



## Original Article

# Saving DNA from museum specimens: The success of DNA mini-barcodes in haplotype reconstruction in the genus *Anastrepha* (Diptera: Tephritidae)

Sandra M. Velasco-Cuervo<sup>a,\*</sup>, Elkin Aguirre-Ramirez<sup>a</sup>, Jenny Johana Gallo-Franco<sup>a</sup>, Ranulfo González Obando<sup>b</sup>, Nancy Carrejo<sup>b</sup>, Nelson Toro-Perea<sup>a</sup>

<sup>a</sup>Departamento de Biología, Sección de Genética, Universidad del Valle, Calle 13 # 100-00, Cali, Colombia

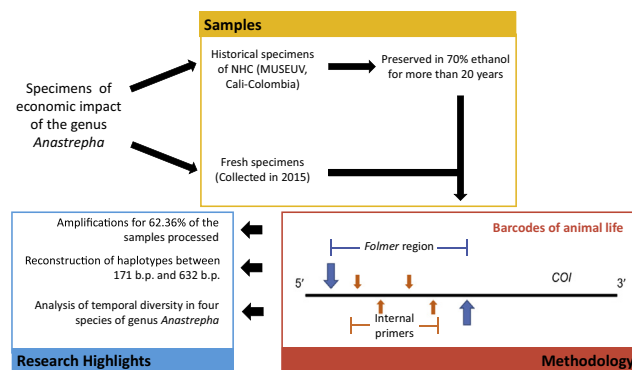
<sup>b</sup>Departamento de Biología, Sección de Entomología, Universidad del Valle, Calle 13 # 100-00, Cali, Colombia



## HIGHLIGHTS

- We contributed to knowledge of NHC's through the use of museum specimens.
- We reconstruct haplotypes had lengths between 171 and 632 bp of the COI gene.
- The amplification rate success was 62.36% of the total processed specimens.
- We made a temporal comparison of haplotypes for four species of the genus *Anastrepha*.
- Mini-barcode primers from the COI region in the genus *Anastrepha* were designed.

## GRAPHICAL ABSTRACT



## ARTICLE INFO

## Article history:

Received 18 August 2018

Revised 1 November 2018

Accepted 4 November 2018

Available online 23 November 2018

## Keywords:

COI

Museum insects

Internal primers

Fruit flies

Natural history collections

## ABSTRACT

The fragmentation of DNA in historical specimens is very common, so obtaining sequences that allow molecular identification and the study of diversity is quite challenging. In this study, we used preserved and fresh specimens of the fruit fly genus *Anastrepha*, a genus of economic impact of fruit crops of the Neotropics. From these specimens, we evaluated: (1) the success PCR amplification rates of mini-barcodes fragments of the cytochrome *c oxidase* subunit I (COI) gene, and (2) the usefulness of mini-barcodes in the reconstruction of haplotypes for the identification of species and the diversity analysis. We used 93 specimens from 12 species, which had been preserved in 70% ethanol for more than 20 years. Internal primers were designed in the COI region and primers available in the literature were also evaluated. We obtained amplifications for 62.36% of the samples processed, and reconstructed haplotypes between 171 bp and 632 bp. Variable amplification rates between combinations of primers and between species were obtained, and molecular identification of some museum specimens was achieved. It was also possible to compare the haplotypes obtained in four species from which both fresh and museum samples were available. Our results also show the importance of the adjustment of the primers for the amplification, allowing to amplify fragments of up to 400 bp. The use available resources in biological collections is key to increasing knowledge of species of interest, and by means of the amplification of mini-barcodes, short sequences can be obtained that allow the molecular identification of specimens and the reconstruction of haplotypes with multiple purposes.

© 2018 The Authors. Published by Elsevier B.V. on behalf of Cairo University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer review under responsibility of Cairo University.

\* Corresponding author.

E-mail address: [sandra.velasco@correounivalle.edu.co](mailto:sandra.velasco@correounivalle.edu.co) (S.M. Velasco-Cuervo).

<https://doi.org/10.1016/j.jare.2018.11.001>

2090-1232/© 2018 The Authors. Published by Elsevier B.V. on behalf of Cairo University.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Introduction

Specimens deposited in museums and Natural History Collections (NHCs), together with molecular biology approaches, are important resources for the generation of knowledge regarding the population dynamics and population genetics. Historical specimens from NHCs are an important source of data that can contribute to a wide range of studies [1], including studies on changes in geographic ranges (spatial and altitudinal), phenological and ecological changes, and evolutionary changes (both genetic and morphological). Regarding evolutionary changes, museum specimens allow historical comparisons of genetic variability as well as the determination of population structures and phylogenetic analyses [2]. Therefore, Holmes et al. [3] have proposed the use of NHCs as a direct tool in studying evolutionary change. Although museum specimens constitute a vast repository of genetic information [4], their DNA is expected to be highly degraded, fragmented and diluted, in addition to having a high risk of contamination by bacteria and fungi [2,3].

Due to DNA fragmentation in museum specimens, PCR amplification of historical DNA has been restricted to short amplicons [2], and various methodologies have been designed for the storage, extraction and amplification of DNA from such samples. Among the methodologies that have been used in attempts to increase the amplification of degraded DNA, mini-barcode amplification represents one of the most popular approaches. This methodology consists of the amplification of fragments with sizes of 100–300 base pairs (bp) in the *cytochrome c oxidase* subunit I (COI) gene, which has been established as the most popular barcode marker [5]. The sequencing of short overlapping fragments has allowed reconstruction of the entire region after assembly, resulting in an effective methodology for species identification [6]. Mini-barcodes have been used to amplify ancient DNA from several taxonomic groups. For example, Patel et al. [7] and Sonet et al. [8] analysed the utility of mini-barcodes in bird orders and species, respectively, solving issues associated with taxonomy and interspecific relationships. In Lepidoptera, the use of mini-barcodes has made it possible to identify cryptic species complexes [9]. Additionally, Francoso and Arias [10] tested mini-barcodes on several bee genera from the Meliponini tribe and were able to recover 619-bp regions for species delimitation for conservation purposes. In black flies (Diptera: Simuliidae), mini-barcodes allowed delimitation species, and the authors highlight the importance of studying the diversity available in NHCs [11]. Büsse et al. [12] analysed mini-barcodes in the nuclear genes (28S rDNA) of 46 arthropod taxa using universal primers and achieved successful amplification in all samples evaluated. Finally, using mini-barcodes, Van Houdt et al. [13] retrieved the COI region of several genera of fruit flies and observed a difference in the success of amplification depending on the date of collection of the specimens. However, Van Houdt et al. [13] did not analyse fruit flies of the genus *Anastrepha* (Schiner 1988).

Fruit flies (Diptera: Tephritidae) are a group of insects that includes numerous species of economic impact around the world. In this context, the genus *Anastrepha* comprises the species with the greatest economic impact found in the Neotropics; these fruit flies attack a wide range of fruit crops and species cultivated throughout South America, Central America, the southern United States, and the Caribbean Islands [14–19]. Due to the impact of these insects on fruit growing, it is important to characterise the diversity of fruit flies. Thus, DNA barcoding analysis can be used in a wide range of applications, such as larval and adult species identification, genetic diversity, demographic history, phylogeography and population structure [17–20].

Obtaining DNA from specimens of the *Anastrepha* genus preserved in museum collections and the subsequent reconstruction of the DNA barcode can allow the reconstruction of haplotypes in fruit flies from the Neotropics. In turn, these haplotypes can facilitate species identification and delimitation and make it possible to analyse their genetic diversity and population dynamics. Internal primers designed for fruit flies by Van Houdt et al. [13] have not been tested in the genus *Anastrepha*; thus, data obtained in these flies by taking advantage of NHCs could expand the scope of study on the evolutionary scale and allow scientists to delve into the historical and spatial changes of *Anastrepha* species of economic impact such as *A. fraterculus* (Wiedemann), *A. obliqua* (Macquart), *A. striata* (Schiner), and *A. grandis* (Macquart), among others, and into their identification by molecular methods. The use of molecular data in the analysis of biological diversity is crucial in the new era of information. In the Neotropics, the importance of the biological resources available in NHCs to the development of knowledge on biodiversity and its uses in research and teaching [21], including species of economic impact and agricultural importance such as fruit flies, has been highlighted.

In this study, various combinations of primers were used to implement mini-barcode amplification in museum specimens of the genus *Anastrepha* that have been preserved in ethanol for 20 years. Mini-barcoding was used to obtain the region of the standard DNA barcode in these animals, that is, a 658-bp fragment located at the 5' end of the COI gene. Haplotypes and species were determined based on the data obtained. These strategies allow better use of museum specimens and permit their possible incorporation in studies of genetic diversity and population dynamics.

## Material and methods

### Museum and fresh specimens

The Museo de Entomología de la Universidad del Valle - MUSENUV is located in southwestern Colombia (Cali, Colombia) and preserves biological specimens of arthropods of great diversity, mounted in ethanol and in the dry state. The *Anastrepha* species available in the MUSENUV that were used in this study were collected and identified between 1993 and 1995 [16]. The collections were obtained in the biogeographic valley region of the Cauca River, at the foothills of the Western and Central Mountain Ranges and on the Pacific coast of southwestern Colombia. The specimens are preserved in 70% ethanol. We worked with 12 species of the genus *Anastrepha* and 93 specimens. Table 1 presents the collection data for species of the genus *Anastrepha* available at MUSENUV that were included in this study. Additionally, fresh samples (collected in 2015) of some *Anastrepha* species from southwestern Colombia (Table 1) were used to compare genetic information on the COI gene with specimens of the same species that had been preserved in ethanol for approximately two decades. Fresh samples were preserved in 96% ethanol and identified individually based on the key of Steyskal (1977).

### Obtaining mini-barcodes from museum specimens

Since the historical samples used in this study had been preserved in 70% ethanol for approximately 20 years, DNA fragmentation was expected [4]. Extraction of DNA from the legs and heads of fruit flies was conducted using a commercial kit (DNeasy Blood and Tissue, QIAGEN, Hilden, Germany) with certain modifications. Briefly, (i) PBS buffer (pH 8) was used instead of ATL buffer because it has been recommended for high-performance purification of animal tissues (QIAGEN, Hilden, Germany); (ii) the samples were incubated for 72 h at 56 °C for cell lysis.

**Table 1**

Collection information for historical and fresh specimens of the *Anastrepha* genus included in this study. The number of individuals processed, the number of individuals for whom PCR products were obtained and sequences useful for the analysis are shown. The positive result of amplification and sequencing corresponds to at least one mini-barcode.

Species	Collection information			Number of individuals processed	Number of individuals with successful amplification	Number of individuals with successful sequencing	
	Locality	Collection coordinates	Collection date (month/year)				Form of collection (Fruit or trap)
<i>A. bahiensis</i>	Pavas	3°40'30.0"N, 76°35'0.96"W	04/1994	Trampa	3	1	0
	Puerto frazadas (Tulúa), Finca Las Camelias	4°05'0.6"N, 75°59'56.39"W	10/1994	Trampa	3	3	3
<i>A. coronilli</i>	Alto Anchicaya	3°31'59.88"N, 76°52'03"W	06/1995	Trampa	5	1	0
<i>A. distincta</i>	Puerto frazadas (Tulúa), Finca Las Camelias	4°05'0.6"N, 75°59'56.39"W	09/1994	Trampa	5	5	5
	Chagres	3°07'1.9"N, 76°36'.19.1"W	06/1995	Trampa	5	3	3
	Sombrerillo	3°47'45.3"N, 76°46'25"W	03/1994	Trampa	5	5	5
	Pavas	3°40'30.0"N, 76°35'0.96"W	01/1994	Trampa	2	0	0
	Alto Anchicaya	3°31'59.88"N, 76°52'03"W	06/1995	Trampa	1	0	0
	Cali	3°22'18.48"N, 76°31'51.6"W	03/2015	<i>Inga edulis</i> (Guama)	3	3	3
<i>A. flavipennis</i>	Mulalo	3°38'7.00"N, 76°28'27.11"W	09/1994	Trampa	5	1	0
<i>A. fraterculus</i>	Puerto frazadas (Tulúa), Finca Las Camelias	4°05'0.6"N, 75°59'56.39"W	09/1994	Trampa	2	2	1
	Vereda el Engaño	3°34'41.3"N, 76°47'15.17"W	06/2015	<i>Psidium guajava</i> (Guayaba)	1	1	1
	La Elsa	3°34'13.7"N, 76°46'14.00"W	06/2015	<i>Psidium guajava</i> (Guayaba)	2	2	2
	El Queremal	3°31'44.7"N, 76°42'51.6"W	06/2015	<i>Psidium guajava</i> (Guayaba)	3	3	3
	Pichinde	3°26'23.4"N, 76°37'01.9"W	06/2015	<i>Psidium guajava</i> (Guayaba)	2	2	2
	Sevilla (Vereda la Cristalina) - Finca el Vergel	4°15'10.02"N, 75°56'02.7"W	03/2015	<i>Coffea</i> sp. (Café)	3	3	3
<i>A. grandis</i>	Sevilla	4°17'55.02"N, 75°53'59.1"W	11/1994	Trampa	5	4	3
	Buitrera	3°22'20.5"N, 76°34'11.3"W	11/1995	Trampa	5	5	5
	Puerto frazadas (Tulúa), Finca Las Camelias	4°05'0.6"N, 75°59'56.39"W	09/1994	Trampa	5	3	2
	Chagres	3°07'1.9"N, 76°36'.19.1"W	09/1994	Trampa	3	3	1
<i>A. leptozona</i>	Sombrerillo	3°47'45.3"N, 76°46'25"W	11/1994	Trampa	5	5	4
	Chagres	3°07'1.9"N, 76°36'.19.1"W	09/1994	Trampa	5	3	3
	Chagres	3°07'1.9"N, 76°36'.19.1"W	01/2015	<i>Pouteria caimito</i> (Caimito)	9	9	9
	Cajamarca-Roldanillo	4°28'37.99"N, 76°12'50.0"W	02/2015	<i>Quararibea cordata</i> (Sapote)	2	2	2
<i>A. mucronota</i>	Sevilla	4°17'55.02"N, 75°53'59.1"W	11/1994	Trampa	3	1	1
	Cajamarca-Roldanillo	4°28'37.99"N, 76°12'50.0"W	02/2015	<i>Quararibea cordata</i> (Sapote)	3	3	3
<i>A. obliqua</i>	Cerrito	3°41'5.0"N, 76°18'48"W	03/1994	<i>Averrhoa carambola</i> (Carambolo)	5	3	1
	Mulalo	3°38'7.00"N, 76°28'27.11"W	03/1994	<i>Mangifera indica</i> (Mango)	5	0	0
<i>A. pickeli</i>	Sombrerillo	3°47'45.3"N, 76°46'25"W	05/1994	Trampa	2	2	2
	Chagres - La Balsa	3°07'1.9"N, 76°36'.19.1"W	06/1994	Trampa	4	0	0
	Cajamarca-Roldanillo	4°28'37.99"N, 76°12'50.0"W	01/2015	<i>Manihot esculenta</i> (Yuca)	5	5	5
<i>A. serpentina</i>	Chagres	3°07'1.9"N, 76°36'.19.1"W	09/1994	Trampa	5	5	4
<i>A. striata</i>	Anchicaya	3°31'59.88"N, 76°52'03"W	02/1994	Trampa	5	4	2

Since the DNA was indeed fragmented (the COI gene was not able to be amplified with the primers of Folmer et al. [22] in museum specimens), it was necessary to use the approach of mini-barcodes, using internal primers targeting the COI gene. Primers designed by Van Houdt et al. [13] for the family Tephritidae were tested, alongside new primers derived from COI sequences of the genus *Anastrepha* available in GenBank and from other COI sequences obtained in this study from fresh specimens (Table 2, Fig. 1). The design of the primers was done with the Primer3 software (v. 4.0.0) [23,24]. The amplification reagents (30 µL for each reaction) consisted of: 1X of PCR Buffer (Thermo Fisher Scientific, Waltham, Massachusetts, USA), 0.2 mM of dNTPs (Thermo Fisher Scientific, Waltham, Massachusetts, USA), 2 mM of MgCl<sub>2</sub> (Thermo Fisher Scientific, Waltham, Massachusetts, USA), 0.6 U of Taq DNA polymerase (Thermo Fisher Scientific, Waltham, Massachusetts, USA), 0.4 µM of each primer (Thermo Fisher Scientific, Waltham,

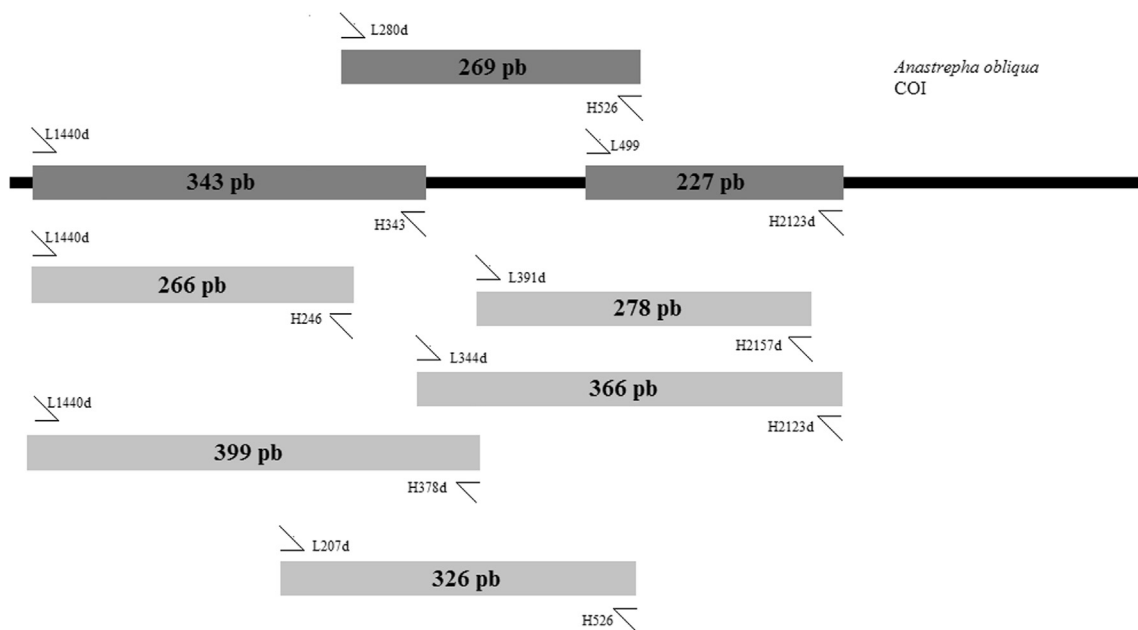
Massachusetts, USA) and 2 µL of DNA for all combinations of primers (at a concentration of 10 ng/µL). The thermal profile consisted of an initial denaturation at 94 °C for 3 min followed by 39 cycles of 94 °C for 30 s, 50 °C for 30 s and 72 °C for 30 s, with a final extension at 72 °C for 7 min. For some pairs of primers, the annealing temperature (50C) was modified as follows: Anas399d, 52 °C, Anas366d and Anas278d, 55 °C.

All primer combinations were evaluated in the studied museum specimens. In addition, all primer pairs were also evaluated in fresh specimens, obtaining 100% amplification (data not shown). However, all COI sequences from fresh specimens were obtained using the universal primers LCO1490 and HCO2198 [22] (Table 2).

The visualization of the PCR products was carried out after electrophoresis in 1.5% agarose gels in 0.5X TBE Buffer (0.045 M tris-borate, 0.001 M EDTA) at 100 V for 30 min, and staining with ethidium bromide, at a final concentration of 0.5 µg/mL, further visu-

**Table 2**  
Combinations of primers evaluated in this study. The % of historical samples amplified for each primer combination is shown. NA: not applicable (historical samples were not evaluated with the Folmer combination).

Marker code (ID)	Primer code	Primer sequence	Ref.	Amplification %
Folmer	LCO1490	GGTCAACAATCATAAAGATATTGG	[22]	NA
	HCO2198	TAAACTTCAGGGTGACCAAAAAATCA	[22]	
Teph343	L1440d	TYTCAACWAATCATAARGATATTGG	[13]	30.9
	H343	CCAGCTCCGTTTTCTACTAT	[13]	
Teph269d	L280d	CGAATAAATAATATAAGATTTTGAYT	[13]	29.0
	H2123d	TAWACTTCWGGRTGWCCAAAAAATCA	[13]	
Teph227	L499	ATTAATATACGATCAACAGGAAT	[13]	11.3
	H2123d	TAWACTTCWGGRTGWCCAAAAAATCA	[13]	
Anas278d	H2157d	TAAAATTGGRTCWCCTCCTC	This research	8.3
	L391d	RTGGAGGAGCTTCACTAGAT	This research	
Anas266d	H246	TCGTGGGAATGCTATATCAGG	This research	30.6
	L1440d	TYTCAACWAATCATAARGATATTGG	[13]	
Anas366d	H2123d	TAWACTTCWGGRTGWCCAAAAAATCA	[13]	29.4
	L344d	TGGTACAGGWTGAACGTGTT	This research	
Anas399d	H378d	TCCTCCATGAGCRATTARAGA	[13]	6.3
	L1440d	TYTCAACWAATCATAARGATATTGG	This research	
Anas326d	H526	ACAAATAAAGGTATTCGGTCAAA	[13]	46.1
	L207d	GGATTMGGAAATTGAWTAGT	This research	



**Fig. 1.** Amplification strategy of the COI gene with internal primers. The black bar represents the COI gene. Different regions of mini-barcodes are represented as dark gray boxes [13] or light gray (Combinations of primers Van Houdt et al., 2010 [13] - This study). The length of each region of mini-barcode is given in base pairs (bp). The position of the primers is indicated by the arrows at the ends of the bars. (Modified, with permission, from Van Houdt et al. [13].)

alised under UV light transilluminator. The purification of the PCR products and sequencing of the DNA fragments was carried out by contracting the services of a specialized provider (MACROGEN Rockville, Maryland, USA).

The sequences obtained were manually edited using Sequencher software (Gene Codes Corporation, Ann Arbor, MI, USA). For each historical specimen for which more than one mini-barcode was obtained, these were concatenated to form contigs of the barcode fragments (Fig. 2a). The sequence contigs of specimens of the same species were aligned and compared to determine haplotypes by species according to the polymorphisms found in the overlapping region (Fig. 2b). Additionally, priming mismatch tests were performed because degraded DNA may contain deficient junctions that can decrease the amplification rate. In this case, each primer's sequence was aligned with each haplotype found (from historical and fresh samples), and the number of mismatches, measured as different base pairs, was counted. An average number of mismatches was obtained for each primer based on the total number of aligned nucleotides on each haplotype. Amplification differences between species and between primer combinations were determined based on the amplification rates.

#### Mini-barcoding of museum specimens for molecular identification of species and diversity assessment

A neighbour-joining (NJ) analysis was performed using the haplotypes of historical samples and fresh samples obtained in this study as well as haplotypes reported for *A. obliqua* [18] and for *A. striata* [19]. We used data reported by Aguirre-Ramirez et al. [18] and Gallo-Franco et al. [19] because they were obtained from fresh specimens collected in the same geographic area as our sample, which allowed us to compare the diversity in the species over time. Using reconstruction through NJ, we verified whether the haplotypes reconstructed from mini-barcodes of museum specimens were more closely related to sequences of the same species obtained from fresh samples. This analysis also allowed us to study

some of the intra- and interspecific relationships of the species analysed.

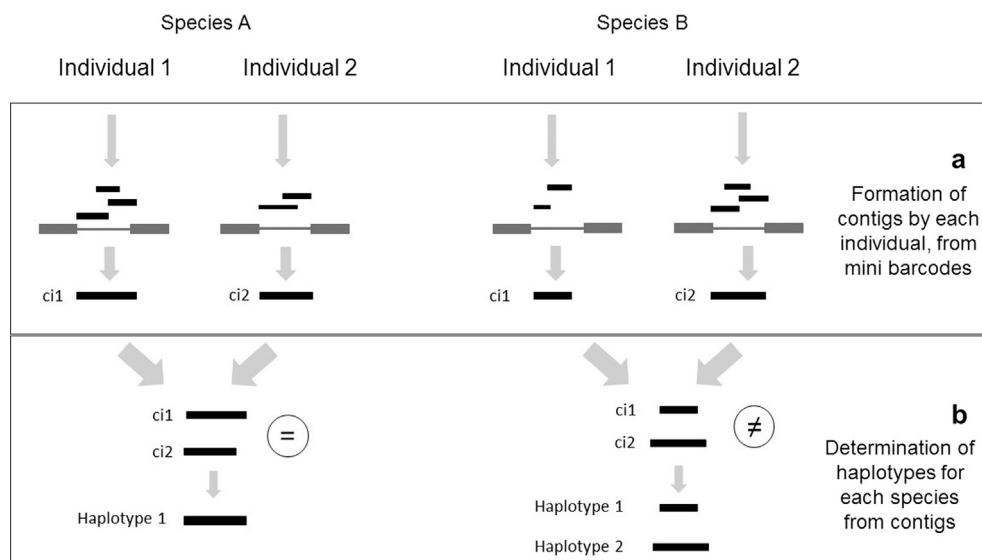
To determine the haplotypic changes (i) according to the spatial distribution of the species, and (ii) over time in the same geographical region, a comparative analysis with the reconstructed haplotypes of the historical specimens was carried out. The comparative analysis consisted of performing an alignment of each historical haplotype (i) with sequences reported in the GenBank database, using the BLASTn platform (NCBI), and (ii) with haplotypes of fresh samples obtained in this study of several species of the genus *Anastrepha*. The comparative analyzes were also done to evaluate the validity of the mini-barcodes in the identification at the species level. This knowledge contributes to both our understanding of how species can change over time and their genetic diversity, taking into account historical specimens deposited in NHC's.

## Results

DNA was obtained from 58 (62.36%) of the 93 processed museum samples (Table 1). In total, 12 species of the genus *Anastrepha*, morphologically identified and available at the MUSENUV (Colombia), were analysed (Table 1). In general it was possible to obtain DNA fragments more than 150 bp in length from the historical specimens, allowing the reconstruction of COI haplotypes with lengths between 171 bp and 632 bp. The size differences of the haplotypes are due to the different amplification rates obtained for the individual fragments and to the fragment combinations that were used for each individual.

#### Obtaining mini-barcodes from museum specimens

Using only the primers and amplification parameters proposed by Van Houdt et al. [13], we achieved amplification rates of 11.3–30.9%; fragment Teph343 (343 bp) was the most highly amplified (30.9%) (Table 2). For some species such as *Anastrepha flavipennis* (Greene), *Anastrepha mucronota* (Stone) and *Anastrepha striata*, no amplification was obtained with these primers, whereas for other



**Fig. 2.** Workflow for the determination of haplotypes. (a) The line between dark gray boxes represents the COI region that is flanked by the primers designed by Folmer et al. [22]. For each individual of different species, different fragments or mini barcodes corresponding to the COI region were obtained. The fragments were overlapped and concatenated to give origin to the contig for each individual of each species, in the figure this corresponds to the black boxes marked with ci1 and ci2. Therefore, a contig was generated for each individual according to the available fragments. (b) The contigs of the individuals of the same species were aligned and compared, and haplotypes were determined for each species according to the polymorphisms found in the overlapped region. Finally, all the haplotypes obtained from historical specimens were contrasted with haplotypes of fresh samples and those available in the GenBank database. In the figure: ci1 (contig of the individual 1) and ci2 (contig of the individual 2).



species such as *A. obliqua*, *Anastrepha pickeli* (Lima), *Anastrepha serpentina* (Wiedemann) and *A. coronilli* (Carrejo & González), the amplification rate for fragment Teph343 varied between 15.4% and 20%. Due to the low amplification performance of these primers, new primers were designed and combined with the primers proposed by Van Houdt et al. [13]. In this way, amplification rates of up to 46.1% were obtained (Table 2); this showed that implementing or designing specific primers for the genus can increase amplification rates because Van Houdt et al. [13] did not include the genus *Anastrepha* in their study. Table 2 and Appendix A (Supplementary Material) show the success rates for amplification and sequencing with each primer combination for each museum specimen analysed. Fragment Anas326d (primer combination H526-L207d, 326-bp DNA fragment) was obtained the greatest number of times, with an amplification rate of 46.1% (Table 2).

The amplification rate also varied among species; *A. leptozona* was the species with the highest amplification rate (73.3% of amplified fragments) (Table 3). The species with the lowest ampli-

**Table 3**

Amplification rates of at least one fragment per species and the number of historical haplotypes identified for each of the 12 species included in this study. Although all the species presented amplified fragments, in some cases sequencing success was not achieved and therefore no haplotypes could be identified.

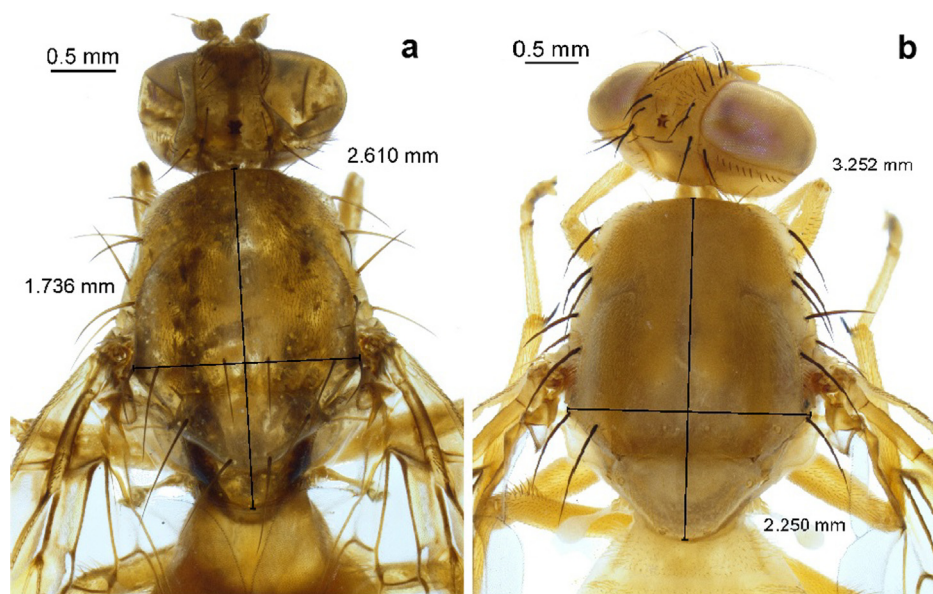
Species	Number of museum individuals processed	Amplification success %	Number of haplotypes identified
<i>A. bahiensis</i>	6	22.77	2
<i>A. coronilli</i>	5	3.33	0
<i>A. distincta</i>	18	31.77	4
<i>A. flavipennis</i>	5	2.85	0
<i>A. fraterculus</i>	2	33.33	1
<i>A. grandis</i>	18	35.46	2
<i>A. leptozona</i>	10	73.33	5
<i>A. mucronota</i>	3	12.5	1
<i>A. obliqua</i>	10	3.91	2
<i>A. pickeli</i>	6	23.8	1
<i>A. serpentina</i>	5	49.76	3
<i>A. striata</i>	5	10	1

fication rates were *A. flavipennis*, *A. coronilli* and *A. obliqua*, with 2.9%, 3.3% and 3.9% amplification, respectively; we were able to obtain haplotypes only for *A. obliqua* (Table 3). In species with low amplification success such as *A. obliqua*, the tissue of the specimens preserved in ethanol (and processed in this study) was deteriorated, mainly in the thoracic region (Fig. 3a). In addition, species with bigger body size, such as *A. leptozona*, had higher amplification rates than other species (Fig. 3b). The tissues showed a hyaline appearance with symptoms of dehydration due to storage in ethanol.

With the adjustment tests of the primers and their corresponding binding region in the COI gene, it was found that the average number of non-complementary base pairs between historical samples and the primers H343 and H526, i.e. the primers with which they obtained the highest amplification rates (Fragments Teph343 and Anas326d, respectively), was 0.5 and 2.58, respectively (Table 4). The H343 primer showed a lower mismatch with respect to the data obtained by Van Houdt et al. [13], who obtained an average of 1.14 mismatches for this primer, which could explain the higher amplification and sequencing rates we obtained, for the Teph343 fragment. The primers designed by Van Houdt et al. [13] had not previously been tested on species of the genus *Anastrepha*, so this result corroborates the importance of this particular primer: H343. For the H526 primer, our mismatch is greater than that obtained by Van Houdt et al. [13], who obtained an average mismatch of 1.79, this being the highest value they reported.

Although the mismatch for H526 was on average 2.58 (Table 4), the results of amplification and sequencing obtained were acceptable (with amplified fragments that could be assembled with others to form contigs and identify haplotypes) so we decided to review in detail the mismatch of this primer with each haplotype (Table 4). We found that the mismatch with the historical haplotype of *A. fraterculus* was 8 bp, and that by removing this measure to calculate the average, primer mismatch was reduced to 1.78, equal to that reported by [13], which coincides with the high amplification rate of the Anas326d fragment.

The primers in our study with the highest number of average mismatches were H2157d with 5.6, followed by L499 with 4. These



**Fig. 3.** Comparison of tissue state in fruit flies of the genus *Anastrepha* collected in 1994. (a) Extension of the thorax of a specimen of *A. obliqua* collected in 1994 in the locality of Chagres, with measures of length of the thorax (2.610 mm) and width of thorax (1.736 mm). The deterioration of the tissue due to dehydration can be appreciated. (b) Extension of the thorax of a specimen of *A. leptozona* collected in 1994 in the locality of Sombrierillo, with measurements of thorax length (3.252 mm) and thorax width (2.250 mm). When comparing the images (a and b) the differences in tissue deterioration can be appreciated even though both samples were collected more than 20 years ago.

**Table 4**

Mismatching of the internal primers in each reported haplotype. The question marks indicate that there is no sequence information available to evaluate the fit of the primers. Accession No. in the GenBank is shown for each sample that presents the haplotype and the size of each haplotype. The type of sample where each haplotype was reported is indicated with HS for historical samples and FS for fresh samples. \*The haplotypes of HS and FS coincide in all the available bases compared.

Species	Code	Accession number	Haplotype Size (bp)	Sample Type	Primers									
					H343	L280d	L499	H2157d	L391d	H246	L344d	H378d	H526	L207d
<i>Anastrepha bahiensis</i>	AbaMus01	MG992505	205	HS	2	0	9	7	1	0	0	0	?	?
	AbaMus02	MG992506	440	HS	0	0	9	7	1	0	0	0	8	1
<i>Anastrepha distincta</i>	AdiMus01/	MG992507 /	613/658*	HS/FS	0	0	4	4	0	0	0	0	1	1
	AdiMus02	MG992508	404	HS	0	0	3	5	0	0	0	0	1	0
	AdiMus03	MG992509	421	HS	0	0	4	5	0	0	0	0	1	0
	AdiMus04	MG992510	171	HS	?	5	?	?	?	?	?	?	?	?
<i>Anastrepha fraterculus</i>	AfrMus01	MG992511	427	HS	0	1	3	?	0	0	0	0	8	0
	Af01_ENG	MK084505	658	FS	0	0	3	0	0	0	0	0	1	0
	Af01_QUE	MK084508	658	FS	0	0	3	0	0	0	0	0	1	0
	Af01_PIC	MK084511	658	FS	0	0	3	0	0	0	0	0	1	0
	Af03_SEV	MK084513	658	FS	1	1	3	0	0	0	0	0	1	0
<i>Anastrepha grandis</i>	AgrMus01	MG992512	613	HS	1	1	4	6	0	3	0	0	2	2
	AgrMus02	MG992513	193	HS	1	1	?	?	0	?	0	0	?	?
<i>Anastrepha leptozona</i>	AleMus01	MG992514	259	HS	?	1	?	?	?	2	?	?	?	0
	AleMus02	MG992515	255	HS	?	1	?	?	?	2	?	?	?	0
	AleMus03	MG992516	261	HS	0	1	?	?	?	2	?	?	?	0
	AleMus04	MG992517	439	HS	0	1	2	?	2	2	1	0	?	0
	AleMus05	MG992518	232	HS	?	1	?	?	?	2	?	?	?	0
	AI02_CHA	MK084478	658	FS	0	1	2	1	2	1	1	0	3	0
	AI01_CHA	MK084477	658	FS	0	1	2	1	2	2	1	0	3	0
	AI07_CHA	MK084482	658	FS	0	1	2	1	2	2	1	0	2	0
	AI08_CHA	MK084483	658	FS	0	1	2	1	2	2	1	0	2	0
	AI10_CHA	MK084485	658	FS	0	1	2	1	2	2	1	0	2	0
	AI04_CHA	MK084480	658	FS	0	1	2	2	2	2	1	0	2	0
	AI03_CHA	MK084479	658	FS	0	1	2	1	2	2	1	0	2	0
	AI11_CAJ	MK084489	658	FS	0	1	2	0	1	0	0	0	2	0
	AI12_CAJ	MK084490	658	FS	0	1	2	0	1	0	0	0	2	0
<i>Anastrepha mucronota</i>	AmuMus01	MG992519	251	HS	0	0	?	?	0	0	0	0	?	?
	An01_CAJ	MK084486	658	FS	1	1	5	0	4	1	1	4	2	2
<i>Anastrepha obliqua</i>	AobMus01	MG992520	194	HS	?	?	3	6	1	?	0	0	0	6
	AobMus02	MG992521	224	HS	?	?	3	6	0	?	?	0	1	7
<i>Anastrepha pickeli</i>	ApiMus01/	MG992522/	632/658*	HS/FS	0	0	3	5	2	0	0	?	3	0
	Ap01_CAJ	MK084491												
<i>Anastrepha serpentina</i>	AseMus01	MG992523	411	HS	1	0	3	5	0	1	0	1	2	?
	AseMus02	MG992524	366	HS	1	0	3	6	0	?	0	1	2	?
	AseMus03	MG992525	421	HS	1	0	3	6	0	2	0	1	2	?
<i>Anastrepha striata</i>	AstMus01	MG992526	229	HS	2	0	?	?	1	0	0	0	?	?

are high values if we consider that these primers have 20 bp and 23 bp respectively. These high mismatch values seem to have an effect on the amplification rate. For example, for the fragment Thep227 (a combination of primers L499 and H2123d) an amplification percentage of only 11.3% was obtained, while for the fragment Anas278d (a combination of primers H2157d and L391d) an amplification percentage of 8.3% was obtained. These two primers (L499 and H2157d) were only used for the amplification of the mentioned fragments.

In general, the primers that had the higher number of mismatches worked better, having higher amplification rates. Some of the primers designed in this study, such as H246, L207d and L344d, had mismatches of 0.96, 0.67 and 0.30 respectively. These mismatches are considered low, and they explain the amplification rates of the fragments Anas266d, Anas326d and Anas366d greater than 29%. This demonstrates the need to design specific primers to improve performance in the amplification of museum specimens.

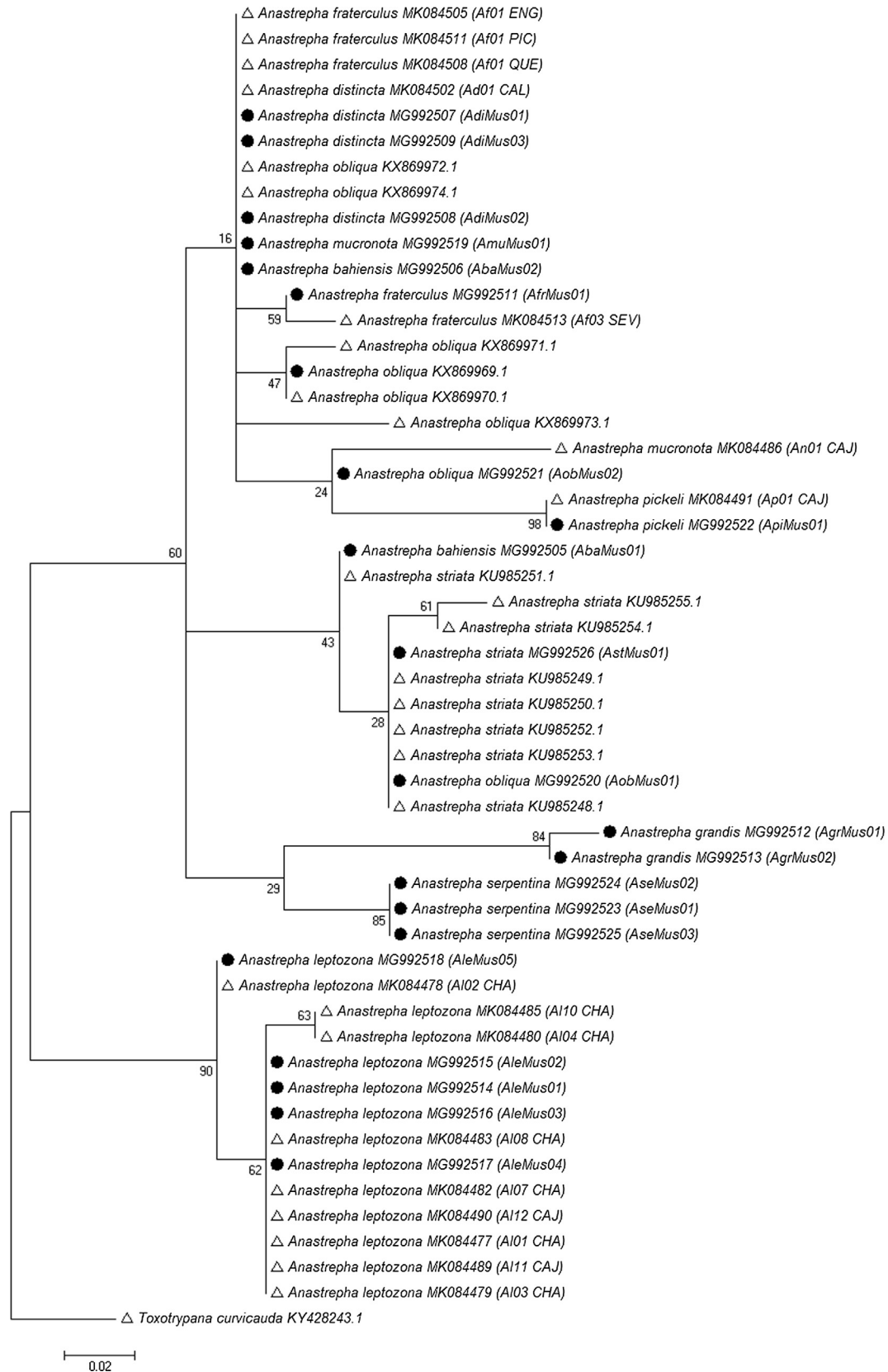
#### Use of the mini-barcodes of museum specimens for the molecular identification of species and the assessment of diversity

NJ analysis (Fig. 4) corroborated that the sequences obtained from historical specimens were generally grouped with sequences of the same species obtained from fresh samples. This result indicates a good molecular identification of museum specimens using mini-barcodes for the reconstruction of haplotypes. With the

haplotypes of fresh and museum specimens, it was possible to establish phylogenetic relationships at the intra- and interspecific levels. For the species *A. grandis*, *A. serpentina*, *A. striata* and *A. pickeli*, monophyletic groups were formed (Fig. 4). For *A. leptozona*, one group was formed and no haplotype was shared between the historical and fresh samples. Finally, for *A. bahiensis*, *A. fraterculus*, *A. obliqua*, *A. distincta* and *A. mucronota*, no monophyletic groups were formed. In *A. fraterculus*, *A. obliqua*, *A. distincta* and *A. bahiensis*, it is common to find paraphyletic relationships, because these species have been classified as a cryptic species complex.

When we compared the haplotypes obtained from the museum specimens with the haplotypes reported in sequence databases (GenBank - NCBI), we obtained identities between 95% and 100%. The results for each species are presented below.

- ***A. bahiensis***: Two haplotypes were found, one of 205 bp (Accession No. MG992505) and another of 440 bp (Accession No. MG992506). The 205-bp haplotype displayed 95% identity with sequences of *A. obliqua* and *A. schultzi* (*fraterculus* group). On the other hand, the 440-bp haplotype displayed 97% and 99% identity with sequences of *A. bistrigata* and *A. striata* (*serpentina* group), respectively. The results showing identities between 95% and 99% with the sequences of different species groups could be due to the short length of the sequences in Accession No. MG992505.



**Fig. 4.** NJ analysis demonstrates the utility of mini-barcodes to obtain haplotypes of historical specimens that allow the discrimination and identification of species. This analysis is based on 22 haplotypes of historical samples and 14 haplotypes of fresh samples. The haplotypes are identified with black circles when they correspond to historical specimens, and with white triangles when they correspond to fresh specimens. The analysis was performed using the MEGA 7.0.2 software with the Kimura-2-parameter nucleotide substitution model.



- ***A. obliqua*, *A. distincta* and *A. fraterculus***: The coincidence was 99% with sequences reported for *A. obliqua*, *A. fraterculus*, *A. distincta* and *A. canalis*. All the species mentioned belong to the *fraterculus* group. The species were kept together in the same clade.
- ***A. grandis***: Two haplotypes were obtained, one of 193 bp and another of 613 bp. For both, identities greater than 98% with sequences of the same species deposited in GenBank were obtained. Identity with sequences from other species was below 95%.
- ***A. leptozona***: We found a minimum identity of 95% (for both historical and fresh specimens) with sequences of the same species deposited in GenBank. Even the shortest haplotypes (232 bp) had identities between 95% and 99% with sequences of the same species. We reconstructed 5 haplotypes from 8 museum specimens, and 9 haplotypes from 10 fresh specimens. Despite the high number of haplotypes in so few samples, the divergence between haplotypes (K2P) was not higher than 0.3% (Supplementary Material, Appendix B).
- ***A. mucronota***: The haplotype found in historical specimens (251 bp) (Accession No. MG992519) displayed greater identity with sequences of the *fraterculus* group (>95%). The incongruence between the taxonomic and molecular identifications may be due to the short length of the sequence. However, we must emphasise that no *A. mucronota* sequences are available in GenBank. The haplotype obtained from fresh samples of *A. mucronota* coincides with some species of the same group (*mucronota* group), with identities lower than 92%. In this study, we found *A. leptozona* and *A. mucronota* infesting sapote fruits of the same tree. However, the fruits were not individualised, and whether the individuals emerged from the same fruits or from different fruits is unknown.
- ***A. pickeli***: An identity of 95% was obtained for the *A. pickeli* species when the haplotype obtained in this study was compared with haplotypes reported in GenBank. For this haplotype of *A. pickeli*, we found identities between 97% and 100% with *A. manihoti* sequences. However, the coverage was between 88% and 100% with sequences of *A. manihoti* available in GenBank.
- ***A. serpentina***: We found an identity of 98–99% only with sequences of the same species reported in GenBank.

Given the low number of processed samples, it is difficult to perform temporal diversity analysis. However, to determine whether the reconstruction of haplotypes from historical samples can be used for this purpose, we analysed our available material through a temporal comparison. The COI sequences of fresh and museum specimens of four *Anastrepha* species (*A. distincta*, *A. pickeli*, *A. nunezae* and *A. leptozona*) were compared. For *A. distincta*, 13 historical specimens with sequence sizes (fragments or contigs) between 171 bp and 607 bp were processed. Four haplotypes, one of which matched the only haplotype of the fresh samples, were determined among these 13 museum samples. For *A. pickeli*, two contigs of 276 bp and 632 bp were obtained from museum samples. The two contigs corresponded to a single haplotype that matched the haplotype found in fresh samples. For *A. mucronota* (museum sample), we obtained only one contig of 251 bp. This haplotype did not match any of the haplotypes found in the fresh samples (three in total), while in the NJ analyses, it was grouped with samples of the *fraterculus* group. The BLAST analysis also yielded an identity of 98–99% with sequences reported for species of the same group. The sequences reported in this study for *A. mucronota* represent the first entry of the COI gene of this species in GenBank (Accession No. MG992519).

Finally, for *A. leptozona*, five historical haplotypes were found; none of these is shared with the fresh samples, for which new and different haplotypes were determined. This haplotypic difference may suggest high diversity of *A. leptozona* in the sampling area.

## Discussion

### Mini-barcoding of museum specimens of genus *Anastrepha*

Using the primers described in Folmer et al. [22], we obtained unsuccessful amplification rates. However, primer annealing was used in the internal regions to obtain the COI fragments. Although all study samples were stored in 70% ethanol as a preservative, the percentage of ethanol may have decreased over time since it is more volatile than water. This may have resulted in degradation of the nucleic acids. It has been shown that DNA degradation in ethanol at low concentrations occurs relatively rapidly [25]. For example, Reiss et al. [26] found that Coleoptera specimens preserved in ethanol (95%) showed high DNA degradation after six weeks, showing that ethanol is useful only for short-term preservation.

Additionally, a difference in DNA degradation by species was observed. For example, in *A. obliqua*, which had low amplification rates, most of the samples were in poor condition; the tissue had dehydrated, mainly in the thoracic region, and had a hyaline appearance. In contrast, in *A. leptozona*, which had the highest amplification rate, the samples were better conserved and in a state similar to fresh samples. Likewise, Patel et al. [7] reported different amplification rates with the same pair of primers when evaluating different orders of birds. Obtaining different amplification rates with the same pair of primers in different species could be (among other possible reasons) due to DNA degradation and fragmentation; this can lead to (i) the production of increasingly smaller fragments, allowing only short amplicons of 150–300 bp to be obtained, and (ii) a lack of stable sites for primer binding, causing an imbalance in the sequence of interest.

Considering that a decrease in fragment size due to degradation is inversely proportional to amplification success, degradation may impact the success of amplification in museum samples [13]. However, we observed that fragments of 343 bp and 326 bp had the highest amplification rates, in contrast to smaller fragments between 200 and 300 bp. Other authors, such as Hernández-Triana et al. [11], reported similar results with 271 specimens from 36 species of black flies (Simuliidae) with collection dates ranging from 1937 to 2004. Hernández-Triana et al. [11] found that fragments of 388 bp had a higher amplification success rate (25.8%) than that of fragments of 307 bp (11.8%). Obtaining higher amplification rates with longer fragments suggests that the differences in amplification rates between primers and species are due to other factors, such as the binding of the primers to the target sequence [11].

The adjustment of the primers has an important role in museum DNA amplification. In fact, our results show an inverse correlation: the higher the number of mismatches, the lower the amplification rate for historical samples. The primers used in this study are specific for the family Tephritidae [13] and for the genus *Anastrepha*; this is an important factor that reduced mismatching and increased the probability of obtaining amplicons from degraded DNA. For the fragments Teph343 and Anas326d, we used the primers H343 and H526, respectively; these primers, which were designed by Van Houdt et al. [13], delivered amplification rates of 30.9% and 46.1%, respectively. Thus, primers H343 and H526 seem to be quite useful as universal primers in the family Tephritidae. However, other

primers did not work optimally, and the design of specific primers allowed us to increase the amplification rates in regions that could not be amplified with less specific primers. For example, using primers designed for Tephritidae we obtained an amplification rate of 11.3% for the fragment Teph227 (primer pairs L499 - H2123d) with a length of 227 bp. By designing a specific primer (L344d) closer to the binding region of the first L499, we obtained a longer amplicon (366 bp) and a higher amplification rate (29.4%). The difference in the results obtained using primers L499 and L344d (Table 4) (4 and 0.3, respectively) shows how the adjustment of the primers to specific DNA regions confers greater amplification success in historical samples. Therefore, we emphasise the importance of designing specific primers in cases in which primers targeting other organisms lead to low amplification rates. Additionally, the design of specific primers allowed us to amplify fragments of greater length for some samples; this is considered an important feature in preventing unwanted DNA co-amplification [11].

Although the amplification rate was not 100%, results similar to ours were obtained in other studies. Dillon et al. [27] tested various methods of DNA extraction and found different amplification rates for different species of Hymenoptera (Ichneumonidae) that had been preserved for different lengths of time. The results of Dillon et al. [27] agree with our results, in which we obtained different amplification rates for species of the genus *Anastrepha*. Most recently, studies such as those of Hernández-Triana et al. [11] have obtained amplicons from 45% of processed samples of Coleoptera, a rate similar to that reported in this study.

Based on the literature and on our results, it should be noted that the quality of the degraded DNA varies depending on the preservation method and the species analysed. It is important to consider adjusting the primers to the amplification regions of interest in historical samples to increase the probability of hybridisation of the primer with the DNA to be amplified. Appropriate specimen preservation for future analyses involving molecular methods is also important, and the use of ethanol preservation at high concentrations should be considered.

Finally, other technologies could be used to increase the amplification success rate in historical samples. We have already shown that DNA amplification from museum specimens using classical PCR mini-barcodes and sequencing techniques is affected by DNA degradation. We are able to obtain amplification rates of approximately 50%, and the amplification rates differ depending on the species and primer sets used. The application of technologies such as high-performance sequencing or next-generation sequencing (NGS) could solve these problems. The greatest advantage of the use of NGS technologies with historical samples is the low amount of template DNA required [28]. In addition, the use of these technologies reduces the time required for obtaining the DNA region of interest and incurs lower total cost than does traditional sequencing [29]. However, NGS technologies are still more expensive. An interesting perspective would be provided by analysing the amplification rates of the primers designed in this study by high-performance sequencing.

#### Mini-barcoding allows molecular identification of species and the evaluation of *Anastrepha* diversity

Mini-barcode amplification allowed us to obtain fragments of the COI gene that were long enough to permit the reconstruction of haplotypes and the identification of species from historical specimens. The NJ tree showed how haplotypes obtained from historical samples can be used to effectively determine the relationships among species because, in general, haplotypes of historical and fresh samples of the same species were grouped independently of the length of the fragments that generated the haplotypes. In general, the results of the NJ analyses are consistent with the results obtained by other authors [8], who were able to identify species using fragments up to 135 bp

in length and to demonstrate evolutionary relationships between individuals using sequences of various sizes, as in this study. Hernández-Triana et al. [11] also found consistency between historical and fresh specimens when they performed NJ phylogenetic analyses and confirmed the use of mini-barcodes for species delimitation and for the study of interspecific relationships.

In BLAST comparisons with the sequences available in GenBank, identities of 95–100% were obtained. The identities obtained for most of the species exceeded the 3% divergence threshold proposed by Hebert et al. [5], which permits more accurate identification from barcodes obtained in museum specimens. In some cases, the mini-barcodes did not allow determination of the taxonomic identity of the specimen. For example, it was not possible to differentiate with certainty the sequences of the species of the *fraterculus* group. The haplotypes obtained in this study for the species *A. obliqua* and *A. fraterculus* had 99% identity with GenBank sequences belonging to several species, including *A. fraterculus*, *A. obliqua*, *A. canalis* and *A. distincta*. Scally et al. [30] showed that the use of mitochondrial genes in the *fraterculus* group does not allow total separation of the species; they hypothesized that introgression and hybridisation occur between different species of this group. This would suggest the formation of paraphyletic clades. Even the use of nuclear genes has only allowed the separation of the species *A. obliqua* from the remaining species in the *fraterculus* complex [30]. Mengual et al. [31] used three nuclear regions and three mitochondrial regions and also found monophyletic groups for the species mentioned. Finally, Barr et al. [32] evaluated a 539-bp segment of the COI gene and were unable to delimit species such as *A. fraterculus* and *A. obliqua*. Therefore, longer fragments should be used to assess species identity, and the use of local haplotype databases could confer greater confidence for species delimitation in this group.

For species of *A. mucronota*, we found incongruence in molecular identification in which the historical haplotype is more closely related to haplotypes of species of the *fraterculus* group in historical and fresh samples. This haplotype of *A. mucronota* had a length of 251 bp; therefore, we propose that it may be a very short fragment of a preserved region, which hinders a successful identification at the molecular level. In recent samples, we found another haplotype of *A. mucronota* that had an identity of less than 92% with sequences belonging to specimens of the *mucronota* group. However, because no sequences of *A. mucronota* for the COI gene are reported in GenBank, a comparison of intraspecific diversity was not possible. For this species, it should be considered that the historical specimens used in this study were identified in 1994 as *A. nunezae*. Norrbom & Korytkowski [33] reviewed the specimens and concluded that *A. nunezae* is a synonym of *A. mucronota*. The fresh specimens collected in 2015 morphologically agree with the historical specimens available in the MUSENUV.

In terms of identification, *A. pickeli* offers an interesting case for analysis. We obtained only one haplotype that, in addition, was the same for historical and fresh specimens. Through BLAST analysis, this haplotype could be identified as *A. pickeli* but with only 95% identity. Bomfim et al. [34] reported a high degree of genetic variation among populations of this species, and the variation is congruent with their geographic distribution, suggesting that a cryptic species complex may exist. Well-differentiated clades from Brazil and Paraguay have been defined, also with identities of 95%. Our data corroborate the hypothesis of high diversity within what is considered the nominal species, *A. pickeli*, and suggest that a review of this species throughout its distribution could provide more evidence regarding the population relationships that exist within it and its taxonomic status and resolve whether it comprises a single species or a species complex. The *A. pickeli* haplotype also had an identity between 88% and 100% with *A. manihoti* sequences. *A. manihoti* and *A. pickeli* are sister species that mainly infest the stem and fruit of cassava (*Manihot esculenta*), respectively. The historical spec-

imens used in this study were collected using a trap, while the fresh specimens were obtained from cassava fruits harvested in the field. Therefore, we consider that there could be two possible scenarios: (i) since they are sister species and therefore belong to the same group of species, it is possible that barcodes within the COI region do not allow the delimitation of *A. pickeli* and *A. manihoti*, a case similar to that of the *fraterculus* group; and (ii) introgression between *A. pickeli* and *A. manihoti* could also occur in the wild, meaning that they can share mitochondrial haplotypes.

The availability of samples of the species *A. leptozona* from the same geographic region but with a temporal separation of 21 years allowed us to have an idea of the diversity of haplotypes. Five historical haplotypes and seven recent haplotypes were found in 15 analysed specimens. The historical haplotypes were all different from the recent haplotypes. This separation could suggest that there is wide diversity in the sampling area and/or that genetic diversification of this species may have occurred in the collection area. Thus, it is emerging as an interesting species in studies of diversity and population dynamics in regions of high fruit production such as southwestern Colombia.

In general, our results highlight the importance and usefulness of historical samples deposited in NHCs for studies of the genetic diversity and population dynamics of species. In recent years, the importance of the data that can be obtained by analysing historical samples around the world has been highlighted [7]. In this work, it is evident that it is possible to use historical samples not only for species identification but also for the analysis of population dynamics (if the number of samples allows it).

## Conclusions

Our results are successful compared to other results found in the literature; we obtained amplification rates of up to 46.1% in museum specimens, and haplotypes between 171 bp and 632 bp could be reconstructed. In some cases, we achieved species identification, demonstrating that this is a good methodology for obtaining barcodes for the identification of animal life. However, reconstruction of the COI region using mini-barcodes still faces the traditional complications associated with the use of this genetic marker, and some species cannot be delimited. It is important to always consider the quality of the DNA and possible modification of the primers; we suggest the use of specific, degenerate primers to increase the probability of amplification of the regions of interest. Obtaining data from historical samples is possible and, among other things, facilitates the performance of phylogenetic, diversity, and population dynamics analyses, so it is expected that sequences obtained from historical samples will be included in future studies. However, it is important to plan for adequate preservation of the specimens deposited in the collections, and we urge the NHCs to implement systems to maximise the possibility of using samples for molecular analysis. This would increase the use of available resources and expand the knowledge of the population dynamics of species over time.

## Conflict of interest

*The authors have declared no conflict of interest.*

## Compliance with Ethics Requirements

*All Institutional and National Guidelines for the care and use of animals (insects) were followed.*

## Acknowledgements

This research was internally funded by Universidad del Valle with the project entitled “Barcodes of fresh and museum specimens of the fruit fly *Anastrepha striata* and *Anastrepha obliqua*: Contribution to the knowledge of its demographic history and population structure in Valle del Cauca” with information code 7976. We thank the Images Laboratory of the Postgraduate Biology Department of “Universidad del Valle” and the Biologist Juan Felipe Ortega for taking photographs of the specimens of the *Anastrepha* genus. Thanks to Dr José Julián Tavera Vargas of the Universidad del Valle and Dr Enrique Arbeláez Cortés of the Universidad Industrial de Santander for their comments.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jare.2018.11.001>.

## References

- [1] Lister AM, Brooks SJ, Fenberg PB, Glover AG, Johnson KKG, Lister AM, et al. Natural history collections as sources of long-term datasets. *Trends Ecol Evol* 2011;26(4):154–5.
- [2] Wandeler P, Hoeck PEA, Keller LF. Back to the future: museum specimens in population genetics. *Trends Ecol Evol* 2007;22(12):634–42.
- [3] Holmes MW, Hammond T, Wogan GOU, Walsh RE, LaBarbera K, Wommack EA, et al. Natural history collections as windows on evolutionary processes. *Mol Ecol* 2015;25(4):864–81.
- [4] Zimmermann J, Hajibabaei M, Blackburn DC, Hanken J, Cantin E, Posfaj J, et al. DNA damage in preserved specimens and tissue samples: a molecular assessment. *Front Zool* 2008;5:18.
- [5] Hebert PDN, Ratnasingham S, deWaard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc R Soc B* 2003;270:S96–9.
- [6] Hajibabaei M, McKenna C. DNA Mini-barcodes. *Methods Mol Biol* 2012;858:339–53.
- [7] Patel S, Waugh J, Millar CD, Lambert DM. Conserved primers for DNA barcoding historical and modern samples from New Zealand and Antarctic birds. *Mol Ecol Resour* 2010;10:431–8.
- [8] Sonet G, Breman FC, Lenglet G, Louette M, Montanes G, Nagym ZT, et al. Applicability of DNA barcoding to museum specimens of birds from the Democratic Republic of the Congo. *Bonn Zool Monogr* 2011;57:117–31.
- [9] Hajibabaei M, Smith MA, Janzen DH, Rodriguez JJ, Whitfield JB, Hebert PDN. A minimalist barcode can identify a specimen whose DNA is degraded. *Mol Ecol Notes* 2006;6:959–64.
- [10] Franco E, Arias MC. *Cytochrome C oxidase I* primers for corbiculate bees: DNA barcode and mini-barcode. *Mol Ecol Resour* 2013;13:844–50.
- [11] Hernández-Triana LM, Prosser SW, Rodríguez-Perez MA, Chaverrí LG, Hebert PDN, Gregory TR. Recovery of DNA barcodes from blackfly museum specimens (Diptera: Simuliidae) using primer sets that target a variety of sequence lengths. *Mol Ecol Resour* 2014;14:508–18.
- [12] Büsse S, Von Grunmkow P, Mazanec J, Tröster G, Hummel S, Hörnschemeyer T. Note on using nuclear 28S rDNA for sequencing ancient and strongly degraded insect DNA. *Entomol Sci* 2017;20:137–41.
- [13] Van Houdt JKJ, Breman FC, Virgilio M, DeMeyer M. Recovering full DNA barcodes from natural history collections of Tephritid fruitflies (Tephritidae, Diptera) using mini barcodes. *Mol Ecol Resour* 2010;10:459–65.
- [14] Selivon D, Perondini ALP, Morgante JS. A genetic-morphological characterization of tow cryptic species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae). *Ann Entomol Soc Am* 2005;98(3):367–81.
- [15] Martínez JO. Nuevos registros en el género *Anastrepha* (Diptera: Tephritidae) para Colombia. *Rev Colomb Entomol* 2007;33(1).
- [16] Carrejo NS, González R. Lista preliminar de las moscas de la fruta del genero *Anastrepha* (DIP.: TEPHRITIDAE) en el departamento del Valle del Cauca Cali, Colombia. *Bol Mus Ent Univ Valle* 1994;2(1,2):85–93.
- [17] Velasco-Cuervo SM, Toro-Perea N. Análisis del gen COI y su contribución a la filogenia y a la diversidad genética espacial y temporal del género *Anastrepha* (Diptera: Tephritidae), usando especímenes frescos, de museo y bases de datos de secuencias. Cali (Colombia): Tesis de Maestría. Universidad del Valle; 2016.
- [18] Aguirre-Ramírez EJ, Velasco-Cuervo SM, Gallo-Franco EJ, González R, Carrejo NS, Toro-Perea N. Genetic diversity and population structure of *Anastrepha obliqua* (Diptera: Tephritidae) in southwestern Colombia. *Entomol Exp Appl* 2017;164(3):291–304.
- [19] Gallo-Franco JJ, Velasco-Cuervo SM, Aguirre-Ramírez E, González-Obando R, Carrejo NS, Toro-Perea N. Genetic diversity and population structure of *Anastrepha striata* (Diptera: Tephritidae) in three natural regions of southwestern Colombia using mitochondrial sequences. *Genetica* 2017;145(1):79–89.

- [20] Meeyen K, Nanork P, Pramual P. Population structure, population history and DNA barcoding of fruit fly *Bactrocera latifrons* (Hendel) (Diptera: Tephritidae). *Entomol Sci* 2013. <https://doi.org/10.1111/ens.12049>.
- [21] Ossa PA, Giraldo JM, López GA, Díaz LG, Rivera FA. Colecciones biológicas: Una alternativa para los estudios de diversidad genética. *Boletín Científico Centro de Museos Museo de Historia Natural* 2012;16(1):143–55.
- [22] Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 1994;3(5):294–9.
- [23] Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. *Bioinformatics* 2007;23(10):1289–91.
- [24] Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3 – new capabilities and interfaces. *Nuclei Acids Res* 2012;40(15):e115.
- [25] Li Y, Zhou X, Feng G, Hu H, Niu L, Heber PDN. COI and ITS2 sequences delimit species, reveal cryptic taxa and host specificity of fig-associated *Sycophila* (Hymenoptera, Eurytomidae). *Mol Ecol Resour* 2009;10(1):31–40.
- [26] Reiss RA, Schwert DP, Ashworth AC. Field preservation of Coleoptera for molecular genetic analyses. *Environ Entomol* 2015;24:716–9.
- [27] Dillon N, Austin AD, Bartowsky E. Comparison of preservation techniques for DNA extraction from hymenopterous insects. *Insect Mol Biol* 1996;5:21–4.
- [28] Prosser SWJ, Dewaard JR, Miller SE, Hebert PDN. DNA barcodes from century-old type specimens using next-generation sequencing. *Mol Ecol Res* 2016;16:487–97.
- [29] Shokralla S, Gibson JF, Nikbakht H, Janzen DH, Hallwachs W, Hajibabaei M. Next-generation DNA barcoding: using next-generation sequencing to enhance and accelerate DNA barcode capture from single specimens. *Mol Ecol Res* 2014;14:892–901.
- [30] Scally M, Into F, Thomas DB, Ruiz-Arce R, Barr NB, Schuenzel EL. Resolution of inter and intra-species relationships of the West Indian fruit fly *Anastrepha obliqua*. *Mol Phylogenetic Evol* 2016;101:286–93.
- [31] Mengual X, Kerr P, Norrbom AL, Barr NB, Lewis ML, Stapelfeldt AM, et al. Phylogenetic relationships of the tribe Toxotrypanini (Diptera: Tephritidae) based on molecular characters. *Mol Phylogenetic Evol* 2017;113:84–112.
- [32] Barr NB, Ruiz-Arce R, Farris RE, Silva JG, Lima KM, Dutra VS, et al. Identifying *Anastrepha* (Diptera: Tephritidae) species using DNA barcodes. *J Econ Entomol* 2018;9(11):405–21.
- [33] Norrbom AL, Korytkowski CA. New species of and taxonomic notes of *Anastrepha* (Diptera: Tephritidae). *Zootaxa* 2011;2740:1–23.
- [34] Bomfim ZV, Lima KM, Silva JG, Costa MA, Zucchi RA. A morphometric and molecular study of *Anastrepha pickeli* Lima (Diptera: Tephritidae). *Neotrop Entomol* 2011;40(5):587–94.