie. **Objetivo**: Identificar especies de *Encarsia* mediante la secuenciación del gen *CO1*. Con ello, se buscó aumentar la información disponible para este grupo en el GenBank y así facilitar la comprensión taxonómica revelando posibles especies crípticas. **Metodología**: Durante 2015 y 2016, recolectamos ninfas de Aleyrodidae y Diaspididae de diversas ubicaciones, Tampico Alto y Ciudad Cuauhtémoc en Veracruz, Altamira, Tamaulipas y Saltillo, Coahuila. La identificación taxonómica de las especies *Encarsia citrina*, *Encarsia perplexa* y *Encarsia tamaulipeca* se realizó junto con la secuenciación del gen *CO1*. Las secuencias obtenidas se depositaron posteriormente en el GenBank con las claves de acceso MF444685, MF444686 y MF444687, respectivamente. **Resultados**: Mediante la aplicación de la secuenciación del gen *CO1*, se identificaron con éxito tres especies de *Encarsia* en las regiones investigadas. Cabe destacar que el registro de las dos últimas especies marcó su primera presencia en el GenBank, aumentando aún más la base de conocimientos para este género a nivel mundial. **Implicaciones**: La inclusión de estas nuevas secuencias en el GenBank representa un avance significativo para Aphelinidae, específicamente *Encarsia,* en México. Esta expansión de datos servirá como una herramienta valiosa para validar los métodos de identificación tradicionales, explorar las variaciones intra e interespecíficas, y arrojar luz sobre especies crípticas previamente desconocidas. **Conclusiones**: Mediante la utilización de la secuenciación del gen *CO1*, este estudio identificó y registró con éxito tres especies de *Encarsia* en el GenBank, incluidas dos especies que no se habían registrado previamente a nivel mundial. Los nuevos datos genéticos serán fundamentales para mejorar nuestra comprensión de este grupo de gran importancia ecológica, contribuyendo a evaluaciones taxonómicas más precisas y alentando investigaciones adicionales sobre la diversidad y distribución de Aphelinidae en México y el mundo.
Palabras clave: Aphelinidae; Encarsia; secuenciación del gen *CO1;* identificación de especies.

INTRODUCTION

Traditional taxonomy has allowed species classification, determination, and differentiation taking into account morphological characteristics of each specimen and is currently the basis for classification of all living organisms. Some aspects that can complicate taxonomic identification are morphological variations among individuals that may be due to climate, habitat, diet or many other environmental factors; the examination of genomic variations in these same individuals combined with correlations to environmental factors can aid in determining whether such variations are simply normal individual differences within a species, the emergence of a subspecies, or even warrant classifying some as belonging to a previously unidentified (and thus cryptic) species (Brusca, 2002). About 97% of insect species lack a precise taxonomic description of characters that allow them to be differentiated from each other (Martin-Piera, 2000).

This is exemplified by the Eulophidae in Mexico where studies of the local fauna provided new taxonomic keys for this group and resulted in the identification of new genera and species that provide important progress in the taxonomy of the group (Hansson, 1997). Previously, keys based on fauna found in the United States of America were used, but, often found not to characterize many Mexican specimens of interest, even at the genus level. These newly available keys provide an opportunity to compare and contrast classical identifications with genomic data also obtained from the Mexican fauna. It should be mentioned that, species identification through traditional taxonomy is essential, and the use of DNA-based techniques should be considered as a useful and complementary tool in species identification and classification. The barcode initiative that began in Canada (Hebert *et al.*, 2003) remains as an ambitious project of global importance, which it is intended to identify any species in the world through the sequencing of its DNA (similar to a barcode used for supermarket products). These sequences are being

collected on the different GenBank data platforms and available to make comparisons.

Different methods and techniques have facilitated insect species identification using small fragments of cellular material (Hajibabaei et al., 2006); however, it remains a problem the lack of taxonomic descriptions of most species, which limits the sequencing and registration of these species. DNA barcodes are based on the identification of species from the mitochondrial gene cytochrome c oxidase subunit 1 (*CO1*) fragment, allowing a high variation among species of the same genus (Hebert et al., 2003). The conserved DNA sequence of *CO1* include six regions: Region 1 (567-639), Region 2 (591-666), Region 3 (595-674), Region 4 (700-793), Region 5 (709-823) and Region 6 (732-830).

Aphelinidae are important parasitoids of Aleyrodidae, Coccidae, Pseudococcidae, and Aphididae; they have a cosmopolitan distribution and include more than 1120 species in 40 recognized genera. In Mexico, 13 of these are recorded and *Encarsia* is the best represented with 96 species (Myartseva et al., 2012). The keys of Myartseva *et al* (2012) provide a basis for comparing and contrasting these species with genomic data we obtain in the research reported here, to further refine our understanding of this genus in Mexico. In this research, the identification of species of *Encarsia* using the amplification and sequencing of the *CO1* gene was performed.

MATERIALS AND METHODS

Sample collection

Fourth stage nymphs of Aleyrodidae and nymphs of Diaspididae were collected during 2015 (March-October) and 2016 (March-November) in Tampico Alto (22° 06' 47" N, 97 ° 50' 53" W) and Ciudad Cuauhtémoc (22° 12' 11.46" N, 97 ° 50' 16.04" W) in Veracruz; Altamira (22° 33' 42.53" N, 98° 09' 12.96" W) in Tamaulipas and Saltillo (25° 26' 34.51" N, 100° 59' 22.93" O) in Coahuila. Leaves of *Psidium guajava* (L), *Citrus sinensis* (L) and *Hedera helix* (L) were collected and placed in Petri dishes labeled and transferred to the laboratory to allow the emergence of adult parasitoids. Samples were checked daily. Emerged adults were collected with a fine hair brush wet with 70% ethanol; 90% of samples collected were stored in 2 mL Eppendorf tubes containing 70% ethanol at 4°C (labeled with field data), Noyes (1982), the other 10% was used for taxonomic and molecular identification (Noyes, 1982).

Taxonomic identification

Five percent of the remaining ten percent of specimens were mounted on slides, with Canada balm (Noyes, 1982) and the key for aphelinids of Mexico by Myartseva *et al*., (2012) was used for taxonomic identification. Voucher specimens were deposited in the Museo de Insectos de la Facultad de Agronomía de la Universidad Autónoma de Tamaulipas (MIFA-UAT), in Ciudad Victoria, Tamaulipas, México.

DNA extraction

The remaining specimens (5%) were processed in the Laboratory of Molecular Biology, Department of Food Research (DFR) of the Autonomous University of Coahuila (UA de C). DNA extraction was performed using two techniques, the non-destructive (HotShot) one of Truett *et al*., (2000) and the Doyle and Doyle (1987) technique (with adaptations) (Doyle and Doyle, 1987; Truett *et al*., 2000). With the first technique, the entire specimen can be preserved for slide mounting unlike the second, which is a completely destructive method to carry out DNA extraction.

HotShot technique for *Encarsia tamaulipeca* (Myartseva and Coronado-Blanco, 2002)

The first two lysis buffers, 5 µL of 100 mM NaOH and 15 µL of 0.26 mM EDTA disodium were mixed in a 0.2 mL Eppendorf tube; introducing one specimen per tube; if the specimen was in alcohol, it was dried completely on blotting paper for one hour at room temperature prior to amplification; a light vortex shake (Prism mini, C1801) to the tube was done to mix both buffers together with the insect and introduced to the thermocycler (Axygen) for 30 minutes at 95 °C; after this, 20 µl of 40 Mm Tris-HCl was added to the tube as a buffer solution; and the liquid was then transferred to a new tube (Truett et al., 2000).

Doyle and Doyle technique (with adaptations), for *Encarsia citrina* (Craw, 1891) and *Encarsia perplexa* (Huang and Polaszek, 1998)

The A matrix and beads were homogenized in 1.5 ml centrifuge tubes and ten specimens were placed in each tube; lysis buffer (Tris HCl pH 8 (100 mM), EDTA pH 8 (50 mM), NaCl (50 mM), 2% SDS) was included and the tubes were transferred to a tissue disruptor (MP Fast-Prep 24) for 60 seconds. The tubes were then removed and placed in a centrifuge (Spectrofuge 24D) for three minutes at 13,300 × g to reduce froth, then placed in the disruptor for 60 seconds and again in the centrifuge to obtain the total solution, which was placed in a new 1.5 ml Eppendorf tube; 400 ml of Phenol-chloroform-isoamyl alcohol solution (25:24.1 v/v) was added and centrifuged at 13,300 × g for 15 minutes. After this, the procedure mentioned in the

original technique was used, but the centrifugation processes were performed under cold conditions (Doyle and Doyle, 1987).

PCR amplification

PCR amplification, HotShot technique

The polymerase chain reaction (PCR) was performed with a final volume of 25 µL, using an Axygen-type thermocycler and for the Cytochrome Oxidase Subunit 1, gene primers C1-J-1718 (5'-GGAGGATTTGGAAATTGATTAGTTCC-3') and C1-N-2191 (5'-CCCGGTAAAATTAAAATATAAACTTC-3 '), specific for insects (Simon et al., 1994). The PCR mixture consisted of 11.3 µL of Q-water, 5 µL of 5X buffer, 2 µL of Mg Cl_2 (25 mM), 0.5 µL of dNTPs (10 mM), 2 µL of BSA (10 mg/ml), 1 µL of each primer (10 µM), 0.2 µL Taq DNA polymerase (5U/ µL) and 2 µL of DNA (100 ng/ µL). DNA was quantified using the Epoch spectrophotometer (BioTek Instruments) conditioned with the Take3 plate and the Gen 5 (5.1.11v) software. The thermocycler program used was: 2 min at 94°C followed by six cycles of 30 s at 94°C; 90 s at 45°C and 70 s at 72°C; later, 36 cycles of 30 s at 94°C; 90 s at 51°C and 70 s at 72°C with five min at 72°C at the end of the last cycle (Truett *et al*., 2000). The PCR product was separated by electrophoresis for 30 min on 105 volt agarose gel at 1.5% with the molecular ladder 100 bp marker (Promega), stained with 3 µl ethidium bromide (0.3 □g/□L), the gel was observed and photographed on a UV light transilluminator.

PCR amplification, Doyle and Doyle technique

The same final volume and primers of the previous method were used. The PCR mixture was: 14.5 µL of Mili-Q water; 2.5 µL of 10X buffer (10 mM $MgCl_2$); 0.5 µl dNTPs (10 mM); 2 µL of each primer (10 µM); 0.5 µl of Taq DNA polymerase (5U/µL) and 2 µL DNA (14.31 ng/µL). The temperature program was: ten min at 94°C, followed by 35 cycles of 1 min at 94°C; 1 min at 55°C; 1 min at 72°C and the final extension of five minutes at 72°C. The PCR product was separated with the same procedure as in the previous method. The PCR product of each species (22 µL) was transferred to a 1.5 ml tube adding 200 µL of 1X TE buffer pH 8, 220 µL of sodium acetate; two volumes of cold 99.9% (440 µL) ethanol were added and allowed to incubate for 60 minutes at -20 °C, and centrifuged at 16,300 × *g* for ten minutes at 4 °C. The pellet was then washed with 300 µL of cold 70% ethanol, centrifuged at 16,300 × *g* for five minutes at 4 °C; the supernatant was decanted and the pellet was allowed to dry for at least four hours and re-suspended in 22 µL of sterile Mili-Q water.

Sequencing and analysis

The amplified products were sequenced on a Perkin Elmer Applied Biosystems Model 3730 by the Taq FS Dye Terminator Cycle Sequencing Fluorescence-Based Sequencing method. The sequences obtained were compared to the sequences deposited in the NCBI database using nucleotide BLAST (Basic Local Alignment Search Tool) for highly similar sequences. The sequences were edited and aligned using the BioEdit software. Each sequence was loaded in the PAK2 (version 1.0.0.0) software to obtain the bar code of each species.

Determination of conserved regions and estimation of synonymous and nonsynonymous changes

To identify the conserved regions and the synonymous/nonsynonymous substitutions of DNA, a Multiple Sequence Alignment of 23-mtDNA *CO1* sequences was done using the online server MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle/) (Edgar, 2004a and 2004b). The Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution was calculated by MEGA X software Version 10.0.5 (Kumar, *et al.*, 2018; Tamura, Nei, and Kumar, 2004). When substitution-rate variation exists among nucleotide sites but is ignored, there are possibilities to underestimate biases in the rates of change among nucleotides (Wakeley, 1994). The alignment was downloaded and saved in a compatible format for the next analysis. To determine conserved regions and assess synonymous and nonsynonymous changes, we conducted a comprehensive analysis of the coding regions of 23 aligned sequences until the detection of a termination codon using specialized software. The analysis was perform with the DNA Sequence Polymorphism (DnaSP) software Version 6.1 (Rozas *et al*., 2017).

Phylogenetic tree construction

In order to estimate the phylogenetic status of *CO1* sequences from *Encarsia* species, 23 mt-DNA sequences deposited in the NCBI nucleotide database (http://www.ncbi.nlm.nih.gov/nucleotide/) were aligned by MUSCLE method, and after that, the alignment was used for Neighbor-Joining tree construction using MEGA X (Kumar *et al*., 2018; Saitou and Nei, 1987). The evolutionary distances were computed using the Maximum Composite Likelihood method employing the same software (Tamura *et al*., 2004).

RESULTS AND DISCUSSION

Taxonomic identification

The species collected in the field were *E. citrina*, *E. perplexa* and *E. tamaulipeca* which were previously described in Mexico by (Myartseva and Coronado-Blanco, 2002). Taxonomical identification was performed using keys established by (Craw, 1891; Huang and Polaszek, 1998; Myartseva *et al*., 2012; Myartseva and Coronado-Blanco, 2002).

Molecular identification

E. tamaulipeca: The technique HotShot, worked well for this specie, obtaining a well-defined amplification of the fragment of 518 bp (Fig. 1). Also, the specimens spent less than four months preserved in ethanol, which allowed DNA of good quality to be obtained with minimal degradation by ethanol.

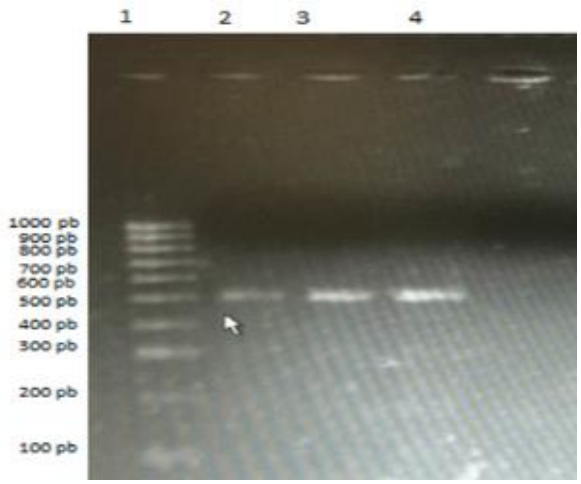

Figure 1. DNA amplification by PCR of the *CO1* region from *Encarsia tamaulipeca*. Lane 1= molecular marker Invitrogen 100 bp ladder. Lanes 2, 3 and 4= samples of the same species, DNA was separated using an agarose gel 1.5%.

E. citrina and *E. perplexa*: The Doyle and Doyle technique was better adapted to be used with both species, which are smaller in size compared to *E. tamaulipeca.* Specimens were preserved for more than five months in ethanol and amplification of the 518 bp (Figs 2 and 3) was obtained for both species as well as for *E. tamaulipeca* noted earlier. The sizes of the amplicons obtained in this study match those reported by De León et al. (2010) for the genus *Encarsia*, suggesting affiliation with this genus, as according to De León (2010), the amplicon size varies depending on the genus.

When comparing the sequence of *E. citrina* in the GenBank, a 98.45% identity was obtained with the registered sequence of *E. citrina* from Japan and the accession key KF778511. The sequence of these species was sent for registration and it is considered the first record in the GenBank database for Mexico with the access code MF444685.

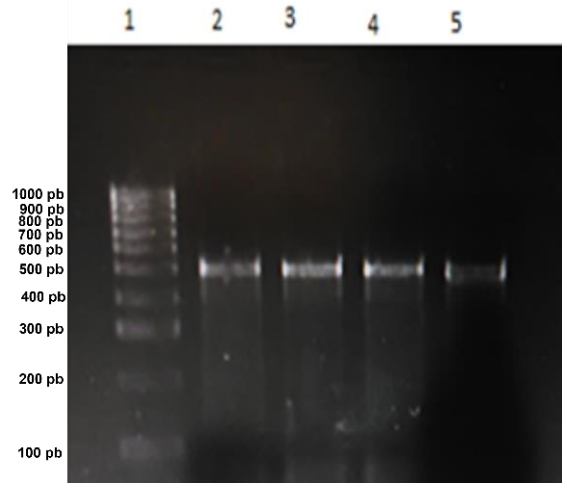

Figure 2. PCR product amplification of the *CO1* DNA region from *Encarsia spp*, (lane 1) molecular weight marker Invitrogen 100 bp ladder, lane 2).- *E. citrina*, lane 3.- *E. perplexa* and lanes 4 and 5.- *E. tamaulipeca*. DNA was isolated using the Doyle and Doyle technique. DNA was separated using an agarose gel 1.5%.

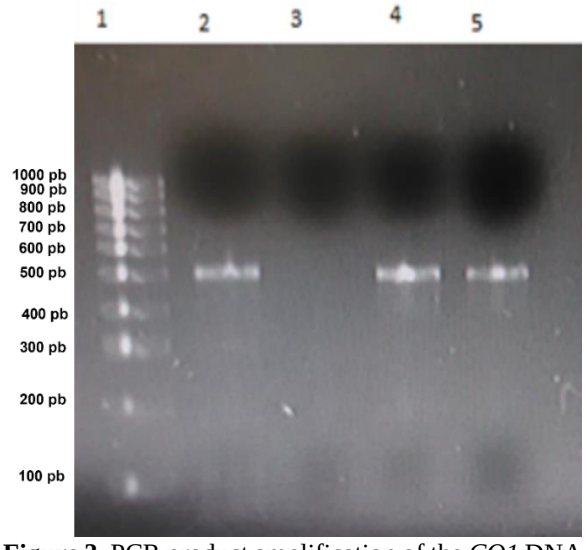

Figure 3. PCR product amplification of the *CO1* DNA region from *Encarsia perplexa*, (lane 1) molecular weight marker Invitrogen 100 bp ladder, lane 2, 4 and 5= samples of the same species. DNA was isolated using the Doyle and Doyle technique. DNA was separated using an agarose gel 1.5%.

The comparison of *E. perplexa* sequence had 93.39% and *E. tamaulipeca* 92.67% identity with the *E. brimblecombei* (Girault), registered in Japan with access code KF778461 in the GenBank database. Since to date, there is no record of the *CO1* gene sequencing of both species in that database, they were sent for registration on the GenBank platform with access codes MF444686 and MF444687 respectively. Both species are new records in the world on the GenBank. Zhang *et al*., (2014) carried out adult extractions of the apple aphid *Eriosoma lanigerum* Hausmann 1802, parasitized by *Aphelinus mali* (Haldeman) from which the *CO1* fragment of the DNA was extracted and amplified with primers F (5'-TCTCATATAATTTGTAATGAAAG-3') and R (5'-TGATAACTAGGAGGAAAATTTAT-3'), where they obtained 648 bp fragments, which served to differentiate two parasitoid populations; compared to this research, a difference of 130 bp was obtained, with the difference considered due to our work being with a different genus (Zhang *et al*., 2014).

There are studies related to amplification and sequencing of species of *Encarsia*, in which by amplification and sequencing of the 28S rDNA region of seven races of *Encarsia formosa* (Gahan), and two of *Encarsia luteola* (Howard), the amount of genetic variation within and between species at this locus was estimated (Babcock and Heraty, 2000). In addition, Wethersbee *et al.,* (2004) performed PCR amplification of the 18S r DNA gene from three species of parasitoids *Lysiphlebus testaceipes* (Cresson), *Lipolexis scutellaris* (Mackaver) and *Aphelinus gossypii* (Timberlake) from the citrus brown aphid *Toxoptera citricida* (Kirkaldy), for species detection and differentiation by specific primers, which allowed amplification of DNA of each parasitoid even within its host (Weathersbee et al., 2004). However, it should be mentioned that the levels of genetic variation differ between loci of the same genome, because the mutation rates are different, likewise, the genetic variation is different between loci of two different genomes (nuclear and chloroplastic, as in this case). By this reason, this study aims to identify the species based on the variation in the chloroplast gene *CO1*, where the non-destructive HotShot technique was employed, in this case, DNA flows through mouth, anus, and spiracles, worked for *E. tamaulipeca*, which we believe, was due to these specimens being preserved for less than three months of curation in 70% ethanol; the two remaining species were curated in ethanol for more than five months after being collected from the field. Guzman *et al.,* (2016) obtained bands and fragments that were better defined in different species of Trichogrammatidae, Mymaridae, and Eulophidae than ours and we opine this may be due to their being processed after one day to four months preserved in 70 and 100% ethanol (Guzmán-Larralde *et al.,* 2016). For *E. citrina* and *E. perplexa*, it was not possible to obtain DNA with the minimally invasive HotShot technique. It has been reported that very little DNA is usually obtained from a single insect; among the factors that influence these results are: the collection, transportation and preservation procedures of the specimens (Rivero et al. 2007). Storage temperature is a key factor; in the short term, it is better to store insects at 4°C, because freezing-thawing processes usually promote contamination and DNA degradation (Gómez and Uribe 2007). The completely destructive Doyle and Doyle technique successfully obtained DNA by including the most sclerotized structures of the insect (legs, antennae, and head) by complete maceration despite the specimens having been preserved in ethanol for more than four months.

Determination of *CO1* sequence conservation

The nucleotide frequencies of the *CO1* Encarsia sequences are 31.85% (A), 44.59% (T/U), 10.38% (C), and 13.18% (G). The transition/transversion rate ratios were k_1= 1.371 (purines) and k_2= 3.703 (pyrimidines). The overall transition/transversion bias is R= 0.925, where $R = [A * G * k_1 + T * C * k_2]/[(A + G) * (T + C)]$. This analysis involved 23 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All ambiguous positions were removed for each sequence; there were 1221 positions in the final dataset. In the nucleotide composition of the *CO1* gene sequence, our results revealed a high A/T content, in agreement with previous studies on insect mitochondrial DNA and another study about the variability in the *Encarsia* genus using also de *CO1* gene (Monti *et al*., 2007).

The ratio of transitions to transversions is higher in mitochondrial DNA, this effect can be attributed to the effect of methylation (Keller *et al.,* 2007). In our results, the transitional substitutions are higher than the transversions (Table 1).

Table 1. Pattern of Nucleotide Substitution (Transition vs. Transversion) in *CO1* Gene Sequences of the *Encarsia* Genus.

	A	T	C	G
A	-	10.18	2.33	4.13
T	7.25	-	7.5	2.97
C	7.25	32.82	-	2.97
G	10.08	10.08	2.33	-

Estimation of Synonymous and Nonsynonymous Changes

To estimate the substitutions, we used the alignment of the 23 sequences of *CO1* (1-1278 bp). Sites with gaps

in the complete data file were excluded. The DnaSP software estimate 26.05 Synonymous sites and 84.95 NonSynonymous sites. Table 2 shows the substitutions for each sequence, and in Table 3 are shown the Ks/Ka ratios of *Encarsia citrina* strain DAG1, *Encarsia perplexa* strain DAG2, and *Encarsia tamaulipeca* strain DAG3 *CO1* sequences. In general, the non-synonymous nucleotide substitutions cause an amino acid change (Gonzales *et al.*, 2002). In our results, the number of nonsynonymous sites is higher to the synonymous sites (Table 2).

Table 2. Comparison and distribution of Synonymous and Non-Synonymous nucleotide substitutions across 23 *CO1* Sequences.

Protein Coding Region	Synonymous sites	Non-Synonymous sites
MH475901.1	25.33	85.67
MH475900.1	25.33	85.67
MH475899.1	24.33	86.67
MH129875.1	25.67	85.33
MH536771.1	24.67	86.33
KF778511.1	25.00	86.00
MH475903.1	26.33	84.67
KT748705.1	25.67	85.33
MG263528.1	25.67	85.33
KT748716.1	25.00	86.00
KT748725.1	25.00	86.00
MH475902.1	24.67	86.33
KF778401.1	25.00	86.00
KF778464.1	24.33	86.67
KF778411.1	24.33	86.67
KF778461.1	26.00	85.00
KF778433.1	26.00	85.00
MF444685.1	27.67	83.33
KY607910.1	27.67	83.33
MH456759.1	28.67	82.33
JQ083712.1	29.33	81.67
MF444686.1	29.00	82.00
MF444687.1	28.50	82.50

Nevertheless, it has been reported that many of the nucleotide polymorphisms within the *Encarsia genus* were due to synonymous changes (Monti *et al*., 2007). By another side, the non-synonymous/synonymous mutation ratio (Ka/Ks) in our results is minor to 1 (Table 3). This rate is of great significance in phylogeny and evolutionary dynamics of protein-coding sequences across closely related and yet diverged species. Indicates negative (purifying) selection when Ka is less than Ks (Zhang *et al*., 2006). According to the neutral theory, the synonymous substitutions that does not change the amino acid in the protein will be tolerated, but nonsynonymous substitutions will be removed by purifying selection. However, if a nonsynonymous substitution confers some selective advantage, then it will be rapidly fixed in the population by positive selection. Consequently, nonsynonymous substitutions will be higher than synonymous substitutions (Chu and Wei, 2019).

Evolutionary relationships of taxa

The evolutionary history was inferred using the Neighbor-Joining method. Fig. 4 shows the optimal tree calculated with the sum of branch length = 1.44404312. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. This analysis involved 23 nucleotide sequences with accession number and localization described in Figure 4. The mitochondrial DNA cytochrome oxidase c subunit 1 (*CO1*) gene is one of the most popular markers for population genetic and phylogeographic studies across the animal kingdom (Monti *et al*., 2007). In our phylogenetic tree, 2 major clades are shown. Note that taxa *Encarsia citrina* strain DAG1, *Encarsia perplexa* strain DAG2 and *Encarsia tamaulipeca* strain DAG3 have a common ancestor belonging to the same clade or lineage. These species are also related to *E. lounsburyi* and *Encarsia perniciosi*. The mitochondrial *CO1* gene sequences analysis showed a close relationship between the sequences of interest.

Table 3. Non-Synonymous/Synonymous Mutation Ratios (Ka/Ks) for *CO1* Sequences of *Encarsia citrina* strain DAG1, *Encarsia perplexa* strain DAG2, and *Encarsia tamaulipeca* strain DAG3.

Sequence 1	Sequence 2	Ks	Ka
MF444685.1	'MF444686.1'	0.2488	0.0500
MF444685.1	'MF444687.1'	0.3030	0.0499
'MF444686.1'	'MF444687.1'	0.0356	0.0503

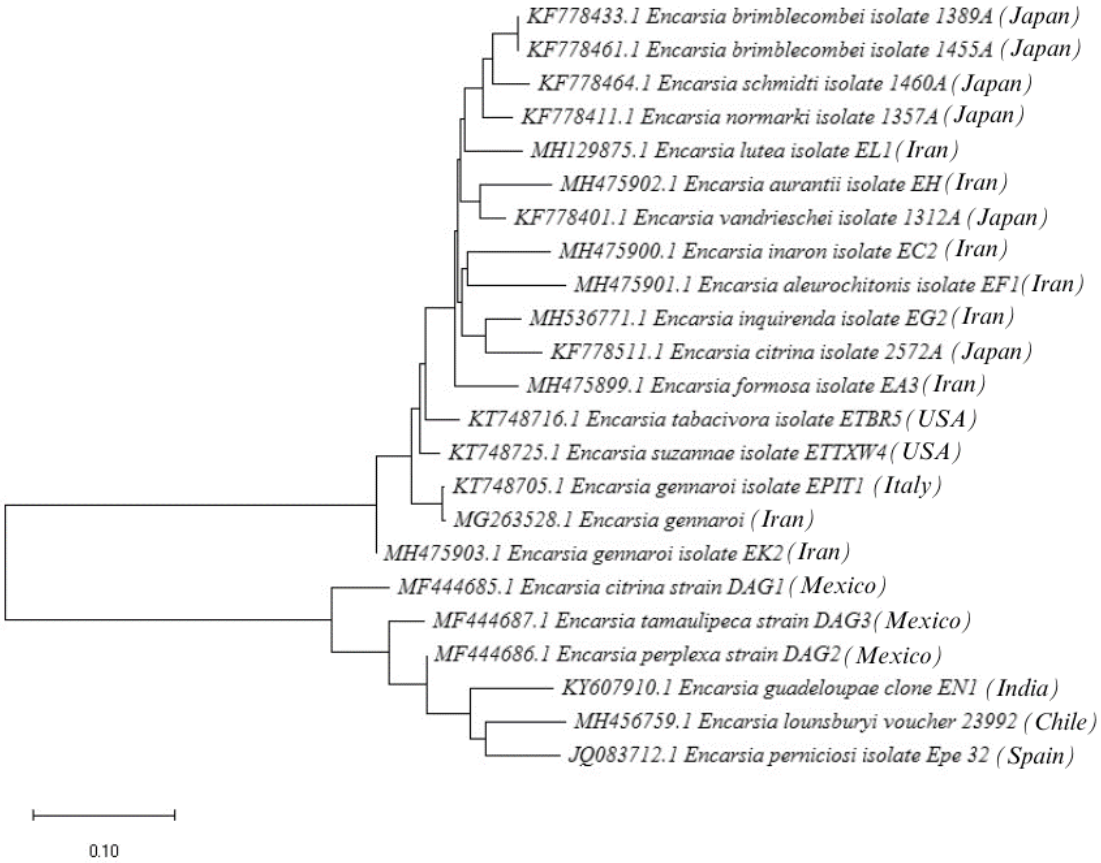

Figure 4. Phylogenetic Tree Based on Mitochondrial *CO1* Gene Sequences of the *Encarsia* Genus. This phylogenetic tree was constructed using the Neighbor-Joining method with evolutionary distances calculated by the Maximum Composite Likelihood method. The tree displays the relationships among 23 nucleotide sequences (with accession numbers and locations), emphasizing the evolutionary history and lineage relationships within the *Encarsia* genus. The tree illustrates two major clades, showing the close relationship between *Encarsia citrina* strain DAG1, *Encarsia perplexa* strain DAG2, and *Encarsia tamaulipeca* strain DAG3, as well as their association with other species.

CONCLUSION

In this research, three different species of the *Encarsia* genus were identified by taxonomic keys, later these species were also identified by DNA amplification and sequencing of the *CO1* gene. These species were: *Encarsia citrina*, *Encarsia perplexa*, and *Encarsia tamaulipeca*, their *CO1* gene sequences were registered in the GenBank, including two species previously unrecorded worldwide. In the nucleotide composition of the CO1 gene sequence, results revealed: high A/T content, transitional substitutions are higher than the transversions and the number of nonsynonymous sites is higher to the synonymous sites. The newfound genetic data will be instrumental in enhancing our understanding of this ecologically important group, contributing to more precise taxonomic assessments and encouraging further investigations into the diversity and distribution of Aphelinidae in Mexico and beyond.

Acknowledgments

To. We thank Dr Adriana Guzmán Larralde (postdoctoral student) in Universidad Autonóma de Nuevo León and M.C. Gerardo Arcos Cavazos, from Centro de Investigación Regional del Noreste, at Campo Experimental Las Huastecas for their technical assistance.

Funding. Consejo Nacional de Ciencia y Tecnología for a scholarship (CVU 478580) to D. A. García-Guerrero during his Ph.D. program, and Universidad Autonóma Agraria Antonio Narro, through financing the projects of O. Garcia Martinez's graduate students.

Conflict of interest. The authors declare no conflict of interest.

Compliance with ethical standards. Not applicable.

Data availability. The data is available upon request, with the corresponding author raul.rodriguez@uadec.edu.mx

Author contribution statement (CRediT). D.A. García-Guerrero – Conceptualization, Formal analysis, Methodology, Writing-original draft. **O. García-Martínez** – Conceptualization, Formal analysis, Methodology, Writing-original draft. **S. Nikolaevna-Myartseva** – Conceptualization, Formal analysis, Methodology, Writing-original draft. **L. Palomo-Ligas** – Data curation, Validation, Supervision, Resources, Writing–review & editing. **L.A. Aguirre-Uribe** – Data curation, Validation, Supervision, Resources, Writing–review & editing. **M. Flores-Dávila** – Data curation, Validation, Supervision, Resources, Writing–review & editing. **A.C. Flores-Gallegos** – Data curation, Validation, Supervision, Resources, Writing–review & editing. **R. Rodríguez-Herrera** – Funding acquisition, Supervision, Validation, Resources, Writing–review & editing. **H. Hernández-Torres** – Supervision, Validation, Resources, Writing–review & editing

REFERENCES

Babcock, C.S. and Heraty, J.M., 2000. Molecular markers distinguishing *Encarsia formosa* and *Encarsia luteola* (Hymenoptera: Aphelinidae). *Annals of the Entomological Society of America*, 93(4), pp.738-744. http://doi.org 10.1603/0013-8746(2000)093[0738:MMDEFA]2.0.CO;2

Brusca, R.C. and Brusca, G.J., 2003. Invertebrates. 2ª ed. Sinauer Associates, Massachusetts, 936 pp. *Ciencia y Mar*, 9(25), pp.45-47.

Chu, D. and Wei, L., 2019. Nonsynonymous, synonymous and nonsense mutations in human cancer-related genes undergo stronger purifying selections than expectation. *BMC Cancer*, 19, pp.359. http://doi.org/10.1186/s12885-019-5572-x

Craw, A., 1891. Internal parasites discovered in the San Gabriel Valley; recommendations and notes. California State Board. *Horticultural Bulletin*, 57, pp.1-7

Doyle, J.J. and Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19, pp.11-15

Edgar, R.C., 2004a. Muscle: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, 5, pp.113. http://doi.org/10.1186/1471-2105-5-113

Edgar, R.C., 2004b. Muscle: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), pp.1792-1797. http://doi.org/10.1093/nar/gkh340

Gómez, M. and Uribe, S., 2010. Optimización de un procedimiento de extracción de ADN para mosquitos anofelinos. *Revista Colombiana de Entomología*, 36(2), pp.260-263. http://doi.org/10.25100/socolen.v36i2.9156

Gonzales, M.J., Dugan, J.M. and Shafer, R.W., 2002. Synonymous-non-synonymous mutation rates between sequences containing ambiguous nucleotides (Syn-SCAN). *Bioinformatics*, 18(6), pp.886-887. http://doi.org/10.1093/bioinformatics/18.6.886

Guzmán-Larralde, A.J., Suaste-Dzul, A.P., Gallou, A. and Peña-Carrillo, K.I., 2016. DNA recovery from microhymenoptera using six non-destructive methodologies with considerations for subsequent preparation of museum slides. *Genome*, 60(1), pp.85-91. http://doi.org/10.1139/gen-2015-0172

De León, J., Neumann, G., Follett, P. and Hollingsworth, R., 2010. Molecular markers discriminate closely related species *Encarsia diaspidicola* and *Encarsia berlesei* (Hymenoptera: Aphelinidae): Biocontrol candidate agents for white peach scale in Hawaii. *Journal of Economic Entomology*, 103, pp.908-916. http://doi.org/10.1603/EC09316

Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W. and Hebert, P.D.N., 2006. DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America*, 103(4), pp.968-971. http://doi.org/10.1073/pnas.0510466103

Hansson, C., 1997. Survey of *Chrysocharis Förster* and *Neochrysocharis Kurdjumov* (Hymenoptera, Eulophidae) from Mexico, including eight new species. *Miscel·lània Zoològica*, 20, pp.81-95.

Hebert, P.D., Cywinska, A., Ball, S.L. and deWaard, J.R., 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London B: *Biological Sciences*,

270(1512), pp.313-321. http://doi.org/10.1098/rspb.2002.2218

Huang, J. and Polaszek, A., 1998. A revision of the Chinese species of *Encarsia Forster* (Hymenoptera: Aphelinidae): Parasitoids of whiteflies, scale insects, and aphids (Hemiptera: Aleyrodidae, Diaspididae, Aphidoidea). *Journal of Natural History*, 32, pp.1825-1966. http://doi.org/10.1080/00222939800770911

Keller, I., Bensasson, D. and Nichols, R.A., 2007. Transition-Transversion bias is not universal: A counter example from grasshopper pseudogenes. *PLOS Genetics*, 3(2), e22. http://doi.org/10.1371/journal.pgen.0030022

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K., 2018. Mega X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35(6), pp.1547-1549. http://doi.org/10.1093/molbev/msy096

Martín-Piera, F. and Lobo, J.M., 2000. Diagnóstico sobre el conocimiento sistemático y biogeográfico de insectos hiperdiversos (Coleoptera, Hymenoptera y Lepidoptera) en España. En: Hacia un proyecto CYTED para el inventario y estimación de la diversidad entomológica en Iberoamérica: PRIBES 200. m3m-Monografías Tercer Milenio, Vol.1, *Sociedad Entomológica Aragonesa* (SEA), Zaragoza, pp.287-308

Monti, M.M., Nappo, A.G. and Giorgini, M., 2007. Molecular characterization of closely related species in the parasitic genus *Encarsia* (Hymenoptera: Aphelinidae) based on the mitochondrial cytochrome oxidase subunit I gene. *Bulletin of Entomological Research*, 95(5), pp.401-408. http://doi.org/10.1079/BER2005371

Myartseva, S.N., Ruiz-Cancino, E. and Coronado-Blanco, J.M., 2012. Serie Avispas parasíticas de plagas y otros insectos No. 8. Aphelinidae (Hymenoptera: Chalcidoidea) de importancia agrícola en México. Revisión y claves. *Universidad Autónoma de Tamaulipas*, Ciudad Victoria, México. 413 p.

Myartseva, S.N. and Coronado, J.M., 2002. A new parasitoid of whiteflies from Mexico, with a key to New World species of the genus *Encarsiella* (Hymenoptera: Aphelinidae). *Florida Entomologist*, 85(4), pp.620–624. http://doi.org/10.1653/0015-4040(2002)085[0620:ANPOWF]2.0.CO;2

Rivero, J., Zoghbi, N., Rubio-Palis, Y., Urdaneta, L. and Herrera, F., 2007. DNA degradation of *Anopheles darlingi* collected at high relative humidity and preserved in isopropanol. *Boletín de Malariología y Salud Ambiental*, 47(1), pp.149–151. Retrieved December 9, 2024, from http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1690-46482007000100013&lng=es&tlng=en

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E. and Sánchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, 34(12), pp.3299–3302. http://doi.org/10.1093/molbev/msx248

Saitou, N. and Nei, M., 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), pp.406–425. http://doi.org/10.1093/oxfordjournals.molbev.a040454

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. and Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, 87(6), pp.651–701. http://doi.org/10.1093/aesa/87.6.651

Tamura, K., Nei, M. and Kumar, S., 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences of the United States of America*, 101(30), pp.11030–11035. http://doi.org/10.1073/pnas.0404206101

Truett, G.E., Heeger, P., Mynatt, R.L., Truett, A.A., Walker, J.A. and Warman, M.L., 2000. Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *BioTechniques*, 29(1), pp.52–54. http://doi.org/10.2144/00291bm09

Wakeley, J., 1994. Substitution-rate variation among sites and the estimation of transition bias. *Molecular Biology and Evolution*, 11(3), pp.436-442. http://doi.org/10.1093/oxfordjournals.molbev.a040124

Weathersbee, A.A., Shufran, K.A., Panchal, T.D., Dang, P.M. and Evans, G.A., 2004. Detection and Differentiation of Parasitoids (Hymenoptera: Aphidiidae and Aphelinidae) of the Brown Citrus Aphid (Homoptera: Aphididae): Species-Specific Polymerase Chain Reaction Amplification of 18S rDNA. *Annals of the Entomological Society of America*, 97(2), pp.286-292. http://doi.org/10.1603/0013-8746(2004)097[0286:DADOPH]2.0.CO;2

Zhang, R.-M., Zhou, H.-X., Guo, D., Tao, Y.-L., Wan, F.-H., Wu, Q. and Chu, D., 2014. Two putative bridgehead populations of *Aphelinus mali* (Hymenoptera: Aphelinidae) introduced in China as revealed by mitochondrial DNA marker. *Florida Entomologist*, 97(2), pp.401-405. http://doi.org/10.1653/024.097.0209

Zhang, Z., Li, J., Zhao, X.-Q., Wang, J., Wong, G.K.-S. and Yu, J., 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics, Proteomics & Bioinformatics*, 4(4), pp.259-263. http://doi.org/10.1016/S1672-0229(07)60007-2