The Pennsylvania State University

The Graduate School

College of Agricultural Sciences

# A PHYLOGEOGRAPHIC STUDY OF *ANASTREPHA OBLIQUA* AND *A. LUDENS* (DIPTERA: TEPHRITIDAE)

A Dissertation in

Entomology

by

Raul Ruiz



Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

December 2010

The dissertation of Raul Ruiz was reviewed and approved by the following:

Bruce A. McPheron
Professor of Entomology
Dean for the College of Agriculture
Dissertation Adviser
Chair of Committee

Michael Saunders
Professor of Entomology

Shelby Fleischer
Professor of Entomology

David M. Geiser
Associate Professor of Plant Pathology

Norman B. Barr
Molecular Biologist
Special Member
USDA APHIS Center for Plant Health Science and Technology

Gary W. Felton
Graduate Program Office for Department of Entomology

Signatures are on file in the Graduate School.

## ABSTRACT

This thesis examines the phylogeographic variation of two economically important Tephritid pest species, *Anastrepha ludens* (Mexican fruit fly) and *A. obliqua* (West Indian fruit fly). Both of these species are formidable fruit pests that are associated with many plants and are considered economically important pests. Both species have a wide geographic distribution with occasional infestations occurring in the US. This research evaluates the utility of a sequence-based approach for two mitochondrial genes and nuclear markers. We assess these markers as genetic tools for revealing historical processes, for delineating populations, and examining for population structure.

The Mexican fruit fly is a major pest of orange and grapefruit occurring in Mexico and C. America. I analyzed 543 *A. ludens* individuals from 67 geographic collections and found one predominant haplotype occurring in the majority of specimens. I observed 68 haplotypes in all and see a substantial difference among haplotypes belonging to northern and southern groups of collections. The results show that COI and ND6 are useful for phylogeographic studies of *A. ludens* and may be used for pathway analysis.

The West Indian fruit fly is a major pest of mango and plum contributing to millions of dollars in damage each year throughout its geographic range. The current geographic distribution for *A. obliqua* includes areas stretching from N. Mexico, C. America, and S. America as far and N. Argentina, as well as the Caribbean. West Indian fruit fly specimens were examined for genetic variation at two spatial scales. The first study is of specimens gathered from five different host species over three years in the Mexican state of Veracruz represent those geographic collections examined at a local scale. Additional collections gathered from throughout the geographic range represent the second study of this species.

In Veracuz, numerous specimens were examined for variation that may be due to the influence by ecological factors using mtDNA and nuclear markers. Here, five species of infested fruit was collected over three years and fruit fly larvae were allowed to emerge and mature to adults. I sequenced two genes (COI and NDI) and used 17 nuclear loci from these flies and revealed that the populations examined in Veracruz represent one panmictic population. The value of this information is especially important to the management of this pest because it reveals that one management strategy is sufficient. Additionally, this information is used in knowing the diversity of this pest at this window in time.

Fifty-four collections of *A. obliqua* were sequenced in a broader geographic study. Mitochondrial regions from the COI and ND6 genes reveal six populations throughout the distribution of this fly. I recovered 61 haplotypes from the 349 individuals gathered from 54 localities. Many of these haplotypes were either site specific or region specific making this a useful tool for pathway analysis. There were substantial differences seen among the lineages represented by the haplogroups. The amount of divergence seen among these four haplogroups is in line with differences seen among other closely related taxa I analyzed. This would suggest that a review of taxanomic keys is needed. Also, additional analyses perhaps with unlinked loci are needed to evaluate the level of interaction among different lineages that co-occur.

**TABLE OF CONTENTS**

**LIST OF FIGURES**

## LIST OF TABLES

## ACKNOWLEDGMENTS

I am indebted to my thesis committee, Drs. Bruce McPheron, Norman Barr, Shelby Fleischer, David Geiser, and Michael Saunders for the roles they played in inspiring, supporting, and guiding my thesis work. My special thanks to Bruce and the College of Agriculture for the financial support for my education, for gentling nudging me along improving my understanding of evolutionary biology and molecular ecology and more importantly for exemplifying excellence as a scientist and person. I am a better scientist and person because of you Bruce.

I am grateful to Drs. Phillip Berger, Gordon Gordh, Susan McCombs, and Paul Parker at the USDA APHIS Center for Plant Health Science and Technology who were instrumental in providing funding, guidance, and support for this extended sabbatical. They have been very gracious in allowing me to complete this professional and personal goal. I appreciate their confidence in me as a scientist and in my efforts to achieve this degree.

It was providence that led me to Penn State to undergo this chapter in my life. I firmly believe that divine inspiration has functioned as the tangible force behind every portion of this thesis. I have formed new friendships within the Entomology Department and among highly respected colleagues. I am grateful to Dr. Gary Felton and the Entomology Department for their substantial contributions to this work and for use of lab and office space.

My most sincere gratitude is for those who sacrificed the most during this work. It is to them that I dedicate this thesis. I am grateful to my wife, Oralia, who was always a motivating and an unwavering spirit of support. She not only served as a loving wife and soul mate, she also provided excellent technical support for generating and analyzing these data. I am grateful to my daughter Mindy, who has been an inspiration and source of joy and pride for me all her life. She and my wife are my treasures. I am also grateful to my family, especially my brothers Ralph and Robert, who

have instilled in me the value of an education, my sister Rose, who has loved and raised me like a mother, and my brothers Ranulfo and Rudy and sisters Rosario, Janie, and Irma who have always provided me fine examples of Godly men and women.

# Chapter 1: Introduction

This thesis will address three major contributions. First, I will examine the genetic diversity of *Anastrepha ludens* individuals collected from across its geographic range. I will use various methods that delineate populations and may provide insight into historical processes that shape the distribution of populations. In using a sequence-based approach, I will develop methods for distinguishing among regional genetic types and generate information that may be used by action agencies, pest management groups, and in biodiversity studies. Second, I will examine collections of *A. obliqua* gathered within a localized area in Veracruz, Mexico using a multi-locus approach. Here, a suite of markers will be used to determine whether the genetic variation observed may be attributed to influence by spatial, temporal, or host-association. This information may provide insight into the breeding structure, identify ecological traits that contribute to population structure, and reveal the potential of this species for adapting to various hosts. Third, I will use a sequence-based approach for examining the genetic diversity of *A. obliqua* across the geographic range. This information will permit the delineation of geographic populations and help improve management of this pest across the wide geographic distribution. Additionally, the methods developed here may be of use for studies concerned with measuring species richness and ecological trends.

## THE IMPORTANCE OF *ANASTREPHA*

The genus *Anastrepha* (Schiner) contains approximately two hundred described species belonging to the family Tephritidae. These are further grouped into the subfamily Trypetinae and tribe Toxotrypanini. The genera representing this tribe include *Anastrepha, Toxotrypana,* and *Hexachaeta*. Species within *Anastrepha* are further divided into 18 species groups (Norrbom et al.

2003, 1999). These "true" fruit flies are endemic to the Americas and occur mainly in the subtropical and tropical areas of the western hemisphere (Stone, 1942; Aluja 1994) and are most likely to have originated in South America (Hernandez-Ortiz 1992). *Anastrepha* has been considered the most economically important genus in the Neotropics (Aluja 1994; Aluja et al. 1996; Norrbom & Foote 1989) with host plants reported from 143 genera in 54 plant families (Norrbom et al. 1999). White & Elson-Harris (1992) has listed 15 species as significant pests and 28 others have been found to attack economically important plants to a lesser degree from this genus. Seven of these significant pests are considered economically important species and little is known of their biology and population structure (Aluja 1994). These seven are *A. ludens* (Mexican fruit fly), *A. obliqua* (West Indian fruit fly), *A. fraterculus, A. grandis, A. serpentina, A. striata,* and *A. suspensa.*

The economic impact to countries by pests within the genus *Anastrepha* is attributed to direct insect damage, costs in control and disinfestations to commodities, as well as losses to markets as a result of trade restrictions (Bustos at al., 1993). The development of effective pest control programs is needed and would benefit through an understanding and evaluation of evolutionary processes (Shochat 2006; Müller-Schärer 2004) and biology characteristic to pest species.

Mitochondrial and nuclear DNA from *A. ludens* and *A. obliqua* will be used for examining the demographic processes and relationships among individuals from populations sampled for this study. This information will enhance our knowledge of the mechanisms that contribute to population structure for these species. A thorough understanding of fauna adds value to the appreciation of species diversity, richness and conservation. It provides insight to plant/herbivore relationships, knowledge on the effects to plant fitness (Novotny et al. 2005), and quantification to effects from niche fragmentation. The in-depth look at populations may even enhance our awareness of impacts from biotic and abiotic variation due to human-related issues such as global

warming (Bale et al. 2002). As seen in the genus *Rhagoletis*, it may also provide additional evidence of speciation in sympatry (Bush 1969) and hybridization (Schwarz 2005). In the case of *Rhagoletis*, host race formation is explained by allochronic variation (Smith 1988; Feder et al. 1993) which has also been reported in *A. obliqua* (Henning & Matioli 2006; Aluja & Birke 1993). Within agriculture, the delineation of species and accurate identification of individuals to species, strain, biotype, or population-level is essential to pest management programs. However, if improvement of pest management strategies is being considered, the significance of genetic variation as a function of geography and host must be taken into account. These factors may be of particular importance when taking into account methods such as sterile insect technique (SIT) which have successfully been used to control pests (Klassen & Curtis 2005). Knowing the amount of variation is essential to SIT programs since any substantial genetic variation may translate to differences in performance among distinct and biologically incompatible genetic types. This study explores the use of molecular methods for examining genetic variation in species and populations and provides a glimpse into a wide and localized geographic area and examines these collections for selection by ecological factors.

## USE OF MOLECULAR METHODS

Mitochondrial DNA sequences have been used successfully on tephritids for resolving inter-specific relationships (Barr and McPheron 2006; McPheron et al. 1999; Han and McPheron 1999), intra-specific structure (Lanzavecchia et al. 2008; McPheron et al. 1994), and to develop diagnostic methods for species ID (Barr et al. 2006) that may be used for the pathway analysis of populations. Most methods, however, are associated with old world species such as *Bactrocera* (Aketarawong et al. 2007; Armstrong et al. 1997; Nardi et al. 2005) and *Ceratitis* (Barr et al. 2006; Gasparich et al.

1995; McPheron et al. 1994). Only a few inter (Barr et al. 2005; Smith-Caldas et al. 2001; McPheron et al. 1999) and intra-specific (Alberti et al. 2008; Boykin et al. 2006; Steck 1991) studies have been performed on members of the *Anastrepha* genus. A few diagnostic markers and methods (Fritz and Schable 2004; Heath et al 2002) have been developed for members of *Anastrepha* and are particularly concerned with revealing narrow regional variation and focus less on the broader geographic range of any particular member of this genus. Other methods focus on the interest in resolving relationships between closely related taxa such as sibling species within the fraterculus species complex (Steck 1991; Steck and Sheppard 1993; Selivon et al., 2005). There are no comprehensive studies that look at the phylogeographic structure and breadth of *A. ludens* and *A. obliqua* that utilize entire sequences. The work presented here provides an extensive search for molecular markers from nuclear and mitochondrial DNA regions. The advantages of using multiple regions for examining relationships among geographic collections are many and may provide insight on the relationships for these species. This approach may also provide abundant information regarding evolutionary trends at subspecies-level. It may well provide a sound basis for further developing diagnostic markers for those individual species critical to plant pest issues of importance to the agriculture industry.

## JUSTIFYING THE NEED FOR DIAGNOSTIC MARKERS

Many of the specimens used in conducting this study were identified previously by classically-trained systematists and taxonomists. These included Drs. Allen Norrbom, USDA ARS, Systematic Entomological Laboratory, Washington, D.C. and Don Thomas, USDA ARS, Weslaco, TX. who provided identity confirmation of many of those specimens analyzed here. A further in-

depth evaluation of individuals from the species *A. obliqua* will be graciously performed by Dr. Roberto Zucchi, Universidade de São Paulo, Piracicaba, SP, Brazil.

While easily distinguished from other Tephritidae (Norrbom et al., 1999), *Anastrepha* have few morphological characters that permit their accurate identification to species-level. These include body and wing coloration as well as wing patterns that are used to identify adults from this genus. Also, both male and female genitalia such as the female aculeus may be used (Foote et al., 1993) in species identification. In larvae, oral ridges and anterior spiracles are used in determining the species' identity. It is more difficult however, to accurately identify a first or second instar larva. A limited number of species may be identified with characters originating from either the anterior and posterior regions of the egg (Murillo & Jiron 1994). While useful to classically-trained taxonomists, the characters mentioned here present a challenge for the untrained eye. At best these characters may serve as species indicators but not beyond this taxonomic level, certainly not to population-level. At times these characters may be considered unreliable given the variability as a result of specimen condition and variation of characters as a function of age at a particular life stage. The variation may even be due to subjective characterization by humans and inherent differences between individuals occurring within the vast range of geographic distribution for some of the species. Also, these characters provide little utility for distinguishing between individuals belonging to sibling species or those grouped into species complexes. Therefore, I will develop molecular methods that not only provide a view into evolutionary trends and processes but also facilitate the identification of two members from this genus, *A. obliqua* and *A. ludens*.

## REFERENCES


Aketarawong N, Bonizzoni M, Thanaphum S, Gomuiski LM, Gasperi G, Malacrida AR, and Gugliemino CR (2007) Inferences on the population structure and colonization process of the invasive oriental fruit fly, Bactrocera dorsalis (Handel). Mol. Ecol. 16:3522-3532.

Alberti AC, Confalonieri VA, Zandomeni RO, Vilardi JC (2008) Phylogeographic studies on natural populations of the South American fruit fly, *Anastrepha* fraterculus (Diptera : Tephritidae). Genetica 132, 1-8.

Aluja M, Celedonio-Hurtado H, Liedo P, Cabrera M, Castillo F, Guill J, Rios E (1996) Seasonal population fluctuations and ecological implications for management of *Anastrepha* fruit flies (Diptera: Tephritidae) in commercial mango orchards in southern Mexico. *Journal of Economic Entomology* **89**, 654-667.

Aluja M (1994) Bionomics and management of *Anastrepha*. *Annual Review of Entomology* **39**, 155-178.

Aluja M & Birke B (1993) Habitat use by adults of *Anastrepha obliqua* (Diptera: Tephritidae) in a mixed mango and tropical plum orchard. Ann. Entomol. Soc. Am. 86: 799-812.

Armstrong KF, Cameron CM, Frampton ER (1997) Fruit fly (Diptera: Tephritidae) species identification: a rapid molecular diagnostic technique for quarantine application. *Bull. Entomol. Res.* **87**: 111-8.

Bale JS, Masters GJ, Hodkinson ID, Awmack C, Bezemer TM, Brown VK, Butterfield J, Buse A, Coulson JC, Farrar J, Good JE, Harrington R, Hartley S, Jones TH, Lindroth RL, Press MC, Symrnioudis I, Watt AD, Whittacker JB (2002) Herbivory in global climate change research: direct effects of rising temperature on insect herbivores. Global Change Biol. 8, 1–16.

Barr NB, Copeland RS, De Meyer M, Masiga D, Kibogo HG, Billah MK, Osir E, Wharton RA, and McPheron BA (2006) Molecular diagnostics of economically important *Ceratitis* fruit fly species (Diptera : Tephritidae) in Africa using PCR and RFLP analyses. *Bulletin of Entomological Research* **96**, 505-521.

Barr NB and McPheron BA (2006) Molecular phylogenetics of the genus *Ceratitis* (Diptera: Tephritidae). Molecular Phylogenetics and Evolution. **38**, 216–230.

Barr NB, Cui LW, McPheron BA (2005) Molecular systematics of nuclear gene period in genus *Anastrepha* (Tephritidae). *Annals of the Entomological Society of America* **98**, 173-180.

Boykin LM, Shatters RG, Hall DG, Burns RE, Franqui RA (2006) Analysis of host preference and geographical distribution of *Anastrepha suspensa* (Diptera : Tephritidae) using phylogenetic analyses of mitochondrial cytochrome oxidase I DNA sequence data. *Bulletin of Entomological Research* **96**, 457-469.

Bush G (1969) Sympatric host formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera, Tephritidae). Evolution **23**: 237-251.

Bustos RME, Toledo J, and Enkerlin W (1993) Evaluation of irradiation parameters in the quarantine treatment of Mexican mangoes, p. 329–335. *In* International Atomic Energy Agency [eds.], Cost - Benefit Aspects Food Irradiation Processing. Press Vienna.

Feder JL, Hunt TA, and Bush GL (1993) The effects of climate, host plant phenology and host fidelity on the genetics of apple and hawthorn infesting races of *Rhagoletis pomonella.* Entomol. Exp. Appl. **69:** 117–135.

Foote RH, Blanc FL, and Norrbom AL (1993) ‘‘Handbook of the Fruit Flies (Diptera: Tephritidae) of America North of Mexico,’’ Comstock, Ithaca, NY.

Fritz AH and Schable N (2004) Microsatellite loci from the Caribbean fruit fly, *Anastrepha* suspensa (Diptera: Tephritidae). Molecular Ecology Notes 4, 443–445.

Gasparich G E, Sheppard WS, Han H-Y, McPheron BA, and Steck GJ (1995) Analysis of mitochondrial DNA and development of PCR-based diagnostic molecular markers for Mediterranean fruit fly (*Ceratitis capitata*) populations. Insect Molec. Biol. **4**:61-67.

Han, H-Y and McPheron, BA (1999) Molecular data as a tool to test phylogenetic relationships among higher groups of Tephritidae: A case study using mitochondrial ribosomal DNA sequences. Pp. 115-132 in Fruit Flies (Tephritidae): Phylogeny and Evolution of Behavior, M. Aluja and A. L. Norrbom, eds., CRC Press.

Heath M, Kuhn D, Schnell R, and Olano C (2002) Mitochondrial DNA restriction map for the Caribbean fruit fly, *Anastrepha* suspensa, and occurrence of mitochondrial DNA diversity within highly inbred colonies. Biochemical Genetics **40**, 283–292.

Henning F & Matioli SR (2006) Mating time of the West Indian fruit fly *Anastrepha obliqua* (Macquart) (Diptera: Tephritidae) under laboratory conditions. Neotropical Entomology **35(1)**, 145-148.

Hernandez-Ortiz V (1992) El género *Anastrepha* Schiner en México (Diptera: Tephritidae). Taxonomía, distribución y sus plantas huéspedes. Instituto de Ecología Publ. 33. Xalapa, Veracruz, Mexico. 162 pp.

Klassen W and Curtis CF (2005) History of the sterile insect technique, pp. 3-36 *In* V. A. Dyck, J. Hendrichs, and A. S. Robinson. The Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management. Springer, Dordrecht, The Netherlands. 787 pp.

Lanzavecchia SB, Cladera JL, Faccio P, Petit MN, Vilardi JC, Zandomeni RO (2008) Origin and distribution of *Ceratitis capitata* mitochondrial DNA haplotypes in Argentina. *Annals of the Entomological Society of America* **101**, 627-638.

McPheron, B. A., H-Y. Han, J. G. Silva, and A. L. Norrbom. 1999. Phylogeny of the genera *Anastrepha* and *Toxotrypana* (Trypetinae: Toxotrypanini) based upon 16S rRNA mitochondrial DNA sequences, pp. 343-361. In M. Aluja and A.L. Norrbom (eds.) Fruit flies (Tephritidae): phylogeny and evolution of behavior. Boca Raton, Florida, CRC Press, USA.

McPheron BA, Gasparich GE, Han HY, Steck GJ, Sheppard WS (1994) Mitochondrial-DNA restriction map for the Mediterranean fruit fly, *Ceratitis capitata*. *Biochemical Genetics* **32**, 25-33.

Müller-Schärer H, Schaffner U, and Steinger T. 2004. Evolution in invasive plants: implications for biological control. *Trends Ecol Evol* **19**: 417–22.

Murillo T and Jiron LF (1994) Egg morphology of *Anastrepha obliqua* and some comparative aspects with eggs of *Anastrepha fraterculus* (Diptera, Tephritidae). Florida Entomol. 77: 342-348.

Nardi F, Carapelli A, Dallai R, Roderick GK, and Frati F (2005) Population structure and colonization history of the olive fly, *Bactrocera oleae* (Diptera, Tephritidae). Molecular Ecology 14, 2729–2738.

Norrbom AL and Caraballo J (2003) A new species of *Anastrepha* from Amazonia with redescriptions of *A. caudata* Stone and *A. mendeliana* Lima (Diptera: Tephritidae). Insecta Mundi 17: 33-43.

Norrbom AL, Zucchi RA & Hernández-Ortiz V (1999) Phylogeny of the genera *Anastrepha* and *Toxotrypana* (Trypetinae: Toxotrypanini) based on morphology, p. 299-342. In M. Aluja & A.L. Norrbom (eds.), Fruit flies (Tephritidae): phylogeny and evolution of behavior. Boca Raton, Florida, CRC Press, 944p.

Norrbom AL and Foote RH (1989) The taxonomy and zoogeography of the genus *Anastrepha* (Diptera: Tephritidae), pp. 15-26 in A.S. Robinson and G. Hooper, [eds.], World Crop Pests. Fruit Flies, their biology, natural enemies and control. Amsterdam, Elsevier Science. v. 3A.

Novotny V, Clarke AR, Drew RAI, Balagawi S, Clifford B (2005) Host specialization and species richness of fruit flies (Diptera : Tephritidae) in a New Guinea rain forest. *Journal of Tropical Ecology* **21**, 67-77.

Selivon D, Perondini ALP, and Morgante JS (2005) A genetic-morphological characterization of two cryptic species of the *Anastrepha fraterculus* complex (Diptera: Tephritidae). *Ann. Entomol. Soc. Am*., vol. 98, no. 3, p. 367-381.

Shochat EP, Warren S, Faeth SH, McIntyre NE, and Hope D (2006) From patterns to emerging processes in mechanistic urban ecology. Trends in Ecology & Evolution **21:**186–191.

Smith DC (1988) Heritable divergence of *Rhagoletis pomonella* host races by seasonal asynchrony. Nature **336:** 66–67.

Smith-Caldas MRB, McPheron BA, Silva JG, and Zucchi RA (2001) Phylogenetic relationships among species of the *fraterculus* group (*Anastrepha*: Diptera: Tephritidae) inferred from DNA sequences of mitochondrial cytochrome oxidase I. Neotrop. Entomol. **30**: 565-573.

Steck GJ and Sheppard WS (1993) Mitochondrial DNA variation in *Anastrepha fraterculus*. pp. 9–14 in Aluja, M. & Liedo, P. (Eds) Fruit flies: biology and management. New York, Springer-Verlag.

Steck G J (1991) Biochemical systematics and population genetic-structure of *Anastrepha fraterculus* and related species (Diptera: Tephritidae). Ann. Entomol. Soc. Am. 84: 10-28.

Stone A (1942) The fruit flies of the genus *Anastrepha*. U.S. Dept. Agric. Misc. Publ. 439: 1-112.

Via S (1999) Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. Evolution **53**, 1446–1457.

White IM, and Elson-Harris M (1992) Fruit flies of economic significance: their identification and bionomics. International Institute of Entomology, London.

## Chapter 2: Phylogeographic analysis of Mexican fruit fly (Diptera: Tephritidae) using mitochondrial DNA sequence data

Raul Ruiz, Christopher L. Owen, Donald B. Thomas, and Bruce A. McPheron

## ABSTRACT

*Anastrepha ludens* (Loew) (Diptera: Tephritidae), Mexican fruit fly, is a frugiferous pest that occasionally infests fruit outside its native range. It is associated with >50 plants and is a major pest of grapefruit and orange. It has wide distribution in Mexico and Central America, with infestations occurring in Texas, California, and Florida. This research evaluates the utility of a sequence-based approach for two mitochondrial (COI and ND6) gene regions. We assess these markers as genetic tools for their use in bringing to light historical processes, delineating populations, and for further development for diagnostics. We analyzed 543 individuals from 67 geographic collections and found a predominant haplotype occurring in the majority of specimens. We observed 68 haplotypes in all and see a substantial difference among haplotypes belonging to northern and southern groups of collections. Mexico haplotypes differ by few bases possibly as a result of a recent bottleneck event. In contrast, we see that specimens from two southern collections show high genetic variability delineating three mitochondrial groups. We show that COI and ND6 are useful for phylogeographic studies of *A. ludens.*

The Mexican fruit fly, *Anastrepha ludens* (Loew) (Diptera: Tephritidae), a major pest of citrus and mangos (White and Elson-Harris 1992), is distributed from the southern tier states of the United States through Central America (Enkerlin et al. 1989; Hernandez-Ortiz and Aluja 1993). This species is of great interest for quarantine purposes, as it is frequently recovered in citrus production areas in the United States, including the Lower Rio Grande Valley of Texas, Arizona, and California. Less commonly, Mexican fruit fly has been intercepted in Florida (Steck 1998). Due to the importance of this species as a pest of commercial fruit, its frequent human-mediated dispersal, and its relatively recent confirmation at the southern end of the species' distribution (Jiron et al. 1988; Norrbom et al. 2005), an understanding of Mexican fruit fly population structure could contribute to improved management strategies.

The commercial hosts of Mexican fruit fly, citrus and mango, were introduced to the New World. Baker et al. (1944) hypothesized that the pre-Columbian distribution of the species was restricted to the Sierra Madre Oriental in northeastern Mexico, where Mexican fruit fly is still abundant on a native host plant, yellow chapote (*Casimiroa greggii* (S. Watson)). These authors also assembled anecdotal evidence that Mexican fruit fly infestation of citrus and mango was a relatively recent occurrence, perhaps in the 20$^{th}$ century. Recent pest status in isthmian Central America is well documented. While Mexican fruit fly had been reported as far south as Colombia (Murillo 1931; Steyskal 1977), Norrbom et al. (2005) demonstrated that those records were actually attributable to a previously undescribed species closely related to Mexican fruit fly. Stone (1942) cited records of Mexican fruit fly in Costa Rica, but subsequent pest surveys (Jiron and Zeledon 1979; Jiron and Hedstrom 1988) did not recover this species. A comprehensive review of museum specimens combined with trapping and rearing of infested host material (Jiron et al. 1988) identified six Mexican fruit fly specimens out of 15,000 specimens examined; three of these were specimens

referenced by Stone (1942) and the other three were reared from white sapote, *C. edulis* Llave & Lex. The southernmost verified Mexican fruit fly record is in western Panama (Norrbom et al. 2005), and one of the authors, Korytkowski (pers. comm.), confirms that the species is not considered a pest of commercial fruit in Panama. All evidence suggests that Mexican fruit fly has been, at best, uncommon in the southern portion of its current range.

We employ a mitochondrial DNA (mtDNA) sequence-based molecular approach that examines specimens collected from sites throughout the geographic range of the Mexican fruit fly. Mitochondrial DNA has been a marker of choice in many phylogeographic studies (Avise 2004). Mitochondrial DNA genes experience relatively high substitution rates in the absence of genetic recombination, ideal for examining relationships and inheritance. Previous studies utilizing molecular methods have addressed interspecific (Barr et al. 2005, Smith-Caldas et al. 2001) and intraspecific (Alberti et al. 2008, Boykin et al. 2006, Steck 1991) relationships among and within members of the genus *Anastrepha*. This is the first phylogeographic study that examines the genetic diversity of Mexican fruit fly across its full geographic range. Our study tests the hypothesis of an origin of the species in northeastern Mexico (Baker et al. 1944) and subsequent spread throughout the region. Data from individual haplotypes may also provide diagnostic characters useful in tracking introductions from one geographic region to another, including introductions with quarantine significance.

## MATERIALS AND METHODS

**Sample Collection.** *Anastrepha ludens* specimens used in this study were collected in traps (adults) or from host fruit material (larvae) between 1998 and 2006. The 67 sampling sites and host association of samples are described in Table 1, and the sites are depicted graphically in Fig. 1. Individuals were collected from citrus and mango in Costa Rica, where the only previously reported infested host was white sapote, C. edulis (Jiron and Hedstrom 1988). We analyzed samples from three sites in Panama, the southernmost verified edge of Mexican fruit fly distribution (Norrbom et al. 2005). Specimens were identified using morphological characters by trained taxonomists, were either fresh frozen or placed in alcohol, and shipped to the USDA Center for Plant Health Science and Technology, Pest Detection, Diagnostic, and Management Lab in Edinburg, TX, or to Penn State University, Department of Entomology, University Park, PA. Specimens were maintained at –80$^{o}$C at both locations prior to analysis.

Table 1. Geographical origins and information on collections of *Anastrepha ludens* included in this study.

| Site# | Country | State | Locality | Loc code | *n* | Latitude | Longitude | Elev (ft) | Year coll. | Host |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Population | | | | | | | | |
| 1 | Belize | Cayo Dist | Benque Viejo | BEL1 | 3 | 18.266667 N | -88.45 W | 20 | 1999 | Grapefruit |
| 2 | Belize | Corozal Dist | Corozal | BEL2 | 10 | 18.22255 N | -88.32295 W | 59 | 1999 | Grapefruit |
| 3 | Belize | Stann Creek Dist | Dangriga | BEL3 | 8 | 17.183333 N | -88.58333 W | 554 | 1999 | Grapefruit |
| 4 | Belize | Cayo Dist | Georgeville | BEL4 | 7 | 17.183 N | -88.966 W | 505 | 1999 | Grapefruit |
| 5 | Belize | | Pomona | BEL5 | 6 | 16.983 N | -88.366 W | 135 | 1999 | Grapefruit |
| 6 | Costa Rica | Alajuela Prov | Naranjo | CR1 | 5 | 10.09 N | -84.4 W | 3796 | 2000 | Mango |
| 7 | Costa Rica | Puntarenas Prov | Puntarenas | CR2 | 8 | 9.2 N | -84.006 W | 0 | 2005 | Sour orange |
| 8 | Costa Rica | San Jose Prov | Puriscal | CR3 | 10 | 9.8 N | -84.3 W | 3304 | 2000 | Mix |
| 9 | Costa Rica | San Jose Prov | Acosta | CR4 | 7 | 9.805 N | -84.14 W | 4429 | 2000 | Orange |
| 10 | Costa Rica | San Jose Prov | Rosario | CR5 | 9 | 9.68 N | -84.06 W | 3310 | 2000 | Orange |
| 11 | Costa Rica | San Jose Prov | San Isidro | CR6 | 7 | 9.677 N | -84.074 W | 5410 | 2000 | Grapefruit |
| 12 | Guatemala | | Palin Escuintla | GT1 | 8 | 14.298 N | -90.638 W | 2848 | 1998/2001 | Grapefruit |
| 13 | Guatemala | | Sacatepequez | GT2 | 10 | 14.576465 N | -90.72872 W | 5138 | 1999 | Sapote |
| 14 | Honduras | Atlantida | Jutiapa | HON1 | 4 | 15.770928 N | -86.51461 W | 36 | 2001 | Grapefruit |
| 15 | Honduras | Atlantida | Tela | HON2 | 9 | 15.779946 N | -87.45791 W | 30 | 2001 | Orange |
| 16 | Honduras | Cortes | Lago Yojoa | HON3 | 5 | 14.943928 N | -88.015 W | 2169 | 2001 | Orange |
| 17 | Honduras | Yoro | Guaymitas | HON4 | 6 | 15.502526 N | -87.71162 W | 154 | 2001 | Orange |
| 18 | Honduras | | F. Morazan | HON5 | 8 | 14.7309 N | -87.27327 W | 2044 | 2004 | Sour orange |
| 19 | Mexico | Chiapas | Cacahoatan | MX1 | 20 | 14.959421 N | -92.16796 W | 1240 | 1999 | Sour orange |
| 20 | Mexico | Chiapas | Huehuetan | MX2 | 7 | 15.03195 N | -92.38451 W | 407 | 2003 | Sour orange |
| 21 | Mexico | Chiapas | Tapachula | MX3 | 18 | 14.90675 N | -92.26068 W | 594 | | Sour orange |
| 22 | Mexico | Chiapas | Tapanatepec | MX4 | 11 | 16.3722 N | -94.217 W | 266 | 2000 | |
| 23 | Mexico | Chiapas | Union Juarez | MX5 | 18 | 15.065 N | -92.081 W | 4469 | | Sour orange |
| 24 | Mexico | Colima | Colima | MX6 | 9 | 19.245226 N | -103.7335 W | 1598 | 2006 | Sour orange |
| 25 | Mexico | Colima | Tecoman | MX7 | 9 | 19.237 N | -103.704 W | 1680 | 2006 | Mango |
| 26 | Mexico | Guerrero | Acahuizotla | MX8 | 2 | 17.405868 N | -99.44621 W | 3655 | 2005 | Mango |
| 27 | Mexico | Hidalgo | Chalcocotipa | MX9 | 10 | 20.497 N | -98.924 W | 6745 | 2004/2006 | Orange |
| 28 | Mexico | Hidalgo | | MX10 | 7 | 20.497 N | -98.924 W | 6745 | 2004/2006 | Orange |
| 29 | Mexico | Jalisco | Cd Guzman | MX11 | 10 | 19.701 N | -103.466 W | 5020 | 2006 | Sour orange |
| 30 | Mexico | Jalisco | Cofrida | MX12 | 10 | 20.038333 N | -103.9668 W | 3763 | 2006 | Sour orange |
| 31 | Mexico | Jalisco | Mezquitlan | MX13 | 9 | 19.817 N | -104.34 W | 3176 | 2006 | Orange |
| 32 | Mexico | Jalisco | Saucedes | MX14 | 8 | 20.42046 N | -105.5679 W | 1693 | 2000 | Sour orange |
| 33 | Mexico | Jalisco | La Sauceda | MX15 | 10 | 20.42046 N | -105.5679 W | 1693 | 2006 | Mango |
| 34 | Mexico | Jalisco | Tamazulita | MX16 | 7 | 19.684428 N | -103.2495 W | 3757 | 1999 | |
| 35 | Mexico | Mexico | Zacazonapan | MX17 | 10 | 19.574 N | -99.063 W | 7717 | 2006 | Sour orange |
| 36 | Mexico | Michoacan | Apatzingan | MX18 | 4 | 19.0786 N | -102.351 W | 1053 | 2000 | Grapefruit |
| 37 | Mexico | Michoacan | El Salitre | MX19 | 4 | 20.654 N | -100.378 W | 6667 | 2005 | Grapefruit |
| 38 | Mexico | Michoacan | Nueva Italia | MX20 | 9 | 19.73 N | -101.206 W | 6260 | 2006 | Mango |
| 39 | Mexico | Michoacan | Patzcuaro | MX21 | 10 | 19.517 N | -101.61 W | 7057 | 2006 | Grapefruit |
| 40 | Mexico | Michoacan | San J de Lima | MX22 | 9 | 18.566 N | -103.633 W | 361 | 2005 | Orange |
| 41 | Mexico | Morelos | Mazatepec | MX23 | 11 | 18.7255 N | -99.363 W | 3222 | 1999 | Sour orange |
| 42 | Mexico | Nayarit | San Blas | MX24 | 9 | 21.541 N | -105.284 W | 26 | 2003 | Sapote |

Table 1. (cont.)

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 43 | Mexico | Nayrit | Jala | MX25 | 10 | 21.102 N | -104.5157 W | 4974 | 2005 | Sour orange |
| 44 | Mexico | Nayrit | Jalisco | MX26 | 8 | 20.66 N | -105.224 W | 23 | 2006 | Sour orange |
| 45 | Mexico | Nuevo Leon | Allende | MX27 | 7 | 25.28466 N | -100.0404 W | 1578 | 2003 | Sour orange |
| 46 | Mexico | Nuevo Leon | Cola de Caballo | MX28 | 8 | 25.31364 N | -100.1336 W | 3858 | 1999 | Sapote |
| 47 | Mexico | Nuevo Leon | Linares | MX29 | 9 | 24.85512 N | -99.56573 W | 1158 | 2000 | Grapefruit |
| 48 | Mexico | Nuevo Leon | St Rosa Canyon | MX30 | 18 | 26.03975 N | -100.0409 W | 1929 | 1999 | Mix |
| 49 | Mexico | Puebla | La Ceiba | MX31 | 2 | 19.1 N | -98.22 W | 7093 | 1999 | Grapefruit |
| 50 | Mexico | Queretaro | Ahuacatlan | MX32 | 7 | 20.384 N | -99.9576 W | 6535 | 2005 | Sour orange |
| 51 | Mexico | Queretaro | Jalpan | MX33 | 3 | 20.578 N | -100.3874 W | 6010 | 2005 | Sour orange |
| 52 | Mexico | Quintana Roo | Andres | MX34 | 2 | 19.290705 N | -88.22398 W | 85 | 2001 | Orange |
| 53 | Mexico | Quintana Roo | Bacalar | MX35 | 3 | 18.676757 N | -88.38782 W | 0 | 2001 | Grapefruit |
| 54 | Mexico | Quintana Roo | Senor | MX36 | 3 | 19.811 N | -88.113 W | 72 | 2001 | |
| 55 | Mexico | San Luis Potosi | Huachihuayan | MX37 | 9 | 21.4999 N | -98.95 W | 1631 | 2000 | |
| 56 | Mexico | San Luis Potosi | Picholco | MX38 | 1 | 21.379 N | -98.867 W | 374 | 2004 | Orange |
| 57 | Mexico | Tabasco | Chontalpa | MX39 | 3 | 17.665 N | -93.48 W | 171 | 2006 | Orange |
| 58 | Mexico | Tabasco | Macuspana | MX40 | 6 | 17.698071 N | -92.62833 W | 49 | 2001 | Grapefruit |
| 59 | Mexico | Tamaulipas | Canon Caballero | MX41 | 13 | 24.613 N | -98.559 W | 968 | 2000 | Sapote |
| 60 | Mexico | Tamaulipas | Ocampo | MX42 | 11 | 22.87722 N | -99.25183 W | 676 | 2000 | Orange |
| 61 | Mexico | Tamaulipas | Santa Engracia | MX43 | 8 | 24.029 N | -99.276 W | 807 | 1999 | Grapefruit |
| 62 | Mexico | Veracruz | Alameda | MX44 | 10 | 19.803 N | -96.143 W | 0 | 2005 | |
| 63 | Mexico | Veracruz | Nautla | MX45 | 15 | 20.202049 N | -96.77031 W | 7 | 2001 | Orange |
| 64 | Mexico | Veracruz | Teocelo | MX46 | 10 | 19.387409 N | -96.95669 W | 3766 | 2005 | Sapote |
| 65 | Panama | | Altos de Pacora | PAN1 | 3 | 9.081758 N | -79.29082 W | 52 | 2006 | Trap |
| 66 | Panama | | Boquete | PAN2 | 3 | 8.774567 N | -82.432 W | 3543 | 2006 | Trap |
| 67 | Panama | | Chiriqui | PAN3 | 10 | 8.19187 N | -82.26255 W | 0 | 2006 | Grapefruit |

*n*, number of individual specimens analyzed; feet above sea level. Blank fields indicate missing information.

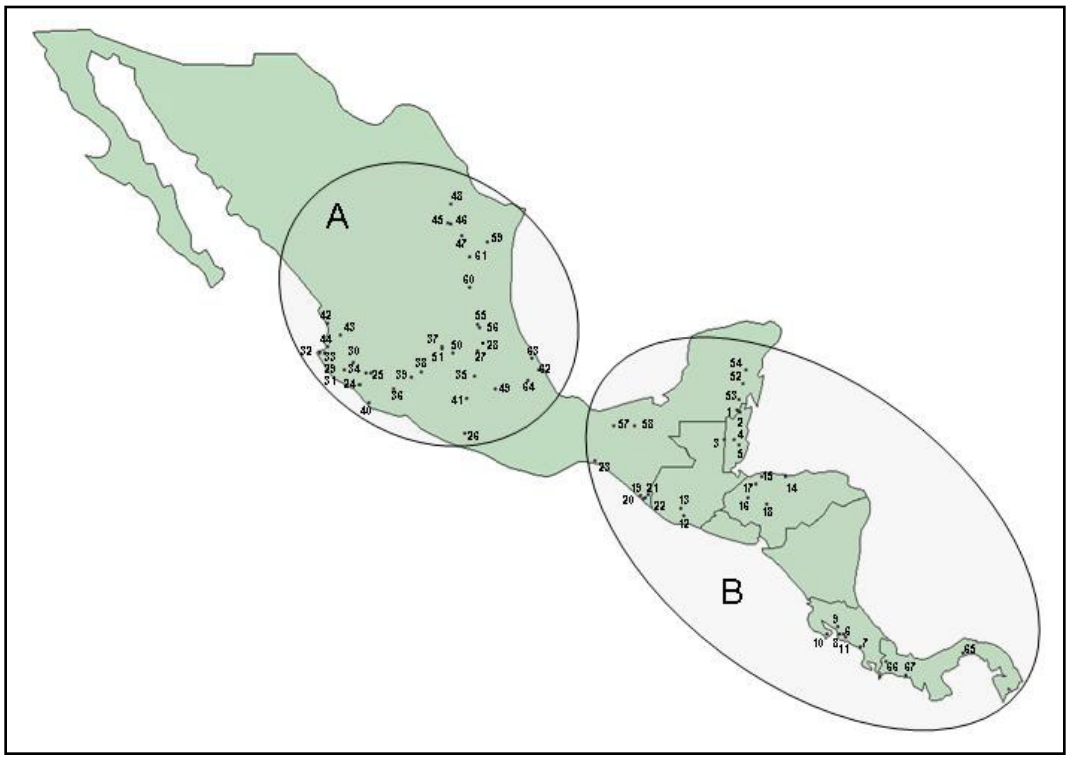


Figure 1. Geographic distribution of the 29 most common haplotypes and respective frequencies among north (A) and south (B) groups of *A. ludens* collections. Pie charts show haplotype observed at site, site number and name corresponding to locality in Table 1.

**DNA Extraction, PCR Amplification, and Sequencing**. DNA was isolated using the DNeasy® Blood and Tissue Kit (Qiagen) following standard DNeasy guidelines for animal tissues. Extractions were stored at –20°C for the duration of the study. Whenever possible, DNA from one specimen per collection was isolated via a non-destructive method (Barr et al., 2006) to serve as voucher. Samples were amplified for a region mitochondrial cytochrome oxidase subunit I (COI) using tRNA-cys2 5'-ACTCCTTTAGAATTGCAGTCTAAT-3' and COId-r 5'-GGGCTCATACAATAAATCCTAAT-3' designed from the Bactrocera abdominalis (GenBank accession number DQ917577) sequence, generating a fragment of approximately 1100 base pairs. A portion of the mitochondrial NADH dehydrogenase subunit 6 (ND6) gene was amplified using TT-J-9886 5'-TAAAAACATTGGTCTTGTAA-3' (Barr et al. 2006) and ND6r 5'-

TTATGATCCAAAATTTCATCA-3' primers and yielded a fragment of approximately 750 base pairs. PCR reactions were performed in 25 μL reactions containing 1 μL of template, 16.875 μL of water, 2.5 μL of buffer (10X conc., Qiagen), 1.5 μL of 25 mM of $MgCl_2$ (Qiagen), 1.25 μL of each primer (5 μM, Operon Technologies), 0.5 μL of 10 mM dNTP mix (Promega), and 0.125 μL of Taq DNA polymerase (Qiagen) to complete the final volume of 25 μL. Amplifications were performed in Applied Biosystems GeneAmp® PCR System 9700. Cycling conditions for amplification of COI and ND6 fragments were 3 min at 94°C followed by 39 cycles of 1 min at 94°C, 1 min at 50°C, 1 min at 72°C, and a final extension of 10 min at 72°C. PCR products were stained with Sybr® Green (Invitrogen) fluorescent dye at 1/10,000X and loaded on 1.5% electrophoresis agarose gels. Documentation of these gels was via a Gel Doc™ (Bio-Rad) imaging system using Quantity One® software. Amplification products were purified with ExoSAP-IT (USB Corp.) prior to sequencing. PCR products were sequenced asymmetrically using 3' BigDye-labeled dideoxynucleotide triphosphates (v 3.1 dye terminators, Applied Biosystems, Foster City, CA) and run on an ABI 3730XL DNA Analyzer with the ABI Data Collection Program (v 2.0) at the Huck Institute's Nucleic Acid Facility at Penn State University.

**Data Analysis**. Sequences were edited with QScreen (PSU: Huck Institute) designed in collaboration with the McPheron Lab (PSU: Dept of Entomology), and Sequencher 4.8 (Gene Codes Corp.). All sequences were aligned with Mega 4 (Tamura et al. 2007) using ClustalW and trimmed to a length of 1090 bp and 736 bp for COI and ND6, respectively. Sequences were concatenated using DnaSP 4.10 (Rozas et al. 2003) and subsequent analyses were carried out on the 1826 bp fragment.

Haplotypes were identified from these combined sequences with DnaSP. The relative haplotype frequencies within populations were calculated using ARLEQUIN 3.11 (Excoffier et al. 2005). Basic molecular summary statistics, the demographic parameters D (Tajima 1989), and F (Fu 1997), and mismatch distributions were estimated with DnaSP. The collections were separated according to northern and southern groups based on their location to the Trans-Mexican Volcanic Belt (TMV). The northern group consisted of collections within close proximity to and north of the TMV. The southern group was of those collections located further south of the TMV. Demographic parameters were estimated for northern and southern groups and their 95% confidence intervals were assessed by 10,000 coalescent simulations to test the hypothesis that all mutations are selectively neutral. ARLEQUIN was used to calculate *F*st (Wright 1951, Weir and Cockerham 1984) measures and to estimate the source of variation among groups, among populations within groups, and within populations using analysis of molecular variance (AMOVA); a permutation method with 10,000 replications was applied. The relationship between mtDNA divergence and geographic distance (isolation by distance) for populations was tested using Mantel tests (Mantel 1967) with 1000 replications in IBD version 1.52 (Bohonak 2002).

A maximum parsimony haplotype network was constructed according to the algorithm described in Templeton et al. (1992) and implemented in the program TCS 1.18 (Clement et al. 2000). TCS was used to generate schematics representative of parsimonious relationships between and among collections according to geographic association. A nested clade phylogeographic analysis (NCPA) (Templeton et al. 1995) was performed using ANeCA (Panchal and Beaumont 2007) to differentiate between population structure and historical events (Templeton 1998). Nested clade contingency tests were performed in GeoDis v2.5, and an inference key was used to determine whether phylogeographic differences were due to allopatric associations, recurrent but restricted gene flow,

or range expansion, thus providing insight into the evolutionary history and population structure of this species.

Phylogenetic tree reconstruction was estimated using Neighbor Joining (NJ), and maximum likelihood (ML). NJ phylogenetic reconstruction was conducted in Mega with 500 bootstrapping replications. The appropriate ML model of sequence evolution was determined with hierarchical likelihood tests (hLRTs) using Modeltest 3.7 (Posada and Crandall 1998) in conjunction with Phyml (Guindon et al. 2005). The TrN+I+G substitution model with a proportion of invariable sites (I) 0.8152 and the shape parameter of the gamma distribution (Γ) 0.8989 was selected using a set of hLRTs implemented in Modeltest. ML bootstrap analyses were conducted with 1000 replicates (Felsenstein 1985). The Phyml ML tree search was initiated using a NJ tree that was constructed with Mega. We rooted these trees with a sequence from one individual *Anastrepha pseudoparallela*.

## RESULTS

**Mitochondrial Variability in *Anastrepha ludens*.** A total of 68 composite (COI+ND6) haplotypes were observed in the 543 individuals of *A. ludens* sequenced. There were 105 (5.7%) polymorphic sites observed (Table 2). Twenty-nine of the haplotypes were recovered from 2 or more individuals (Fig. 1). The most abundant haplotype (AL03) was found in 351 (65%) of individuals sequenced and was observed over a widespread geographic range (Table 3). A second haplotype (AL04) was also relatively common, occurring in 34 (6%) of specimens sequenced, all of which were found only in southern collection sites. The mean number of haplotypes per collection site was 2.7, ranging between 1 and 8 per population. The presence of an extremely abundant haplotype and many low-

frequency, closely-related haplotypes is reflected in both the haplotype (0.576) and nucleotide (0.00184) diversity of the overall sample. We see a substantial difference among populations when collections are analyzed as northern and southern groups. Twenty-six haplotypes were unique to the northern region (group A, hereafter) and 33 to the southern region (hereafter group B) (Table 3 and Fig. 1). There were 9 (13%) haplotypes shared between the geographic areas. The majority of haplotypes 48 (80%) were private from populations with all but 9 being singletons. We observed 20 haplotypes private to fourteen group A populations. Group A collections showed 35 haplotypes (private + shared) with considerably lower haplotype (0.462) and nucleotide diversity (0.00061) (Table 4) when compared to group B. Twenty eight haplotypes were private to eighteen group B populations. Group B collections showed 42 haplotypes (private + shared) and a higher haplotype (0.692) and nucleotide diversity (0.00331) (Table 4) than group A collections.

Table 2. Alignment showing the haplotypes and 105 variable sites for COI and ND6 concatenated sequences observed in *Anastrepha ludens* collections.

| Positions | COI (60) | ND6 (45) |
|---|---|---|
| | 1111 | 111111111111111111111111111111111111111111111 |
| | 11111122222223333344444444455567778889999999990000 | 11122222222333333344444444445555666666777788 |
| | 23445566780234443456899145790112225680390356023034556899127 7 | 45901346789012247803345567889123601236934680 1 |
| | 17472625530910247313316717787170258493902424656217035969980 2 | 41227812705648990090573714694435773828512180 0 |
| AL01 | GTTAACGATGTAGGAAAGCAGTCTAGTAATGGTCGAAAATAAATTGGGTAATCCGGTAAC | CTCACTCCGAATTAAGTGCCTCTTATATTCTTCTCTGGCCGCTAT |
| AL02 | .C.G..A.C..G.A.....C.....T.....C.....A..G..T..... | ..... |
| AL03 | .C.....TG....T.G...C.....A.....TT..... | .....TA.....A..... |
| AL04 | .C.....TG....T.G...C.....A.....TT..... | .....TA.....A....A..... |
| AL05 | .C.G..A.C..G.A.....C.....T.....C.....A..G..T..... | ..T..... |
| AL06 | .C.....TG....T.G...C.....A.....TT..A... | .....TA.....A..... |
| AL07 | .C.....C.....TG....T.G...CG.....A.....TT..... | .....TA.....A..... |
| AL08 | .C.....C.....TG....T.G...C.....A.....TT..... | .....TA.....A..... |
| AL09 | .C.....A.G.....T.....C...C.....TT..... | ..... |
| AL10 | .C.G..ACC..G.A.....C.....T.....C.....A..G..T..... | ..... |
| AL11 | .....GAA.....T.G...C..G.....G.TT..... | .....TA.....T..... |
| AL12 | .....TG....T.G...C.....A.....TT..... | .....TA.....A..... |
| AL13 | .C.....TG....T.G...CG.....A.....TT..... | .....TA.....A..... |
| AL14 | .C.....A.....TG....T.G...C.....A.....TT..... | .....TA.....A..... |
| AL15 | .C.....G.....T.G...C..G.....GGCTT..... | .....TA.....T.....T.T..... |
| AL16 | .....GAA.....T.G...C..G.....G.TT..... | .....TA.....A..... |
| AL17 | .C.....C.....TG....T.G...C.....A.....TT..... | .....TA.....A..... |
| AL18 | .C.....TG....T.G...C.....A.....TT..... | TA..TCTTAG..CGGAC..T..AACCGC.TC.TC..A.T..... |
| AL19 | .....G.....T.....C.....T..... | .....T..... |
| AL20 | .....GAA.....G...C..G.....G.TT....G. | .....TA..... |
| AL21 | .C.....TG....T.G...CG.....A.....TT..... | .....TA.....A.A..... |
| AL22 | .C.....G.....T.....C.GG.....G.TT...G.G | TA.TTCTTAG..CGGAC..T..AACCGC.TC.TC..A.T..... |
| AL23 | .CA.....TG....T.G...CG.....A.....TT..... | .....TA.....A.....A |
| AL24 | .....GAA.....T.G...C..G.....TT..... | .....TA.....T..... |
| AL25 | .C.....G.....T.G...C..G.....TT..... | .....TA.....T..... |
| AL26 | .C.....A.....TG....T.G...C.....A.....TT..... | .....TA.....A..... |
| AL27 | .....G.....T.....A.T.G...C..G.....G.TTTT.... | .....TA.....A..... |
| AL28 | .C.....TG....T.G...C.....A.A...TT..... | .....TA.....A..... |
| AL29 | .C.....A.....TG....T.G...C.....A.....TT..... | .....TA.....A..... |
| AL30 | .C.....TG....T.G...C.....A.....TT..... | T.....TA.....A..... |
| AL31 | .C.....TG....T.....C.....A.....TT..... | .....TA.....A..... |

Table 2. (cont.)

```
AL32  .C.....TG....T.G...C......A.....TT.....  .......TA......A.A.....
AL33  .C.....TG....T.G...C.....AA.....TT.....  .......TA......A.....
AL34  .C.....TG....TAG...C......A.....TT.....  .......TA......A.....
AL35  .C...A..TG....T.G...C......A.....TT.....  .......TA......A.....
AL36  .CA.....TG....T.G...C......A.....TT.....  .......TA......A.....
AL37  T.........GAA.....T.G...C..G.......G.TT.....  .......TA......A.....C..
AL38  ..........GAA.....T.G...C..G.......G.TT..... TA..TCTTAG..CGGAC..T..AACCGC.TC.TC..A.T.....
AL39  .C.....TG....T.G...C......A.....TT.....  .......TA......A.....A...
AL40  .C.....C.....TG....T.G...C......A.....TT.....  .......TA......A.....
AL41  .....TG....T.G...CG.....A.....TT.....  .......TA......A.....
AL42  .C.G..A.C.CG.A.G.......C.....T.....C.......A..G..T.....  .....
AL43  .C.....TG....T.GG..C......A.....TT.....  .......TA......A.....
AL44  .C.....TGC...T.G...C......A.....TT.....  .......TA......A.....A.....
AL45  .C.....GAA.....T.G...C..G.......G.TT.....  .......TA.....T.....
AL46  .C.....TG....T.G...C......A.....TT.....  .......TA.....T.....
AL47  .C.....G....T.A.....CT.....C...C.....TT.....  .....C.....
AL48  ..........G.....T.....C.....T.....  .....C.....
AL49  ..........GAAGT.....T.G...C..G.......G.TT.....  .......TA.....T.....
AL50  .C.....TG....T.G...C......A.....TT.....  .......TA......A..T.....
AL51  .C.....TG....T.G...C....C.A.....TT.....  .......TA......A.....
AL52  .C.....TG....T.G...C......A.....TT.....  .......TA.G....A.....A.....
AL53  .C.....TG.A..T.G...C......A.....TT.....  .......TA......A.....
AL54  .C.....G.....AC.....T.....C.....G.TT.....  .......TA.....T.....T.....
AL55  .C.....G..TG....T.G...C......A.....TT.....  .......TA......A.....
AL56  ..........G.....T.....C.....T..... TA..TCTTAG..CGGAC..T..AACCGC.TC.TC..A.T.....
AL57  .C.....TG....T.G...CG.....A.....TT..... T...TCTTAG..CGGAC..T....C.GC.T..TC.....
AL58  .C.G.....G.....T.....T.G.....GG.....TT.....  .......TA......A.....
AL59  .C.....TG....T.G...C......A.....TT.....  .......TA......A.....G.....
AL60  .C.....TG....T.G.G.C......A.....TT.....  .......TA......A.....
AL61  .C.....TG....T.G...C......A.....TT.....  .......TA......A.....C.....
AL62  .C.....TG....T.G...C......A.....TT.....  ....T..TAG..CGGA...T.....C.
AL63  .C..G.....TG....T.G...C......A.....TT.....  .......TA......A.....
AL64  .C.....TG....T.G...C......A.....TT.....  .......TA......A.....C.....
AL65  .C.....TG....T.G...C......A.....TT.....  .......TAG.....A.....
AL66  AC.....TG....T.G...C......A.....TT.....  .......TA......A.....
AL67  .C.....TG....T.G..GC.....A.....TT.....  .......TA......A.....
AL68  .C.....TG....T.G...C......A.....TT.....  .......TA..A...A.....
```

Table 3. The geographic distribution of 68 haplotypes observed among the *A. ludens* populations examined. Data is sorted by Group where A refers to northern and B to southern collections as shown in haplotypes distribution map (Fig. 1). Population codes (e.g. MX30) correspond to Table 1. GeneBank accession numbers to COI and ND6 sequences that composed the 68 concatenated haplotypes of *A. ludens* are shown.

| Haplotype | COI GeneBank no. | ND6 Genebank no. | Group | Origin of population (*n* = no. of individuals) |
|---|---|---|---|---|
| AL12 | HM538299 | HM538338 | A | MX30 (1), MX43 (1), MX44 (2), MX45 (1) |
| AL23 | HM538307 | HM538348 | A | MX23 (1) |
| AL29 | HM538313 | HM538338 | A | MX7 (6), MX45 (1) |
| AL30 | HM538292 | HM538349 | A | MX42 (1) |
| AL31 | HM538314 | HM538338 | A | MX30 (1) |
| AL32 | HM538292 | HM538346 | A | MX20 (1), MX30 (1) |
| AL39 | HM538292 | HM538351 | A | MX30 (1) |
| AL40 | HM538320 | HM538338 | A | MX15 (1), MX25 (1), MX28 (1) |
| AL45 | HM538325 | HM538341 | A | MX10 (2) |
| AL46 | HM538292 | HM538341 | A | MX10 (1) |
| AL50 | HM538292 | HM538355 | A | MX27 (1) |
| AL51 | HM538328 | HM538338 | A | MX6 (2), MX11 (7), MX24 (1) |
| AL52 | HM538292 | HM538356 | A | MX24 (1) |
| AL55 | HM538331 | HM538338 | A | MX21 (2) |
| AL58 | HM538332 | HM538338 | A | MX26 (1) |
| AL59 | HM538292 | HM538359 | A | MX26 (1) |
| AL06 | HM538293 | HM538338 | A | MX37 (1) |
| AL60 | HM538333 | HM538338 | A | MX26 (1) |
| AL61 | HM538292 | HM538360 | A | MX26 (1) |
| AL62 | HM538292 | HM538361 | A | MX20 (1) |
| AL63 | HM538334 | HM538338 | A | MX20 (1) |
| AL64 | HM538292 | HM538362 | A | MX46 (1) |
| AL65 | HM538292 | HM538363 | A | MX6 (1) |
| AL66 | HM538335 | HM538338 | A | MX44 (1) |
| AL67 | HM538336 | HM538338 | A | MX17 (1) |
| AL07 | HM538294 | HM538338 | A | MX15 (4), MX26 (2), MX37 (1) |
| AL11 | HM538298 | HM538341 | A, B | BEL4 (1), GT1 (1), HON5 (3), MX10 (2), MX35 (2) |
| AL13 | HM538300 | HM538338 | A, B | MX2 (1), MX3 (2), MX4 (1), MX12 (2), MX14 (1), MX17 (1), MX20 (1), MX23 (1), MX24 (1), MX43(1), MX45 (2), MX46 (1) |
| AL14 | HM538301 | HM538338 | A, B | MX3 (1), MX45 (2) |
| AL18 | HM538292 | HM538343 | A, B | BEL5 (2), HON4 (1), MX20 (1) |
| AL28 | HM538312 | HM538338 | A, B | MX1 (1), MX45 (2) |
| AL3 | HM538292 | HM538338 | A, B | BEL1 (2), BEL2 (8), BEL3 (7), BEL4 (2), BEL5 (4), CR2 (1), CR3 (6), CR4 (6), CR5 (3), GT1 (1), GT2 (3), HON1 (2), HON2 (6), HON3 (3), HON4 (2), MX1 (11), MX2 (4), MX3 (10), MX4 (6), MX5 (18), MX6 (6), MX7 (3), MX8 (2), MX9 (10), MX10 (2), MX11 (3), MX12 (8), MX13 (9), MX14 (7), MX15 (5), MX16 (7), MX17 (8), MX18 (4), MX19 (4), MX20 (4), MX21 (7), MX22 (9), MX23 (9), MX24 (6), MX25 (9), |

Table 3. (cont.)

| Haplotype | COI GeneBank no. | ND6 Genebank no. | Group | Origin of population (*n* = no. of individuals) |
|---|---|---|---|---|
| AL03 | HM538292 | HM538338 | A, B | MX26 (2), MX27 (6), MX28 (7), MX29 (9), MX30 (14), MX31 (2), MX32 (7), MX33 (3), MX34 (2), MX36 (3), MX37 (7), MX38 (1), MX39 (3), MX40 (6), MX41 (11), MX42 (10), MX44 (8), MX45 (7), MX46 (7), PAN1 (1), PAN3(8) |
| AL36 | HM538318 | HM538338 | A, B | BEL2 (1), MX21 (1) |
| AL41 | HM538321 | HM538338 | A, B | MX1 (1), MX43 (6), MX46 (1) |
| AL08 | HM538295 | HM538338 | A, B | MX2 (2), MX41 (2) |
| AL01 | HM538290 | HM538337 | B | CR5 (3), CR6 (2) |
| AL10 | HM538297 | HM538337 | B | MX35 (1) |
| AL15 | HM538302 | HM538342 | B | GT2 (1) |
| AL16 | HM538298 | HM538338 | B | BEL4 (1) |
| AL17 | HM538303 | HM538338 | B | BEL4 (2) |
| AL19 | HM538304 | HM538344 | B | HON2 (1) |
| AL02 | HM538291 | HM538337 | B | BEL1 (1), CR3 (3), CR5 (3) |
| AL20 | HM538305 | HM538345 | B | HON2 (1) |
| AL21 | HM538300 | HM538346 | B | HON2 (1), HON5 (1) |
| AL22 | HM538306 | HM538347 | B | HON3 (1) |
| AL24 | HM538308 | HM538341 | B | BEL2 (1) |
| AL25 | HM538309 | HM538341 | B | HON4 (1) |
| AL26 | HM538310 | HM538338 | B | HON4 (1) |
| AL27 | HM538311 | HM538338 | B | HON4 (1) |
| AL33 | HM538315 | HM538338 | B | MX1 (1) |
| AL34 | HM538316 | HM538338 | B | MX1 (2) |
| AL35 | HM538317 | HM538338 | B | MX4 (2) |
| AL37 | HM538319 | HM538350 | B | BEL3 (1) |
| AL38 | HM538298 | HM538343 | B | BEL4 (1) |
| AL04 | HM538292 | HM538339 | B | CR1 (1), CR2 (7), CR3 (1), CR4 (1), CR6 (1), GT1 (2), GT2 (6), HON3 (1), MX1 (1), MX3 (1), PAN1 (2), PAN2 (2), PAN3 (1) |
| AL42 | HM538322 | HM538337 | B | MX1 (1) |
| AL43 | HM538323 | HM538338 | B | MX1 (2) |
| AL44 | HM538324 | HM538352 | B | MX3 (1) |
| AL47 | HM538326 | HM538353 | B | GT1 (1) |
| AL48 | HM538304 | HM538354 | B | GT1 (2), HON5 (2) |
| AL49 | HM538327 | HM538341 | B | GT1 (1) |
| AL05 | HM538291 | HM538340 | B | CR (4) |
| AL53 | HM538329 | HM538338 | B | MX2 (2) |
| AL54 | HM538330 | HM538357 | B | PAN3 (1) |
| AL56 | HM538304 | HM538343 | B | HON5 (1) |
| AL57 | HM538300 | HM538358 | B | HON5 (1) |
| AL68 | HM538292 | HM538364 | B | PAN2 (1) |
| AL09 | HM538296 | HM538337 | B | HON (2) |

Table 4. Diversity estimates for North and South groupings of *A. ludens*

| Region | n | S | π | H | Hd | K |
|---|---|---|---|---|---|---|
| North | 309 | 61 | 0.00061 ± 0.00012 | 35 | 0.46194 | 1.10680 |
| South | 234 | 89 | 0.00331 ± 0.00038 | 42 | 0.69216 | 6.04046 |

**Population Structure, Neutrality Tests and Demographic Inferences**. We examined the source of variation by AMOVA analysis for groups A and B. Overall, an AMOVA of the sequence variation with populations nested within groups revealed that 69.53% of the variation was from within populations, 23.62% from among populations within groups, and 6.85% from between groups (Table 5).

Table 5. AMOVA analysis of *Anastrepha ludens* north and south geographic collections based on mtDNA variation. Distance method used was pairwise difference.

| Source of variation | d.f. | Percentage of Variation |
|---|---|---|
| Among groups | 1 | 6.85 |
| Among populations within groups | 65 | 23.62 |
| Within populations | 476 | 69.53 |

p = 0.0000; One population (MX38) consisting of a single individual was not considered in the analysis.

Additionally, *F*st estimates show significant structure among populations analyzed collectively as well as by groups (Appendix A). The relationship between mtDNA divergence and geographic distance (isolation by distance) for populations was tested using Mantel tests. For the entire dataset, the Mantel test showed values that were positively correlated. These values were shallow but significant ($r = 0.1767$, $P < 0.0020$) indicating isolation by distance. However, when analyzed

separately, this pattern disappears both among group A which showed a negative correlation that was not significant (r = -0.0595, P = 0.2120) and group B which showed a positive correlation that was also not significant (r = 0.0489, P = 0.2750).

Populations within geographic groups A and B deviated significantly from expectations under neutrality using both Tajima's D test and Fu's Fs (Table 6). This deviation can be explained by

Table 6. Neutrality tests and mismatch distribution analysis for North and South groupings of *A. ludens* geographic collections.

| Region | n | Fu's Fs | Tajima's D | Fu and Li's F | Fu and Li's D |
|---|---|---|---|---|---|
| North | 309 | -41.289 | -2.57942 | -5.17895 | -5.80160 |
| South | 234 | -11.374 | -1.82632 | -3.16349 | -3.36189 |

$p<0.05$, $p<0.02$. $p<0.01$

either demographic or selection factors. In order to provide additional resolution into the demographic history, mismatch distributions for these collections were examined. The combined dataset shows a multimodal curve with a single major peak at approximately 11 differences and smaller peaks at 17 and 23 mismatches (Fig. 2a). A constant-sized population is projected to show modality while an expanding population shows a smooth, unimodal distribution (Slatkin and Hudson 1991). The dataset was then partitioned to test geographic groups A and B separately. The mismatch distribution for group A populations (Fig. 2b) showed a smooth curve typifying a recently expanding population and indicated that the observed differences followed the distribution of expected differences under an expanding population model. Group B populations (Fig. 2c) showed the same three peaks observed for the pooled data set. These results suggest the expansion of group B populations at those three time units. Peaks were observed at 11 and 23 mismatches in group A paralleling that seen in group B. Both peaks were considerably reduced in size in collections from

group A, suggesting that a portion of individuals contributing to this comparison within group A are few and may have been introduced from areas belonging to group B at some point in time. The mismatch peak at 17 occurring in group B was not seen in mismatch distributions for group A. The timing of expansion for these populations according to 11, 17, and 23 mismatches may have occurred approximately 1.1, 0.8, and 0.5 million years ago (mya) under the assumption that insect mtDNA clock can be calibrated at $1.15 \times 10^{-8}$ subs/bp/year (Brower 1994). Estimating the time of

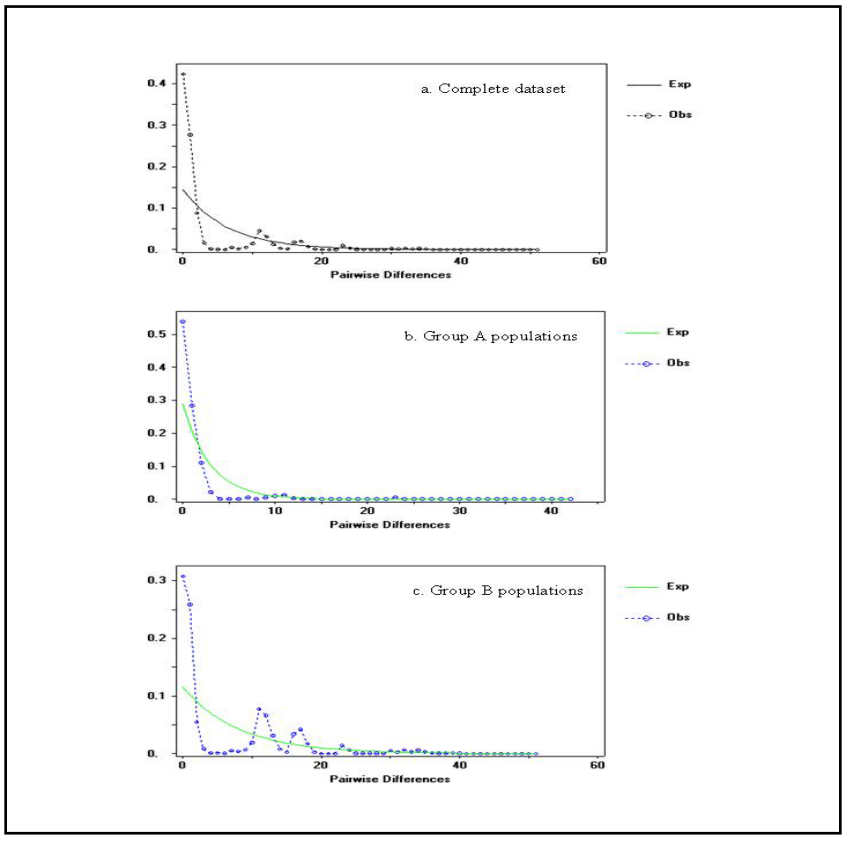


Figure 2. Mismatch distributions for mtDNA haplotypes of A. *ludens.*

introduction into northern areas from the source of origin or divergence for these two collections is based on an average difference among these two groups of 3.36 bp, which yields an approximate divergence time of 0.16 mya. This timing coincides with Pleistocene epoch (1.8 to 0.010 mya)

which has been long argued as a period in time that played a significant role in shaping the diversity within and among species (Bonaccorso et al. 2006; Carstens and Knowles 2007; Jaramillo et al. 2006).

In order to further explore the association among populations and respective haplotypes, we constructed a haplotype network. The haplotype network based upon MP (Fig. 3) showed a star-like arrangement in which most of the unique haplotypes were closely related to the common central haplotype (AL03). This is typical for species that have undergone a bottleneck event and subsequent expansion. Five loops were observed in the haplotype network, and frequency and shared locality were used to resolve these ambiguities (Pfenninger and Posada 2002). For example, the first loop occurring among AL13, AL23, and AL36 was resolved by breaking the connection between AL23 and 36. This allowed AL23 to remain connected to AL13, the most frequent haplotype found within the same geographic location. We were able to use the same criteria for resolving the remaining four loops. NCPA results showed that most alternative associations were not significantly different, suggesting a lack of geographic structure.

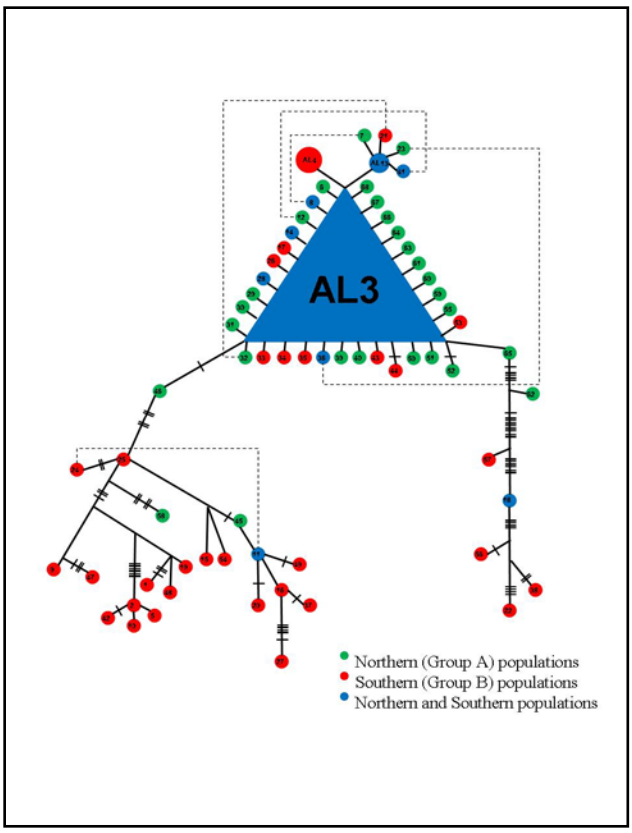


Figure 3. Statistical parsimony network for *Anastrepha ludens*. Dashed lines indicate connections broken in order to resolve network ambiguities. Numbers correspond to 68 haplotypes with mutational steps observed.

Two clades showed significant associations, permitting further phylogeographic inference. Clade 4-4 (Fig. 4) contained a portion of the haplotypes observed in cluster III (Fig. 5), consistent with a hypothesis of restricted gene flow with isolation by distance. Permutation analysis provided significant results for clade 5-2 (Fig. 4) corresponding to clusters III and IV (Fig. 5). For this clade, Templeton's (1998) inference key (updated November 11, 2005) suggested the existence of a contiguous expansion. A phylogenetic reconstruction of the 68 haplotypes identified four major clades (Fig. 5), congruent with the nested clade analysis (Fig. 4). Branch lengths were shallow using both ML and NJ methods. Low bootstrap values on many internal branches in the NJ analysis suggest caution in interpreting topologies shown in Fig. 5. Confidence in phylogenetic analysis was

assessed by Hillis & Bull (1993) who stated that bootstrap values of 70% correspond to a 95% probability that the clade is real.

Twenty two haplotypes clustered within clades I & II with haplotypes mostly originating from southern collections. Within this cluster the ten haplotypes pertaining to clade I varied by an average of 9.6 bases with all of the haplotypes originating from group B (southern) collections except for AL46. Haplotype AL46 was seen in a sample collected from Mexican state of Hidalgo (MX10). Additional haplotypes (AL11 and AL45) pertaining to clade II were recovered at this site (MX10) as well and vary by 8-9 bases when compared to AL46. Haplotype AL03 was also recovered in flies from this site and is seen in clade III.

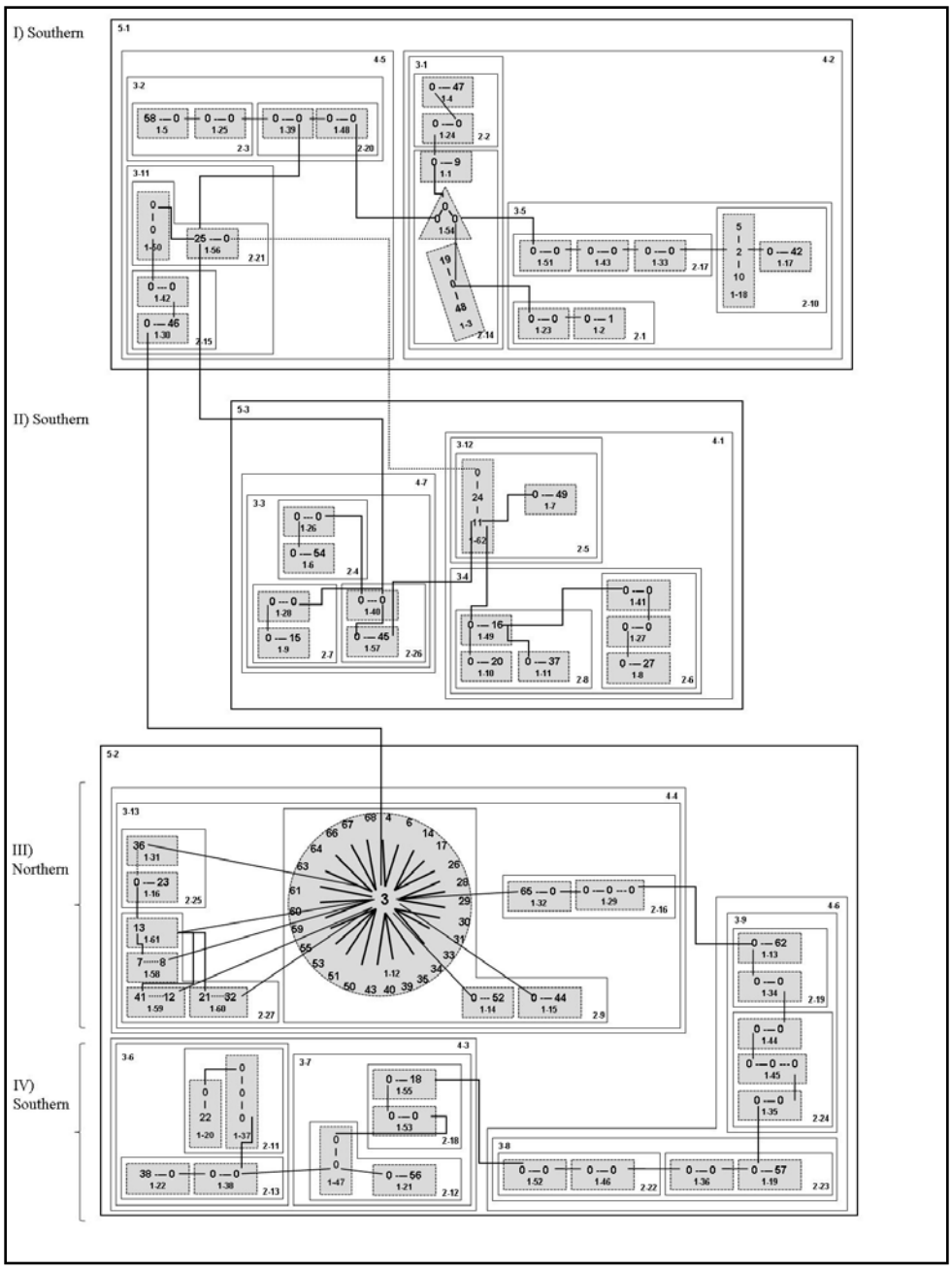


Figure 4. Estimated cladograms and nested design for COI+ND6 haplotypes observed in *Anastrepha ludens* populations sampled. Dashed lines indicate connections broken in order to resolve network ambiguities. Numbers correspond to 68 haplotypes with mutational steps shown as "0". Regional labels correspond to those in text and Fig. 5.

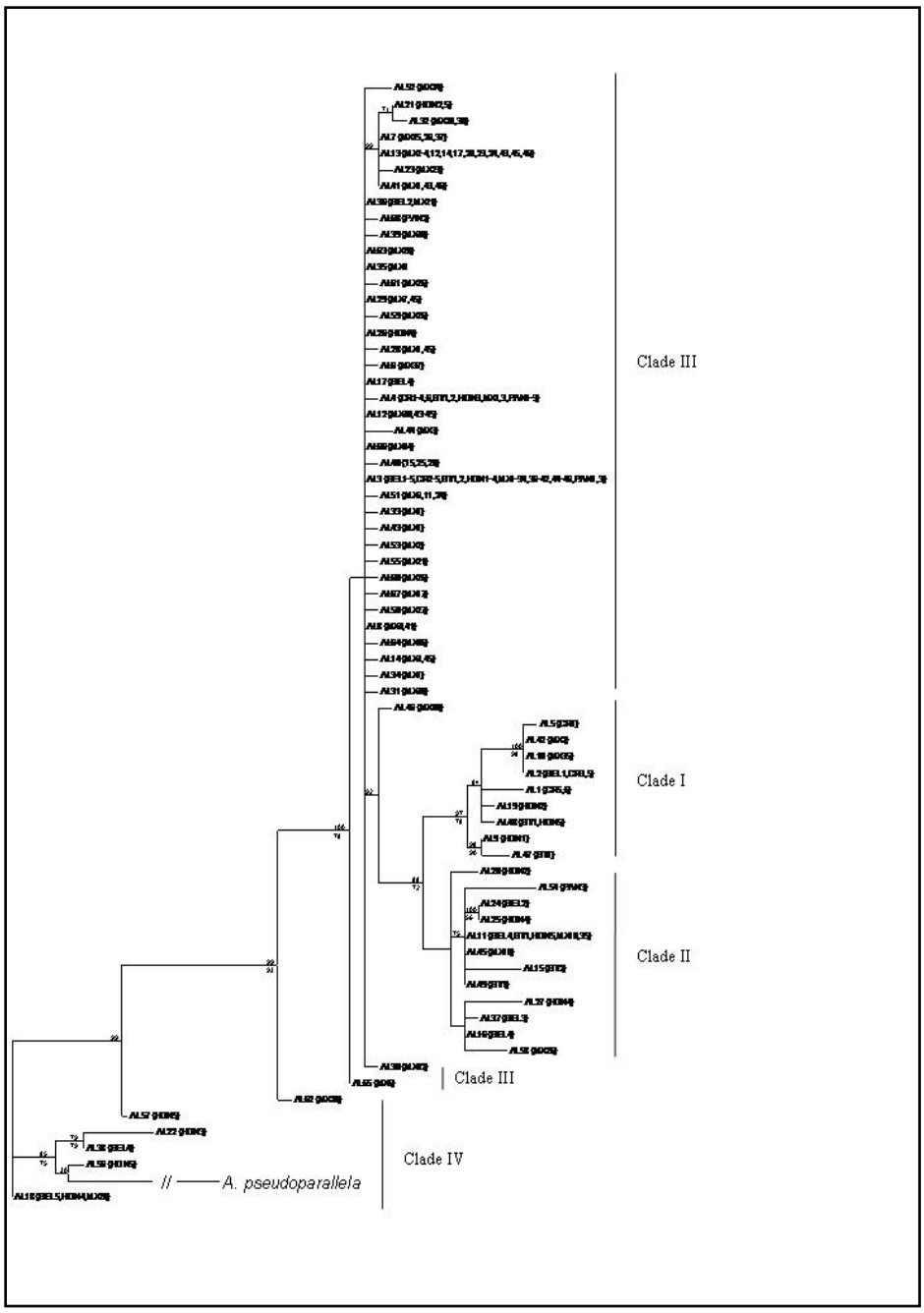


Figure 5. Maximum likelihood tree showing the phylogeny of 68 *Anastrepha ludens* COI+ND6 haplotypes AL01-AL68. Numbers above branches show bootstrap support for ML (1000 replicates) and below for NJ (500 replicates). Asterisks designate unsupported branches with respective method or methods according to Hillis & Bull (1993). Population codes (e.g. MX24) correspond to Table 1. Geographic groups correspond to Fig. 4. Tree was rooted with *A. pseudoparallela*.

Clade II was unsupported with the exception of two internal branches. This clade varied by and average of 3.9 bases. Within clade II both haplotypes AL11 and AL45 were observed in individuals gathered from the Mexican state of Hidalgo (MX10) and AL58 from Nayarit (MX26) both considered from group A (northern) collections. Here, our sampling of MX26 included an additional seven individuals which showed five other haplotypes that differed by 1-3 bases among each other, and 9-10 bases when compared to AL58. A diverse collection of haplotypes as seen in MX10 and MX26 would suggest that perhaps multiple introductions may have occurred at these sites.

Clade III consisted of haplotypes recovered from individuals originating from both geographic groups A and B, with sequences varying by an average of 2.3 bases. In clade III we observed 23 (58%) of the haplotypes were from samples collected from group A, 10 (25%) group B, and 7 (18%) occurring among both. The haplotypes within clade IV represent ancestral haplotypes. Clade IV contained haplotypes that were mostly seen in individuals gathered from Honduras and Belize. These haplotypes varied by an average of 12.8 bases. There were two haplotypes (AL18 and AL62) belonging to one group A collection from Michoacán (MX20) within clade IV. Individuals from this collection showed 6 haplotypes (AL03, AL13, AL18, AL32, AL62, and AL63). Haplotypes AL18 and AL62 differed by 17 bases to each other and by 7 to 23 bases to remaining haplotypes occurring at this collection site. The remaining haplotypes shown by this collection differed by 1 to 3 bases. The haplotypes from collections in MX20 may represent multiple introductions.

In order to estimate the basal haplotype in this assemblage, we rooted the trees with a sequence from *Anastrepha pseudoparallela*, (Fig. 5); the same topology was recovered using other out-group species. The six (AL18, 22, 38, 56, 57, & 62) most ancestral haplotypes seen in cluster IV (Fig. 5), were found in Honduras (4), Belize (2) and Mexico (1). This analysis suggests a potential origin of

this species in Central America, in contrast to Baker et al.'s (1944) assessment of an origin in the Sierra Madre Oriental in northeastern Mexico. The Michoacán samples with the relatively ancestral haplotypes (AL18, AL62) could represent introductions from southern regions under this scenario.

## DISCUSSION

The accurate identification of pests is an essential element in their delineation. Traditionally, pests, such as *A. ludens*, have been identified through classical taxonomic methods. The morphological descriptions require highly trained personnel and methods are not always reliable, more so if the specimen is recovered as an immature. These morphological keys can also be inefficient for distinguishing among cryptic species or individuals belonging to species complexes. Additionally, morphology-based methods are not able to determine the geographic source of interceptions or accurately interpret historical processes because of the limited number of characters as compared to sequencing information. Molecular methods that utilize mitochondrial DNA to identify genetic variation have been used for resolving inter and intra-specific relationships (Alberti et al. 2007, Lanzavecchia et al. 2008), and as a platform for the development of diagnostic methods (Armstrong and Ball 2005; Barr et al. 2006) that provide accurate results within a reasonable time frame. These diagnostic tools have universal application (Folmer et al. 1994) and provide substantial information that permits the identification of intercepted individuals to species and, in some cases, geographic origin. These rapid methods can facilitate time-sensitive pest management decisions.

The phylogeographic analysis of *A. ludens* revealed 68 haplotypes. The presence of one abundant haplotype (AL03) throughout the distribution of the species and the large number of closely-related,

low frequency haplotypes occurring throughout the geographic distribution demonstrate the high degree of homogeneity among populations. A larger percentage of these haplotypes were seen in northern collections. There are, however, substantial differences that suggest a distinction among northern (A) and southern (B) collections.

The results show that both haplotype (A=0.462, B=0.692) and nucleotide (A=0.00061, B=0.00331) diversity estimates for northern and southern collections were significantly different. These data showed the existence of structure among some groups and populations. We observed numerous unique haplotypes characteristic among A and B further suggesting differentiation and limited gene flow among these two groups. Low nucleotide and haplotype diversity, and star-shaped haplotype network with an abundance of low frequency haplotypes all support the notion that populations occurring in the northern range of the species experienced a recent bottleneck event. Nested clade analysis and mismatch distribution views the northern group of haplotypes as undergoing a range expansion and southern collections as stable populations. The significant genetic differentiation and structuring of northern and southern collections could be explained by the restricted gene flow with isolation by distance as was revealed with the nested clade analysis. Differences among these two groups could be attributed to cultural practices and trade restrictions across these regions that might also limit gene flow. The ecological reasons underlying this distinction remain unanswered. The Mantel test showed a pattern of isolation by distance (IBD) among A and B localities collectively. The analysis of the northern (A) and southern (B) groups analyzed individually failed to indicate it. These results could be related to possible reduction in dispersal capability of *A. ludens* and/or due to political barriers and trade restrictions among these regions. The mismatch distribution indicated three periods of expansion for this species when data was analyzed collectively. However, when we analyze groups A and B individually we see that both groups share only two of the three

mismatches, correlating to a time periods, 0.5 – 1.1 mya, well within the Pleistocene epoch. The suggested timing coincides with receding glaciation providing this species the opportunity to expand their geographic range. The high haplotype diversity observed among southern collections suggests this is the source of origin for this fruit fly species. A rooted maximum likelihood analysis showed that the most ancestral haplotype sampled from these data originated in Honduras. Anecdotal evidence using RFLP analysis also suggested that collections from southern Mexico and Central America showed a higher diversity index (Gomes da Silva-Miller unpublished).

In conclusion, we reveal the high degree of homogeneity within the geographic collections with varying levels of structure existing among these populations. The overwhelming presence of one single haplotype minimizes the power of this method for distinguishing interceptions to population-level. Nevertheless, the large number of haplotypes shown by these collections provides an expanded set of geographic-specific diagnostic markers. Other diagnostic methods being developed for this species are examining ISSR (Intersimple sequence repeat) regions (Garza-Farris et al., in publication). The mtDNA data suggests that northern collections along with areas considered the source of this species are still under expansion. We see the divergence of these two groups timed within the Pleistocene, well known for the explosion of flora and fauna in areas inhabited by this species. This growth has provided optimal opportunities for movement of this invasive species into northern areas. Along these lines, we also report the presence of this species on mango and citrus in southern areas. In Costa Rica, trapping of 15,000 specimens showed the absence of *A. ludens* on these hosts as reported by Jirón et al (1988). Those *A. ludens* recovered from white sapote in Costa Rica previously, may have been genetically different, to those individuals recovered on mango and citrus examined here. Also, sixteen specimens gathered in three sites in Panama, where this species was not known to inhabit, showed similar haplotypes as in other southern collections supporting the

idea of a very recent introduction by this species. A better understanding of the diversity, distribution, and historical perspective will lead to more effective methods for managing this invasive pest. We are committed to furthering the application of these and other methods for this and related pest species.

## ACKNOWLEDGEMENTS

We thank Roxanne E. Garza-Farris, Rosita Deleon, Juan D. Vasquez, Lisa A. Ledezma, and Oscar Obregon for providing DNA extracts from the USDA-APHIS-CPHST Mission Laboratory, Mission, TX, Samuel Gum and Oralia Ruiz for technical assistance at The Pennsylvania State University, Deborah S. Grove (Director of the Genomics Core Facility), Ashley Price, and Sheila Plock at The Pennsylvania State University-Huck Institute Nucleic Acid Facility, Jeffrey Feder for providing us microsatellite marker information. Funding for this study came from the USDA-APHIS-CPHST, PIC#M6M02 and USDA/Penn State cooperative agreement 07-8100-1051-CA.

## REFERENCES CITED

**Alberti, A. C., V. A. Confalonieri, R. O. Zandomeni, and J. C. Vilardi. 2008.** Phylogeographic studies on natural populations of the South American fruit fly, *Anastrepha fraterculus* (Diptera: Tephritidae). Genetica 132: 1-8.

**Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990.** Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

**Aluja, M. 1994.** Bionomics and management of *Anastrepha*. Ann. Rev. of Entomol. 39: 155-178.

**Aluja, M., H., C. Hurtado, P. Liedo, M. Cabrera, F. Castillo, J. Guillen, and E. Rios. 1996.** Seasonal population fluctuations and ecological implications for management of *Anastrepha*

fruit flies (Diptera: Tephritidae) in commercial mango orchards in southern Mexico. J. Econ. Entomol. 89: 654-667.

**Armstrong, K. F, and Ball, S. L., 2005.** DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society B-Biological Sciences* **360**, 1813-1823.

**Avise, J.** C. 2004. Molecular Markers, Natural History, and Evolution. Sinauer Associates, Sunderland, MA. 684 pp.

**Baker, A.** C., W. E. Stone, C. C. Plummer, and M. McPhail. 1944. A review of studies on the Mexican fruit fly and related Mexican species. USDA Misc. Publ. 531. U.S. Govt. Printing Office. 155 pp.

**Barr, N.** B., L. Cui, and B. A. McPheron. 2005. Molecular systematics of nuclear gene period in genus *Anastrepha* (Tephritidae). Ann. Entomol. Soc. Am. 98: 173-180.

**Barr, N. B., and B. A. McPheron. 2006.** Molecular phylogenetics of the genus Ceratitis (Diptera : Tephritidae). Mol. Phylogenet. and Evol. 38: 216-230.

**Barr N. B., Copeland R. S., De Meyer M., Masiga D., Kibogo H. G., Billah M. K., Osir E., Wharton R. A., and McPheron B. A. 2006** Molecular diagnostics of economically important *Ceratitis* fruit fly species (Diptera : Tephritidae) in Africa using PCR and RFLP analyses. Bulletin of Entomological Research 96, 505-521.

**Bohonak, A. J. 2002.** IBD (isolation by distance): A program for analyses of isolation by distance. J. Hered. 93: 153-154.

**Bonaccorso, E., Koch, I., and Townsend Peterson, A. 2006.** Pleistocene fragmentation of Amazon species' ranges. Divers. Distrib. 12, 157–164.

**Boykin, L. M., R. G. Shatters, D. G. Hall, R. E. Burns, and R. A. Franqui. 2006.** Analysis of host preference and geographical distribution of *Anastrepha suspensa* (Diptera: Tephritidae) using phylogenetic analyses of mitochondrial cytochrome oxidase I DNA sequence data. Bull. Entomol. Res. 96: 457-469.

**Brower, A. V. Z. 1994.** Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial-DNA evolution. Proceedings of the National Academy of Sciences of the United States of America **91**, 6491-6495.

**Bustos, R., M. E., J. Toledo, and W. Enkerlin. 1993.** Evaluation of irradiation parameters in the quarantine treatment of Mexican mangoes, p. 329–335. *In* International Atomic Energy Agency [eds.], Cost - Benefit Aspects Food Irradiation Processing. Press Vienna.

**Carstens, B. C., Knowles, L. L. 2007.** Shifting distributions and speciation: species divergent

during rapid climate change. Mol. Ecol. 16:619-627.

**Clement, M., D. Posada, and K. A. Crandall. 2000.** TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9: 1657-1659.

**Darwin, C. 1859.** The Origin of Species. John Murray [eds.], London.

**Dowell, R. V., and L. K. Wange. 1986.** Process analysis and failure evidence in fruit fly programs, pp. 43-65. *In* Pest Control: Operation and System Analysis in Fruit Fly management, Springer Verlag 11. Mangel, M., Carey, J. R. [eds.], Plant R. E., NATO, Advanced Science Inst. Ser. G: Ecological Sciences, Berlin.

**Enkerlin, D., L.R. Garcia, and M. F. Lopez. 1989.** Mexico, Central and South America, in Fruit Flies: Their biology, Natural Enemies and Control, pp. 83–90. Vol. 3A, A.S. Robinson &; G. Hooper [eds.], Elsevier Science Publication, Amsterdam, Netherlands.

**Epsky, N. D., R. R. Heath, A. Guzman, and W. L. Meyer. 1995.** Visual cue and chemical cue interactions in a dry trap with food-based synthetic attractant for *Ceratitis capitata* and *Anastrepha ludens* (Diptera: Tephritidae). Environ. Entomol. 24: 1387-1395.

**Eskafi, F. M., and R. T. Cunningham. 1987.** Host plants of fruit-flies (Diptera: Tephritidae) of economic importance in Guatemala. Fla. Entomol. 70: 116-123.

**Excoffier, L., A. Estoup, and J. M. Cornuet. 2005.** Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. Genetics 169: 1727-1738.

**Felsenstein, J. 1985.** Phylogenies from gene-frequencies - a statistical problem. Syst. Zool. 34: 300-311.

**Folmer O., Back M., Hoeh W., Lutz R., and Vrijenhoek R. 1994.** DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **3:** 294–299.

**Foote, R. H., F. L. Blanc and A. L. Norrbom. 1993.** Handbook of the fruit flies (Diptera: Tephritidae) of America north of Mexico. Comstock Publishing Associates, Ithaca.

**Fu, Y. X. 1997.** Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147: 915-925.

**Guindon, S., F. Lethiec, P. Duroux, and O. Gascuel. 2005.** PHYML Online - a web server for fast maximum likelihood-based phylogenetic inference. Nucl. Acids Res. 33: W557-W559.

**Heath, R. R., N. D. Epsky, B. D. Dueben, J. Rizzo, and F. Jeronimo. 1997.** Adding methyl-substituted ammonia derivatives to a food-based synthetic attractant on capture of the Mediterranean and Mexican fruit flies (Diptera: Tephritidae). J. Econ. Entomol. 90: 1584-1589.

**Hernandez, E., D. Orozco, S. F. Breceda, and J. Dominguez. 2007.** Dispersal and longevity of wild and mass-reared *Anastrepha ludens* and *Anastrepha obliqua* (Diptera: Tephritidae). Fla. Entomol. 90: 123-135.

**Hernandez-Ortiz, V., and M. Aluja. 1993.** Lista preliminar del género neotropical *Anastrepha* Schiner (Diptera: Tephritidae) con notas sobre su distribución y plantas hospederas. Folia Entomológica Mexicana 88: 89-105.

**Hillis, D. M. and Bull, J. J. 1993.** An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42: 182–192.

**Jaramillo, C., M. J. Rueda, and G. Mora. 2006.** Cenozoic plant diversity in the Neotropics. Science 311:1893–1896.

**Jirón, L. F., and I. Hedstrom. 1988.** Occurrence of fruit-flies of the genera *Anastrepha* and *Ceratitis* (Diptera: Tephritidae), and their host plant availability in Costa Rica. Fla. Entomol. 71: 62-73.

**Kimura, M. 1983.** Rare variant alleles in the light of the neutral theory. Mol. Biol. and Evol. 1: 84-93.

**Lanzavecchia, S. B., J. L. Cladera, P. Faccio, N. P. Marty, J. C. Vilardi, and R. O. Zandomeni. 2008.** Origin and distribution of Ceratitis capitata mitochondrial DNA haplotypes in Argentina. Ann. Entomol. Soc. Am. 101: 627-638.

**Mantel, N. 1967.** The detection of disease clustering and a generalized regression approach. Cancer Res. 27: 209-220.

**Martinez, A. J., E. J. Salinas, and P. Rendon. 2007.** Capture of *Anastrepha* species (Diptera: Tephritidae) with multilure traps and biolure attractants in Guatemala. Fla. Entomol. 90: 258-263.

**Mayr E. 1942**. Systematics and the Origin of Species. Columbia University Press, New York.

**McPhail, M. 1939.** Protein lures for fruit flies. J. Econ. Entomol. 32: 758-761.

**McPheron, B. A., G. Gasparich, H.-Y. Han, G. J. Steck, and W. S. Sheppard. 1994.** Mitochondrial DNA restriction map for the Mediterranean fruit fly, *Ceratitis capitata*. Biochem. Genet. 32:25-33.

**Nilakhe, S. S., J. N. Worley, R. Garcia, and J. L. Davidson. 1991.** Mexican fruit fly protocol helps export Texas citrus. Subtrop. Plant Sci. 44: 49-52.

**Norrbom, A.** L., C. A. Korytkowski, F. Gonzalez, and B. Orduz. 2005. A new species of *Anastrepha* from Colombia related to Mexican fruit fly (Diptera: Tephritidae). Rev. Colom. Entomol. 31: 67-70.

**Norrbom, A. L., R. A. Zucchi & V. Hernández-Ortiz. 1999.** Phylogeny of the genera *Anastrepha* and *Toxotrypana* (Trypetinae: Toxotrypanini) based on morphology, p. 299-342. *In* M. Aluja and A. L. Norrbom [eds.], Fruit flies (Tephritidae): Phylogeny and evolution of behavior. CRC Press, Boca Raton, Fl.

**Panchal, M., and M. A. Beaumont. 2007.** The automation and evaluation of nested clade phylogeographic analysis. Evol. 61: 1466-1480.

**Pfenninger M. and Posada D, 2002.** Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution Int. J. Org. Evolution*, **56**, 1776–1788.

**Posada, D., and K. A. Crandall. 1998.** MODELTEST: testing the model of DNA substitution. Bioinformatics. 14: 817-818.

**Royama T. 1979.** Effect of adult dispersal on the dynamics of local populations of an insect species: A theoretical investigation. pp. 79-93 *In* A. A. Berryman and L. Satranyik [eds.], Proceedings of the second IUFRO conference on Dispersal of Forest Insects: Evaluation, Theory and Management Implications. Sandpoint, Idaho, 27-31 August 1979. Canadian Forestry Service, Canada.

**Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003.** DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 19: 2496-2497.

**Slatkin, M., and R. R. Hudson. 1991.** Pairwise comparisons of mitochondrial-DNA sequences in stable and exponentially growing populations. Genetics 129: 555-562.

**Smith-Caldas, M. R. B., B. A. McPheron, J. G. Silva, and R. A. Zucchi. 2001.** Phylogenetic relationships among species of the *fraterculus* group (*Anastrepha*: Diptera: Tephritidae) inferred from DNA sequences of mitochondrial cytochrome oxidase I. Neotrop. Entomol. 30: 565-573.

**Steck, G. J. 1991.** Biochemical systematics and population genetic-structure of *Anastrepha fraterculus* and related species (Diptera: Tephritidae). Ann. Entomol. Soc. Am. 84: 10-28.

**Steck, G. J. 1998.** Mexican fruit fly, *Anastrepha ludens* (Loew) (Diptera: Tephritidae). Entomology Circular No. 391, Nov/Dec 1998, Florida Department of Agriculture & Consumer Services. 2 pp.

**Tajima, F. 1989.** Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595.

**Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007.** MEGA4: Mol. Evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24: 1596-1599.

**Templeton, A.R., E. Routman, and C.A. Phillips. 1995.** Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial-DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. Genetics 140: 767–782.

**Templeton, A. R. 1998.** Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Mol. Ecol. 7: 381-397.

**Templeton, A. R., K. A. Crandall, and C. F. Sing. 1992.** A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data .3. cladogram estimation. Genetics 132: 619-633.

**Thomas, D. B., and J. Loera-Gallardo. 1998.** Dispersal and longevity of mass-released, sterilized Mexican fruit flies (Diptera: Tephritidae). Environ. Entomol. 27: 1045-1052.

**Thomas, D. B., T. C. Holler, R. R. Heath, E. J. Salinas, and A. L. Moses. 2001.** Trap-lure combinations for surveillance of *Anastrepha* fruit flies (Diptera: Tephritidae). Fla. Entomol. 84: 344-351.

**Wallace, AR. 1862.** Narrative of a search after birds of paradise. Proc. Zool. Soc. London, 153–161.

**Weir, B. S., and C. C. Cockerham. 1984.** Estimating F-statistics for the analysis of population-structure. Evol. 38: 1358-1370.

**White, I. M., and M. Elson-Harris. 1992.** Fruit flies of economic significance: their identification and bionomics. International Institute of Entomology, London.

**Wright, S. 1951.** The genetic structure of populations, Ann. Eugenics 15: 323–354.

**Zucchi, R. A. 2000.** Taxonomia, pp. 13-24. 327. *In* A. Malavasi & R. A. Zucchi, [eds.], Moscas-das-frutas de importância econômica no Brasil. Conhecimento básico e aplicado. Holos, Riberão Preto.

## Chapter 3: Genetic variation in *Anastrepha obliqua* (Diptera: Tephritidae) in geographic collections from the Mexican state of Veracruz

Raul Ruiz, Md Sajedul Islam, Martin Aluja, and Bruce A. McPheron


### ABSTRACT

There has been considerable interest in understanding evolutionary processes that contribute to diversification of species and populations among tephritids. Only a limited number of studies have examined the genetic diversity of species belonging to the genus *Anastrepha.* To expand our understanding and characterize the influence of ecology, we examine the diversity *A. obliqua.* Specimens for this study were collected during the growing seasons of 1993-1995 in the Mexican state of Veracruz from five host fruit species that have a consecutive phenology. Using sequence-based and microsatellite DNA methods, we examined these collections for discrete differences that may be related to ecological factors. We tested the hypothesis that collections represent a panmictic population vs. populations structured according to spatial, temporal, or host association. This study reveals relationships and historical processes for individuals collected within a small portion of the geographic range of its distribution. We sequenced mitochondrial DNA gene regions from 400 individuals from 52 geographic collections and observed 34 haplotypes. Haplotype and nucleotide diversity were low, with 53% of the individuals showing haplotype OBV01. In order to more effectively detect fine scale differences, we analyzed 719 individuals from 32 of the 52 geographic collections with 17 microsatellite markers. Assignment tests group these individuals into six clusters. Our study reveals moderate genetic diversity with mitochondrial DNA and microsatellite

methods. These data reveal limited genetic structure and suggest that *A. obliqua* populations view a complex host array as a single ecological resource.

## INTRODUCTION

Fruit fly species within the family Tephritidae have a close association with their hosts (Walker *et al*. 2008, Shelly and Villalobos 2004, Forbes *et al*. 2005). A variety of studies has provided substantial evidence that sheds light on the influence of ecology on their evolution (Michel *et al*., 2007, Diegisser *et al*. 2006, Teixeira and Polavarapu, 2003). Ecological mechanisms that stimulate the diversification of individuals within this group have influenced their utilization of alternative resources. These mechanisms have provided avenues for speciation in sympatry (Schwarz *et al*. 2005, Berlocher and Feder 2002). The diversity observed among closely related groups might then be explained in part by the ecological forces at play within close quarters that permit divergence. We see for example that members within Tephritidae occur in tropical, subtropical, and temperate climates. They exhibit a wide range of behavioral and specialized host interactions that facilitate diversification. Host-race formation has been observed among species within this group (Knio *et al*. 2007, Diegisser *et al*. 2006, Bush *et al*. 1989) and reinforces our belief that evolution is mediated by the ecology as well as genetics.

While little is known of ecologically-influenced evolution in the genus *Anastrepha,* much has been studied of the distribution and close-knit relationship between tephritid pest species, their hosts (Novonty *et al*. 2005, Tan & Serit 1994, Malavasi & Morgante 1981) and their behavior (Henning and Matioli 2006, Aluja and Diaz-Fleischer 2006). However, quantifying the level of influence from

the ecology is difficult given the wide range of hosts utilized and habitats where some species within this genus occur. In order to study this phenomenon we examine the relationships of individuals from the species *Anastrepha obliqua* within a relatively small portion of this species' range.

The West Indian fruit fly, occurs from northern Mexico to Southern Brazil and the Caribbean (Steck 2001, Aluja 1994, Foote *et al*. 1993, Hernández-Ortiz and Aluja 1993), and is occasionally found in Texas and California (Epsky *et al*. 2003). This species is designated within the *fraterculus* taxonomic group and is considered a member of the *fraterculus* species complex (Norrbom 2004, Heppner 1991). It is a serious pest of 60 plant species belonging to 24 plant families occurring in tropical and subtropical areas (Aluja *et al*. 1996, Hernández-Ortiz and Aluja 1993, Norrbom and Kim 1988). It's formidable presence throughout the Americas and status as a pest of economic importance throughout this geographic range has warranted the development of various pest management strategies.

There are numerous genetic markers that effectively discriminate among populations and species according to host-association (Cook and Rowell 2007, Berlocher and Feder 2002), and geography (Barr 2009, Hu *et al*., 2008, Gasparich *et al*., 1997) for tephritids. Other methods have been developed to resolve population structure at a microgeographic scale for tephritids (McPheron 1988) and more distantly related dipterans (Bitarello *et al*. 2009, Fraga *et al.* 2003, Ravel *et al.* 2002). This study employs molecular markers that explore for discrete difference among collections of *A. obliqua* gathered from five host species occurring within a transect of approximately 500 km from sites in northern, central, and southern areas in the Mexican state of Veracruz. The three regions display varying ecological characteristics. The central region is typically savanna with a few wetlands, prone to infrequent rain and agriculture is dependent on irrigation (Sanders 1953) and

precipitation is seasonally low with 90 percent of the 1,000 mm of rain occurring during the months of May through October (Garcia 1970). Both northern and southern areas are characterized as humid lowlands contrasting sharply to the central region where a larger portion of the samples were gathered.

Specimens for this study were collected from the host species *Mangifera indica, Spondias purpurea, Spondias mombin, Sterculia apetala* and *Tapirira mexicana*. Ripening of fruit from these host species in Veracruz, Mexico occurs during the months of December through April (*S. apetala*), May through June (*S. purpurea*), June through July (*M. indica*), August through September (*S. mombin*.), and October though November (*T. mexicana*). This consecutive host fruit phenology provides a continual food source for *A. obliqua* within this limited range of its distribution. We propose that one regional population is provided seasonal habitat and migrates from the different host species over time or there are various structured populations existing that pattern their reproductive cycles according to host phenology.

This study examines individuals gathered in Veracruz for population structure resulting from host association and/or other ecological factors. In order to test the hypotheses we analyze fifty two geographic collections. We examine these populations for discrete differences with a sequence-based method. We then increase the number of individuals per collection for thirty two collections and analyze with seventeen microsatellite markers.

## MATERIALS AND METHODS

*Collections and DNA extraction.*

Collections were made during the three growing seasons of 1993-1995 from five host species *Mangifera indica, Spondias purpurea, Spondias mombin, Sterculia apetala* and *Tapirira mexicana* along a 500 km transect from Alamo to Los Tuxtlas in the Mexican state of Veracruz. Infested fruit material was collected from the designated sites and placed in rearing cages. Larvae emerged, pupated, and, upon eclosion, flies were collected and chilled. The specimens were identified morphologically by trained taxonomists, then frozen and shipped to Penn State University, Department of Entomology, University Park, PA. The 52 sample sites are listed in Table 7. A total of 400 individuals were analyzed with mitochondrial DNA (mtDNA) sequencing methods. To increase the likelihood of resolution of population structure, microsatellite loci were also analyzed. A total of 719 individuals from 32 geographic collections were analyzed with microsatellite DNA methods (Table 8). The geographic distribution of collections analyzed with both methods is shown in Fig. 6.

Table 7. Geographical origins and information on collections of *Anastrepha obliqua* gathered in the Mexican state of Veracruz sequenced for this study.

| | Population | | | | | | |
|---|---|---|---|---|---|---|---|
| Site# | Locality | Loc code | *n* | Latitude | Longitude | Month/Year Collected | Host |
| 1 | Actopan | AC3SP | 4 | 19.503N | -96.615W | V-1995 | *Spondias purpurea* |
| 2 | Alamo | AL1SP | 5 | 20.917N | -97.607W | VIII-1995 | *S. purpurea* |
| 3 | Apazapan | AP1SP | 4 | 19.319N | -96.772W | V-1993,VI-1994,?-1995 | *S. purpurea* |
| 4 | Apazapan | AP1CHT | 10 | 19.319N | -96.772W | VII-1994 | trap - *Manilkara sapota* |
| 5 | Apazapan | AP1MT | 10 | 19.319N | -96.772W | VII-1994 | trap – *Mangifera indica* Manila |
| 6 | Apazapan | AP3SP | 9 | 19.324N | -96.721W | VI-1994 | *S. purpurea* |
| 7 | Apazapan | AP3MI | 10 | 19.325N | -96.721W | | *M. indica* var. Criollo |
| 8 | Apazapan | AP4SP | 9 | 19.322N | -96.720W | VI-1994 | *S. purpurea* |
| 9 | Apazapan | AP5MI | 6 | 19.320N | -96.721W | VII-1994 | *M. indica* var. Criollo |
| 10 | Apazapan | AP6SP | 10 | 19.320N | -96.718W | IX-1994 | *S. purpurea* |
| 11 | Chichicaxtle | CHISP | 7 | 19.342N | -96.354W | VI-1994, ?-1994 | *S. purpurea* |
| 12 | Chicoazen | CHI1SP | 10 | 19.504N | -96.597W | VI-1994 | *S. purpurea* |
| 13 | Jalcomulco | JA1SP | 10 | 19.334N | -96.744W | VI-1994, ?-1994 | *S. purpurea* |
| 14 | Jalcomulco | JA2SM | 10 | 19.342N | -96.766W | X-1994 | *Spondias mombin* |
| 15 | Jalcomulco | JA3MI | 5 | 19.330N | -96.763W | VII & VII-1994 | *M. indica* var. Manila |
| 16 | Jalcomulco | JA4SM | 9 | 19.333N | -96.744W | X-1993, X-1994 | *S. mombin* |
| 17 | Llano Grande | LLA1ASM | 9 | 19.363N | -96.865W | X-1994 | *S. mombin* |
| 18 | Llano Grande | LLA1BSM | 7 | 19.362N | -96.865W | X & XI-1994 | *S. mombin* |
| 19 | Llano Grande | LLA1CSM | 3 | 19.363N | -96.864W | X-1994 | *S. mombin* |
| 20 | Llano Grande | LLA1DSM | 2 | 19.363N | -96.865W | X-1994 | *S. mombin* |
| 21 | Llano Grande | LLA1ESM | 3 | 19.362N | -96.865W | X-1994 | *S. mombin* |
| 22 | Llano Grande | LLA1MI1 | 9 | 19.363N | -96.865W | VII-1994 | *M. indica* var. Manila |
| 23 | Llano Grande | LLA1MI2 | 11 | 19.363N | -96.865W | VI & VII-1994 | *M. indica* var. Criollo |
| 24 | Llano Grande | LLA1MI3 | 7 | 19.363N | -96.865W | VII-1994 | *M. indica* var. Melon |
| 25 | Llano Grande | LLA1MI4 | 8 | 19.363N | -96.865W | | *M. indica* var. Kent |
| 26 | Llano Grande | LLA2MI | 7 | 19.359N | -96.866W | VII-1994 | *M. indica* var. Tocotin |
| 27 | Martinez de la Torre | MAR1SM | 8 | 20.060N | -97.055W | X-1994 | *S. mombin* |
| 28 | Monte Blanco | MBSM | 7 | 18.966N | -97.092W | X-1993, X-1994 | *S. mombin* |
| 29 | Monte Blanco | MBTM | 9 | 18.970N | -90.020W | XI-1993 | *Tapirira mexicana* |
| 30 | Playa Azul | PAZMI | 9 | 19.544N | -96.391W | VII & VIII-1994 | *M. indica* var. Kent |
| 31 | Playa Azul | PAZT | 9 | 19.544N | -96.391W | | trap |
| 32 | Palo Gacho | PGSP | 7 | 19.535N | -96.464W | V-1994 | *S. purpurea* |
| 33 | Puente Nacional | PNSP | 7 | 19.339N | -96.494W | VI-1994 | *S. purpurea* |
| 34 | Puente Nacional | PNMI | 5 | 19.339N | -96.494W | VI-1994 | *M. indica* var. Manila |
| 35 | Paso de Ovejas | POSP | 4 | 19.281N | -96.437W | VI-1994 | *S. purpurea* |
| 36 | Plan del Rio | PRSP | 10 | 19.414N | -96.709W | VI-1994 | *S. purpurea* |
| 37 | Paso de San Juan | PSJSP | 10 | 19.196N | -96.338W | VI-1994 | *S. purpurea* |
| 38 | Rinconada | RISP | 8 | 19.359N | -96.572W | VI-1994 | *S. purpurea* |
| 39 | Tejeria | TE1MI1 | 3 | 19.386N | -96.964W | VIII-1994 | *M. indica* var. Manila |
| 40 | Tejeria | TE1MI2 | 10 | 19.386N | -96.964W | VII & VII-1994 | *M. indica* var. Criollo |

Table 7. (cont.)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 41 | Tolome | TOSP | 8 | 19.264N | -96379W | VI-1994 | *S. purpurea* |
| 42 | Tuzamapan | TU1SM | 10 | 19.413N | -96.845W | X-1994 | *S. mombin* |
| 43 | Tuzamapan | TU1MI | 3 | 19.413N | -96.845W | VIII-1994 | *M. indica* var. Criollo |
| 44 | Tuzamapan | TU2SM | 10 | 19.413N | -96.845W | X-1994 | *S. mombin* |
| 45 | Tuzamapan | TU3SM | 8 | 19.413N | -96.845W | X-1994 | *S. mombin* |
| 46 | Los Tuxtlas | TUX1SM | 9 | 18.560N | -95.204W | IX-1994, IX-1995 | *S. mombin* |
| 47 | Los Tuxtlas | TUX2SM | 9 | 18.560N | -95.204W | IX-1994 | *S. mombin* |
| 48 | Los Tuxtlas | TUX2SA | 10 | 18.560N | -95.204W | | *Sterculea apetala* |
| 49 | Los Tuxtlas | TUX3SA | 10 | 18.560N | -95.204W | IX-1994 | *S. apetala* |
| 50 | Los Tuxtlas | TUX4SM | 9 | 18.560N | -95.204W | IX-1994 | *S. mombin* |
| 51 | Los Tuxtlas | TUX5TM | 6 | 18.560N | -95.204W | XI-1994 | *T. mexicana* |
| 52 | Los Tuxtlas | TUX6SM | 8 | 18.560N | -95.204W | XI-1994 | *S. mombin* |

*n*, number of individual specimens analyzed. Blank fields indicate missing information.

Table 8. Geographical origins and information on collections of *Anastrepha obliqua* gathered in the state of Veracruz analyzed with microsatellite DNA methods.

| | Population | | | | | | |
|---|---|---|---|---|---|---|---|
| Site# | Locality | Loc code | *n* | Latitude | Longitude | Month/Year Collected | Host |
| 1 | Actopan | AC3Mi | 24 | 19.503N | -96.615W | V-1995 | *Spondias purpurea* |
| 2 | Alamo | AL1SP | 24 | 20.917N | -97.607W | VIII-1995 | *S. purpurea* |
| 3 | Apazapan | AP1SP | 30 | 19.319N | -96.772W | V-1993,VI-1994,?-1995 | *S. purpurea* |
| 4 | Apazapan | AP3MI | 24 | 19.325N | -96.721W | | *Mangifera indica* var. Criollo |
| 5 | Chichicaxtle | CHISP | 20 | 19.342N | -96.354W | VI-1994, ?-1994 | *S. purpurea* |
| 6 | Chicoazen | CHI1SP | 30 | 19.504N | -96.597W | VI-1994 | *S. purpurea* |
| 7 | Jalcomulco | JA1SP | 30 | 19.334N | -96.744W | VI-1994, ?-1994 | *S. purpurea* |
| 8 | Jalcomulco | JA2SM | 34 | 19.342N | -96.766W | X-1994 | *Spondias mombin* |
| 9 | Jalcomulco | JA1SM | 5 | 19.342N | -96.766W | X-1994 | *S. mombin* |
| 10 | Jalcomulco | JA3MI | 7 | 19.330N | -96.763W | VII & VII-1994 | *M. indica* var. Manila |
| 11 | Jalcomulco | JA4SM | 31 | 19.333N | -96.744W | X-1993, X-1994 | *S. mombin* |
| 12 | Llano Grande | LLA1ASM | 5 | 19.363N | -96.865W | X-1994 | *S. mombin* |
| 13 | Llano Grande | LLA1BSM | 30 | 19.362N | -96.865W | X & XI-1994 | *S. mombin* |
| 14 | Llano Grande | LLA1MI2 | 30 | 19.363N | -96.865W | VI & VII-1994 | *M. indica* var. Criollo |
| 15 | Martinez de la Torre | MAR1SM | 28 | 20.060N | -97.055W | X-1994 | *S. mombin* |
| 16 | Monte Blanco | MBSM | 30 | 18.966N | -97.092W | X-1993, X-1994 | *S. mombin* |
| 17 | Monte Blanco | MBTM | 13 | 18.970N | -90.020W | XI-1993 | *Tapirira mexicana* |
| 18 | Playa Azul | PAZMI | 24 | 19.544N | -96.391W | VII & VIII-1994 | *M. indica* var. Kent |
| 19 | Palo Gacho | PGSP | 24 | 19.535N | -96.464W | V-1994 | *S. purpurea* |
| 20 | Puente Nacional | PNSP | 30 | 19.339N | -96.494W | VI-1994 | *S. purpurea* |
| 21 | Puente Nacional | PNMI | 9 | 19.339N | -96.494W | VI-1994 | *M. indica* var. Manila |
| 22 | Paso de Ovejas | POSP | 13 | 19.281N | -96.437W | VI-1994 | *S. purpurea* |
| 23 | Plan del Rio | PRSP | 24 | 19.414N | -96.709W | VI-1994 | *S. purpurea* |
| 24 | Paso de San Juan | PSJSP | 10 | 19.196N | -96.338W | VI-1994 | *S. purpurea* |
| 25 | Rinconada | RISP | 18 | 19.359N | -96.572W | VI-1994 | *S. purpurea* |
| 26 | Tejeria | TE1MI2 | 21 | 19.386N | -96.964W | VII & VII-1994 | *M. indica* var. Criollo |
| 27 | Tolome | TOSP | 24 | 19.264N | -96379W | VI-1994 | *S. purpurea* |
| 28 | Tuzamapan | TU3SM | 24 | 19.413N | -96.845W | X-1994 | *S. mombin* |
| 29 | Los Tuxtlas | TUX1SM | 25 | 18.560N | -95.204W | IX-1994, IX-1995 | *S. mombin* |
| 30 | Los Tuxtlas | TUX2SM | 24 | 18.560N | -95.204W | IX-1994 | *S. mombin* |
| 31 | Los Tuxtlas | TUX3SA | 24 | 18.560N | -95.204W | IX-1994 | *Sterculia apetala* |
| 32 | Los Tuxtlas | TUX6SM | 30 | 18.560N | -95.204W | XI-1994 | *S. mombin* |

*n*, number of individual specimens analyzed. Blank fields indicate missing information.

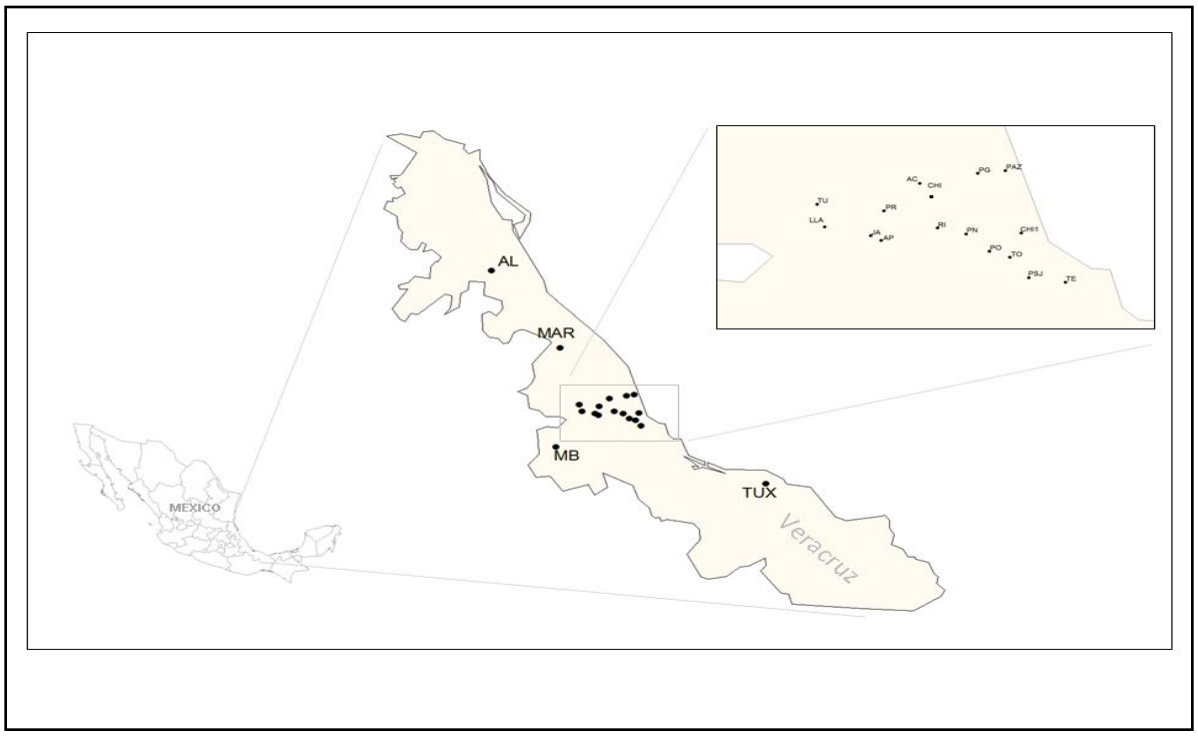


Figure 6. Map showing the distribution of collection sites for *Anastrepha obliqua* specimens used in this study. Rectangle delineates collection sites located in central Veracruz, MX. Collection site ID values (e.g. AL) correspond to locality code prefixes in Tables 7 and 8.

Entire bodies from adult and larvae specimens were maintained at –80$^{o}$C prior to analysis. Specimen DNA was extracted using the DNeasy® Blood and Tissue Kit (Qiagen) following standard DNeasy guidelines for animal tissues. All extractions were stored at –20$^{o}$C for the duration of the study. Whenever possible, DNA from one or more specimens per collection was extracted via a non-destructive method (Barr *et al*. 2006) to serve as a voucher.

*Mitochondrial DNA PCR and sequencing analysis*.

All samples were amplified for the cytochrome oxidase subunit I (COI) region using C1-J-1751 5’-GGATCACCTGATATAGCATTCCC-3’ and C1-N-2191 5’-CCCGGTAAAATTAAAATATAAACTTC-3’(Simon *et al*. 1994) with a fragment of approximately 550 base pairs sequenced. The NADH subunit I (NDI) mitochondrial region was amplified using

N1-J-11861 5'-ATCATAACGAAAYCGAGGTAA-3' and N1-N-12530 5'-CAACCTTTTWGTGATGC-3' primers (Smith *et al*. 1999) and yielded a fragment of approximately 550 base pairs. PCR reactions were performed in 25 µL reactions containing 1 µL of template, 16.875 µL of water, 2.5 µL of buffer (10X conc., Qiagen), 1.5 µL of 25 mM of $MgCl_2$ (Qiagen), 1.25 µL of each primer (5 µM, Operon Technologies), 0.5 µL of 10 mM dNTP mix (Promega), and 0.125 µL of Taq DNA polymerase (Qiagen) to complete the final volume of 25 µL. Amplifications were performed in Applied Biosystems GeneAmp® PCR *System* 9700 thermal cyclers. Cycling conditions for amplification of COI and NDI fragments was 3 min at 94°C followed by 39 cycles of 1 min at 94°C, 1 min at 50°C, 1 min at 72°C, and a final extension of 10 min at 72°C. PCR products were stained with Sybr® Green (Invitrogen) fluorescent dye at 1/10,000X and 3/10,000X then loaded on 1.5% electrophoresis agarose gels. Documentation of these gels was via a Gel Doc™ (Bio-Rad) imaging system using Quantity One® software. Amplification products were purified with ExoSAP-IT (USB Corp.) prior to sequencing. PCR products were sequenced asymmetrically using 3' BigDye-labeled dideoxynucleotide triphosphates (v 3.1 dye terminators, Applied Biosystems) and run on an ABI 3730XL DNA Analyzer with the ABI Data Collection Program (v 2.0) at the Huck Institute's Nucleic Acid Facility at The Pennsylvania State University (PSU).

Sequences were edited with QScreen (PSU: Huck Institute) designed in collaboration with the McPheron Lab (PSU: Dept of Entomology), and Sequencher 4.8 (Gene Codes Corp.). All sequences were aligned with Mega 4.0.2 (Tamura *et al*. 2007) using ClustalW (Thompson *et al*. 1994) and trimmed to a length of 528 bp and 527 bp for COI and NDI, respectively. COI and NDI sequences were concatenated using DnaSP 4.10 (Rozas *et al*. 2003), and subsequent analyses were carried out on the 1055 bp fragment. Haplotypes were identified from these combined sequences with DnaSP. The relative haplotype frequencies within populations were calculated using

ARLEQUIN 3.11 (Excoffier *et al*. 2005). Basic molecular summary statistics and mismatch distributions and genetic distances were estimated with DnaSP.

*mtDNA phylogenetic reconstruction*

Phylogenetic tree reconstruction was estimated using Maximum Likelihood and Maximum Parsimony approaches. The appropriate model of sequence evolution for phylogenetic analysis was determined with hierarchical likelihood tests (hLRTs) using Modeltest 3.7 (Posada and Crandall 1998). The HKY+I+G substitution model was found to be most applicable to the concatenated COI and NDI sequence based on the Akaike ranking. Modeltest revealed a proportion of invariable sites (I) of 0.8138, transition to transversion ratio of 6.841, and the gamma shape distribution shape parameter of 0.4530. MP phylogenetic reconstruction was performed by bootstrapping with 500 replications using Mega. ML bootstrap analyses were conducted with 500 replicates (Felsenstein 1985). The Phyml (Guindon *et al*. 2005) ML tree search was initiated using a NJ tree that was constructed with Mega. We rooted these trees with one individual of *Anastrepha ludens*.

*mtDNA phylogeographic analysis*

A maximum parsimony haplotype network was constructed according to the algorithm described in Templeton *et al*. (1992) and included in the program TCS 1.18 (Clement *et al*. 2000). TCS was used to generate schematics representative of parsimonious relationships between and among collections according to geographic association. A nested clade phylogeographic analysis NCPA (Excoffier *et al*. 1992, Templeton *et al*. 1995) was performed using ANeCA (Panchal and Beaumont, 2007) to differentiate between population structure and historical events (Templeton, 1998). Nested clade contingency tests were performed in GeoDis v2.5 (Posada *et al*. 2000) and the 2005 inference key

(11 Nov 2005 at http://darwin.uvigo.es) was used to determine whether phylogeographical differences were best explained by allopatric associations, recurrent but restricted geneflow, or range expansion. Recently, a debate concerning the legitimacy of NCPA has occurred (Templeton 2008, 2009, Knowles 2008, Beaumont & Panchal 2008, Petit 2008) although Templeton (2009) addressed many of the objections raised by other authors. In order to provide rigorous testing of data and to authenticate results from NCPA we use additional population genetic analyses.

The demographic parameters Tajima's D (Tajima, 1989) and Fu's $F_S$ (Fu, 1997) were estimated for each collection and parameter confidence intervals were assessed by 1,000 coalescent simulations to test the hypothesis that all mutations are selectively neutral (Kimura, 1983) with ARLEQUIN. ARLEQUIN was used to calculate $F_{ST}$ (Wright 1951, Weir and Cockerham 1984) measures and source of variation among geographic groups, among populations within groups, and within populations using analysis of molecular variance (AMOVA, Excoffier *et al*. 2005). Significance testing for variance components was based on 10,000 permutations. The relationship (isolation by distance) between mtDNA divergence ($F_{ST}$) and geographic distance for populations was tested using Mantel tests (Mantel, 1967) with 999 replications in GenAlex 6.2 (Peakall and Smouse 2006).

*Microsatellite PCR and fragment analysis*

DNA was amplified with 17 microsatellite makers (Table 3). Reaction components consisted of 0.5 μL of template, 10.5 μL of water, 1.5 μL of buffer (10X conc., Qiagen), 0.7 μL of 25 mM of $MgCl_2$ (Qiagen), 0.6 μL of each primer (5 μM, Operon Technologies), 0.6 μL of 10 mM dNTP mix (Promega), and 0.08 μL of Taq DNA polymerase (Qiagen) to complete the final volume of 15.08 μL. Amplifications were performed in Applied Biosystems GeneAmp® PCR System 9700 thermal

cyclers. Cycling conditions for amplification of fragments was 3 min at 94$^{o}$C followed by 39 cycles of 1 min at 94$^{o}$C, 1 min at 50$^{o}$C (primer set dependent, see Table 9), 1 min at 72$^{o}$C, and a final extension of 10 min at 72$^{o}$C. PCR amplicons were verified electrophoresis agarose gels prior to fragment analysis. Each forward primer was 5' end-labeled with either 6FAM or VIC (Applied Biosystems) fluorescent dye for laser beam detection of DNA fragments with Applied Biosystems 3730XL. Allele size was scored against the internal-lane 500 ROX standard (Bioventures) using GeneMapper® software (Applied Biosystems).

Table 9. Flanking primers annealing temperature, observed size range, number of alleles observed, and fluorescence dye used for detection on the ABI3730XL and reference for the 17 microsatellite loci used to genotype 719 *Anastrepha obliqua*.

| Locus | Repeat | Primer | Temp | Size | Alleles | Dye | Reference |
|---|---|---|---|---|---|---|---|
| Anob01 | $(CATA)_4(TG)_5..A_{10}$ | 5'-GCACCAGTTCGCAAATAGTCACCG-3'<br>5'-ACACACACACACATATATATAT-3' | 56 | 213-248 | 21 | 6-Fam | Islam *et al.* (in preparation) |
| Anob02 | $(AC)_6(ATAC)_2..A_5$ | 5'-TTTATGGTCGCACGATTTTCC-3'<br>5'-CTCCCACTAGCAATTAGCACC-3' | 56 | 146-162 | 9 | 6-Fam | Islam *et al.* (in preparation) |
| Anob03 | $(CT)_5$ | 5'-CGGTTTGGACTACCCAACTCACTT-3'<br>5'-ACACACACACACAGAGAGAGAG-3' | 56 | 153-182 | 15 | 6-Fam | Islam *et al.* (in preparation) |
| Anob04 | $C_4..C_6..(CT)_7(GT)_4$ | 5'-GGTGAACCTCATCACGCAGCATAA-3'<br>5'-ACACACACACACAGAGAGAGAG-3' | 60 | 170-221 | 26 | 6-Fam | Islam *et al.* (in preparation) |
| Anob06 | $(CT)_5$ | 5'-GGTGAACCAACTCTTGACACGTAA-3'<br>5'-ACACACACACACAGAGAGAGAG-3' | 58 | 156-198 | 21 | 6-Fam | Islam *et al.* (in preparation) |
| Anob08 | $(CT)_6..(GT)_2..(GT)_4$ | 5'-GGTTTCATTAAACCATCATTGGCG-3'<br>5'-ACACACACACACAGAGAGAGAG-3' | 58 | 224-254 | 10 | 6-Fam | Islam *et al.* (in preparation) |
| Anob12 | $(CT)_5..(GT)_3..A_6$ | 5'-CGTCGTTCCAGTCACGTTGGTAAC-3'<br>5'-ACACACACACACAGAGAGAGAG-3' | 54 | 253-287 | 10 | 6-Fam | Islam *et al.* (in preparation) |
| Anob15 | $(AG)_8$ | 5'-GGACAGTCTGACACTGAGTTACTG-3'<br>5'-ACACACACACACAGAGAGAGAG-3' | 58 | 146-194 | 24 | 6-Fam | Islam *et al.* (in preparation) |
| Anob17 | $(AG)_6$ | 5'-CGGCTCTCGTTTTCACGCCTACCA-3'<br>5'-ACACACACACACAGAGAGAGAG-3' | 54 | 54-76 | 11 | 6-Fam | Islam *et al.* (in preparation) |
| Anob18 | $(AT)_3..A_6..(AT)_3..(AC)_5$ | 5'-CTCATTAAAACTGACAACTGTATT-3'<br>5'-AGAGAGAGAGAGACACACACAC-3' | 54 | 214-242 | 13 | 6-Fam | Islam *et al.* (in preparation) |
| A-1 | $(AC)_5..(AC)_6$ | 5'-TGGTCATCTCAGTTTAAGTTCGTT-3'<br>5'-TGTAAAACTTCTTCCGGGCTTA-3' | 50 | 123-131 | 5 | 6-Fam | Feder (unpublished) |
| A-3 | $(GT)_{11}..(GT)_2$ | 5'-ACTAAGTTTTGCGCGGAATG-3'<br>5'-CTGACAGACACTCGGACAGC-3' | 50 | 127-171 | 15 | Vic | Feder (unpublished) |
| A-7 | $(GT)_9$ | 5'-GCAACGTGCTTTTGGCTAAT-3'<br>5'-CGAACGGTAGAAATGTATGCAA-3' | 50 | 255-273 | 11 | 6-Fam | Feder (unpublished) |
| A-20 | $(CA)_{10}$ | 5'-GCTAGCCAAATACACGCACA-3'<br>5'-GAAACAGGCTTTAGCCCACA-3' | 50 | 202-224 | 13 | 6-Fam | Feder (unpublished) |
| A-27 | $(AC)_{13}..(AC)_3$ | 5'-TCCGTGAGTGGCGTACAATA-3'<br>5'-CCCTCGGCTGTAGTTGTGTT-3' | 55 | 153-196 | 19 | 6-Fam | Feder (unpublished) |
| A-28 | $(TG)_9$ | 5'-ACGTACGCGCACAATACAAA-3'<br>5'-GTGGGAGTATGGCGGTAAGA-3' | 55 | 253-283 | 14 | 6-Fam | Feder (unpublished) |
| Asus 1–3B | $(ACTC)_4$ | 5'-CGTTCAGCATTACTTTGA-3'<br>5'-CTTATTTGGAAGTGACTGA-3' | 55 | 104-128 | 10 | Vic | Fritz and Schable (2004) |

*Microsatellite data analyses*

Patterns of allelic diversity were examined at 17 loci. Descriptive statistics were performed with Microsatellite toolkit 3.1.1 (Park 2001). Departures from Hardy-Weinberg and linkage disequilibrium between loci were tested using FSTAT 2.9.3 (Goudet 1995) incorporating the Bonferroni correction method. We tested for deviations from HW and linkage equilibrium by randomizing alleles 1000 times across all loci and samples.

We use assignment tests for these data based on its ability to better separate from among mixed groups as compared to phylogenetic analysis. Our hypothesis while presented in its most simplistic form may not properly represent the outcome of clustering by assignment tests. Since clustering is based on a non-biased assignment of individuals, a more complex combination of ecological scenarios may be at play for these collections. Subclusters may not be detected through phylogenetic approaches. Also, methods based on pairwise genetic distance matrices lose information by collapsing all genotype data into a single value. STRUCTURE has been designed overcome these limitations (Pritchard *et al*. 2000; Falush *et al*. 2003) and been used to infer population structure in numerous species (Falush *et al*. 2003a; Mun *et al*. 2003; Tsuitstui *et al*. 2001).

In order to evaluate differences among samples collected in this region and account for variation among collections resulting from ecological parameters, we employed assignment test methods implemented in STRUCTURE 2.2 (Pritchard *et al*. 2000). STRUCTURE assumes that populations are in Hardy-Weinberg and that loci are not linked. In order to arrive at the optimal *K* value, we first tested our data (*K* 1 through 32) for convergence and appropriate number of

repetitions. Once estimates for convergence and repetition were determined we used the admixture model of ancestry and performed the analysis considering the correlated allele frequency model with a burn in length of 10,000 repetitions and a simulation of 80,000 iterations. We tested our data with STRUCTURE using 10 repetitions of *K*1 through 12 to arrive at the most probable number of putative populations (*K*) that may explain patterns of variation. The optimal *K*-value was chosen according to the maximum likelihood $L(K)$ and $\Delta K$ resulting from analyses by STRUCTURE and statistical methods as outlined by Evanno *et al*. (2005). The optimal probabilities for all individuals were estimated from 10 replicate runs at $K$=6 with permutation analysis using CLUMPP 1.1.1 (Jakobsson and Rosenberg 2007).

We examined these data for source of variation among groups, among populations within groups, and within populations using analysis of molecular variance (AMOVA). Significance testing for variance components was based on 10,000 permutations. AMOVA was performed on various hypothesized models of population structure in order to determine the most likely source of variation.

## RESULTS

### *Sequence variation*

A total of 34 composite (COI+NDI) haplotypes were observed in 400 individuals (Table 7). There were 28 polymorphic sites (Table 10) in the 1055bp concatenated fragment with no gaps or indels detected. The COI (528bp) and NDI (527bp) fragments contributed 10 and 18 polymorphic sites, respectively. Twenty-four of the 34 haplotypes were unique to a single collection (Table 11). The

haplotype and nucleotide diversities were 0.6846 and 0.0017, respectively, with haplotypes differing by 1-8 bases. Haplotype OBV01, the most common, was seen in 53.3% of the individuals, and the next most common haplotype was OBV08 in 14.8% of the individuals (Fig. 7).

Table 10. Alignment showing the 34 haplotypes and 28 variable sites for COI (528bp) and ND1 (527bp) concatenated sequences observed in *Anastrepha obliqua* collections.

| Positions | COI (10) | ND1 (18) |
|---|---|---|
| | | **1** |
| | **111244455** | **555556677778999990** |
| | **1019859912** | **356890214693145573** |
| | **1934423775** | **813404552534032520** |
| OBV01 | CGTGCAATGC | CATTAAAGTCAGTGAGGC |
| OBV02 | .......... | ..C............... |
| OBV03 | .......... | ...C..G.........T |
| OBV04 | ..G........ | ...C..G........... |
| OBV05 | .......... | ...C..G......A.... |
| OBV06 | T........T | ...C..G........... |
| OBV07 | ....T..... | .................. |
| OBV08 | .......... | ...C..G........... |
| OBV09 | .........T | ...C..G........... |
| OBV10 | T........T | ...C..G.........T |
| OBV11 | .........T | .................. |
| OBV12 | .......... | .........T........ |
| OBV13 | .......... | ...C..G.......G... |
| OBV14 | ....T..... | T..C..G........... |
| OBV15 | ......G... | ...C..G....A.....T |
| OBV16 | .......... | ...C..G....A.....T |
| OBV17 | TA.ATG..A. | .................. |
| OBV18 | T......... | ...C..G........... |
| OBV19 | ....T..... | ...C..G.........T |
| OBV20 | TA.ATG..A. | ...C..G........... |
| OBV21 | .......... | ..........A...... |
| OBV22 | .......A.. | .................. |
| OBV23 | .......... | .C................ |
| OBV24 | ....T..... | ...C..G........... |
| OBV25 | .......... | ...C..G........A.. |
| OBV26 | .......... | ............AA.... |
| OBV27 | .......... | ...C..G.........A. |
| OBV28 | .......... | ...C..G....A...... |
| OBV29 | .......... | ...C..G....A.A.... |

Table 10. (cont.)

| Positions | COI (10) | ND1 (18) |
|---|---|---|
| | | 1 |
| | 111244455 | 555556677778999990 |
| | 1019859912 | 356890214693145573 |
| | 1934423775 | 813404552534032520 |
| OBV01 | CGTGCAATGC | CATTAAAGTCAGTGAGGC |
| OBV30 | .......... | ...CG.G.......G... |
| OBV31 | .......... | ...C..GA.......... |
| OBV32 | .......... | ...C..G...G...G... |
| OBV33 | .......... | ...C..G.C......... |
| OBV34 | .......... | .....T............ |

Table 11. The geographic distribution of 34 haplotypes observed among the *A. obliqua* collection localities examined. Population locale (e.g. AC) correspond to Table 7. GeneBank accession numbers to COI and ND6 sequences that composed the 34 concatenated haplotypes of *A. obliqua* are shown.

| Haplotype | COI GenBank no. | NDI GeneBank no. | Origin of population (*n* = number of individuals) |
|---|---|---|---|
| OBV01 | HM592628 | HM592641 | AC(2),AL(1),AP(44),CHI1(4),CHI(4), JA(20),LLA(49),MAR(1),MB(10),PAZ(11),PG(2), PN(1),PO(1),PR(5),PSJ(7)RI(4),TE(10),TO(2),TU(18), TUX(17) |
| OBV02 | HM592628 | HM592642 | AL(1) |
| OBV03 | HM592628 | HM592643 | AL(2),AP(4),CHI(1),JA(1),LLA(1),MAR(3),PAZ(3), PG(1),PN(1),PR(1),RI(3),TE(1), TO(3),TU(1),TUX(5) |
| OBV04 | HM592629 | HM592644 | AL(1),LLA(1),MAR(1),MB(1),TUX(3) |
| OBV05 | HM592628 | HM592645 | LLA(1),MAR(1),TUX(1) |
| OBV06 | HM592630 | HM592644 | AP(5),CHI1(2),JA(1),LLA(4),MB(2),PN(2),TE(1),TO(2),TU(3) |
| OBV07 | HM592631 | HM592641 | AP(1) |
| OBV08 | HM592628 | HM592644 | AC(2),AP(2),CHI1(1),CHI(4),JA(4),LLA(7),MB(1), PAZ(1),PN(5),PO(1),PR(3),RI(1), TE(1),TO(1),TU(6),TUX(19) |
| OBV09 | HM592632 | HM592644 | AP(4),JA(2),PG(1),TU(2),TUX(3) |
| OBV10 | HM592630 | HM592643 | JA(2) |
| OBV11 | HM592632 | HM592641 | JA(1) |
| OBV12 | HM592628 | HM592646 | TUX(1) |
| OBV13 | HM592628 | HM592647 | AP(3),PAZ(2),PSJ(1),TUX(1) |
| OBV14 | HM592631 | HM592648 | TUX(1) |
| OBV15 | HM592633 | HM592649 | TUX(1) |
| OBV16 | HM592628 | HM592649 | AP(4),JA(2),MAR(1),MB(1),PAZ(1),PN(1),PO(2),TUX(2) |
| OBV17 | HM592634 | HM592641 | TUX(1) |
| OBV18 | HM592635 | HM592644 | PR(1),TUX(1) |
| OBV19 | HM592631 | HM592643 | TUX(1) |

Table 11. (cont.)

| Haplotype | COI GenBank no. | NDI GeneBank no. | Origin of population (*n* = number of individuals) |
|---|---|---|---|
| OBV20 | HM592634 | HM592644 | TUX(1) |
| OBV21 | HM592628 | HM592650 | AP(1),PG(3) |
| OBV22 | HM592636 | HM592641 | AP(1) |
| OBV23 | HM592628 | HM592651 | AP(1) |
| OBV24 | HM592631 | HM592644 | AP(1),LLA(1),PN(1) |
| OBV25 | HM592628 | HM592652 | CHI(1) |
| OBV26 | HM592628 | HM592653 | LLA(1) |
| OBV27 | HM592628 | HM592654 | LLA(1) |
| OBV28 | HM592628 | HM592655 | MAR(1) |
| OBV29 | HM592628 | HM592656 | TU(1) |
| OBV30 | HM592628 | HM592657 | PN(1) |
| OBV31 | HM592628 | HM592658 | PSJ(1) |
| OBV32 | HM592628 | HM592659 | PSJ(1) |
| OBV33 | HM592628 | HM592660 | JA(1) |
| OBV34 | HM592628 | HM592661 | MB(1) |

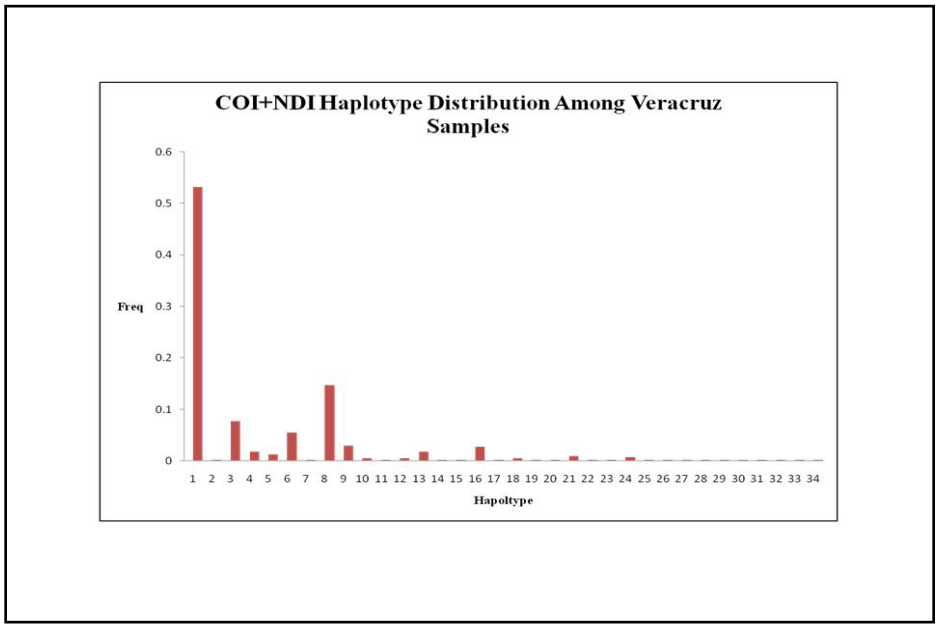


Figure 7. COI+NDI haplotype frequencies observed in *A. obliqua* specimens from Veracruz.

*Phylogenetic reconstruction with mtDNA*

Phylogenetic reconstructions with Maximum Likelihood (ML) (Fig. 8) and Maximum Parsimony (MP) methods provide concordant results showing similar topology, with moderate statistical support for containing single branch leading to haplotypes OBV17 & 20. These haplotypes were each recovered from single individual flies collected from *S. mombin* at Los Tuxtlas. Remaining branches were shallow and not supported according to bootstrap results from either method.

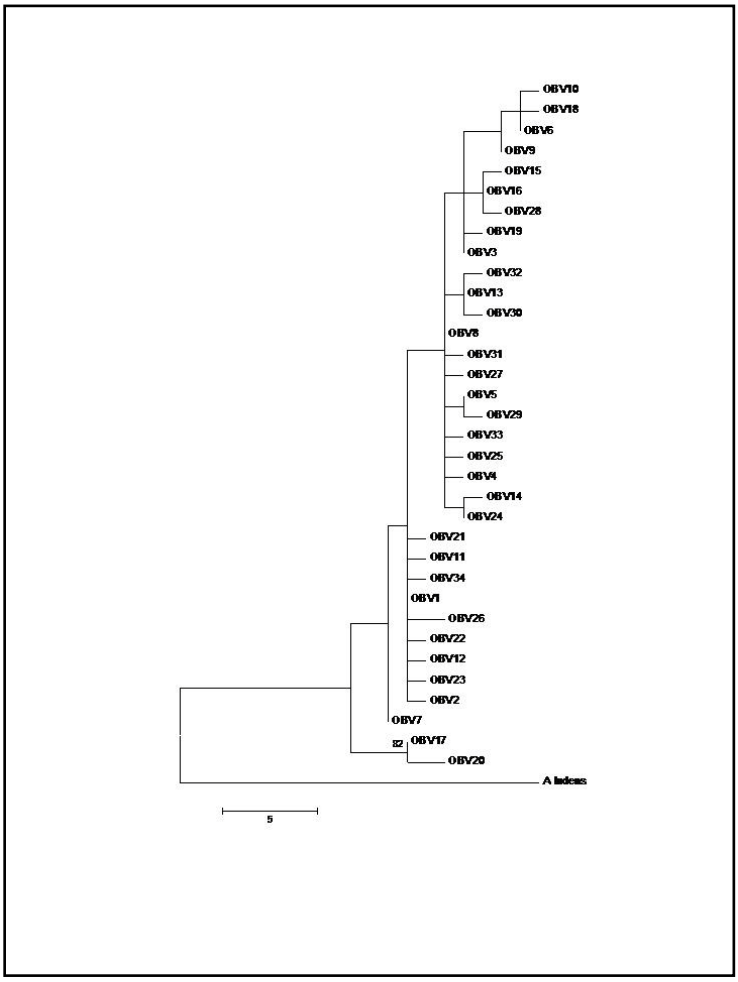


Figure 8. Phylogeny of 34 *Anastrepha obliqua* COI+ND1 haplotypes OBV01-OBV34, constructed using a Maximum Likelihood approach. Value above branch shows bootstrap score. Tree was rooted with *A. ludens*.

*Phylogeographic analysis of mtDNA*

We employed a nested clade phylogeographic analysis (NCPA) in order to examine whether haplotype distribution was significantly associated with geography among collections and to gain a more thorough understanding of complex biological processes and ecological trends influencing the distribution of *A. obliqua* in Veracruz.

Haplotypes separated by up to 14 mutational steps were connected following the hierarchical nesting procedure. NCPA partitioned the 34 haplotypes into two higher level clades (3-1 and 3-2 in Figs. 9 and 10). We have graphically superimposed geographic location (Fig. 9) and host (Fig. 10) onto the NCPA topologies to look for patterns from these two variables. Relatively little structure is obvious from this approach, although clade 3-2 contains primarily haplotypes found in Los Tuxtlas (Fig. 4) and haplotypes found in individuals reared from *S. mombin* (Fig. 10). There were five loops identified in the haplotype network. Both frequency and shared locality were used to resolve these ambiguities (Pfenninger and Posada 2002). For example, the first loop occurring among OBV01, OBV21, and OBV28 was resolved by breaking the connection between OBV21 and OBV28. This allowed OBV28 to remain connected to OBV01, the most frequent haplotype found within the same geographic location. We were able to use the same criteria for resolving the remaining four loops. Thirty four haplotypes and three nesting levels were observed in the cladogram. Twenty eight haplotypes were grouped in clade 3-1 and six in 3-2.

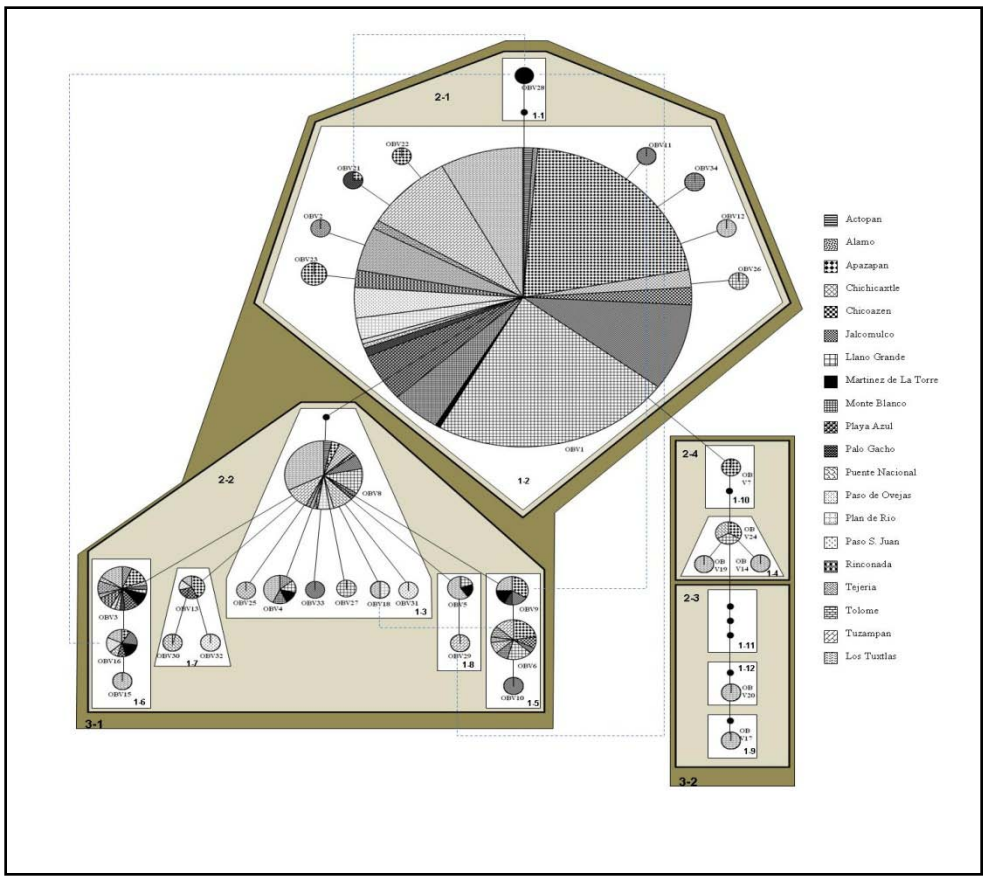


Figure 9. *Anastrepha obliqua* haplotype network with nested clades howing relationships among Veracruz, Mx. collections based on statistical parsimony. Size of the pie charts is proportional to haplotype frequency among populations (see Table 7). Each line between haplotypes indicates one mutational step. Black ovals without letter names are haplotypes not sampled but which are necessary to connect sampled haplotypes.

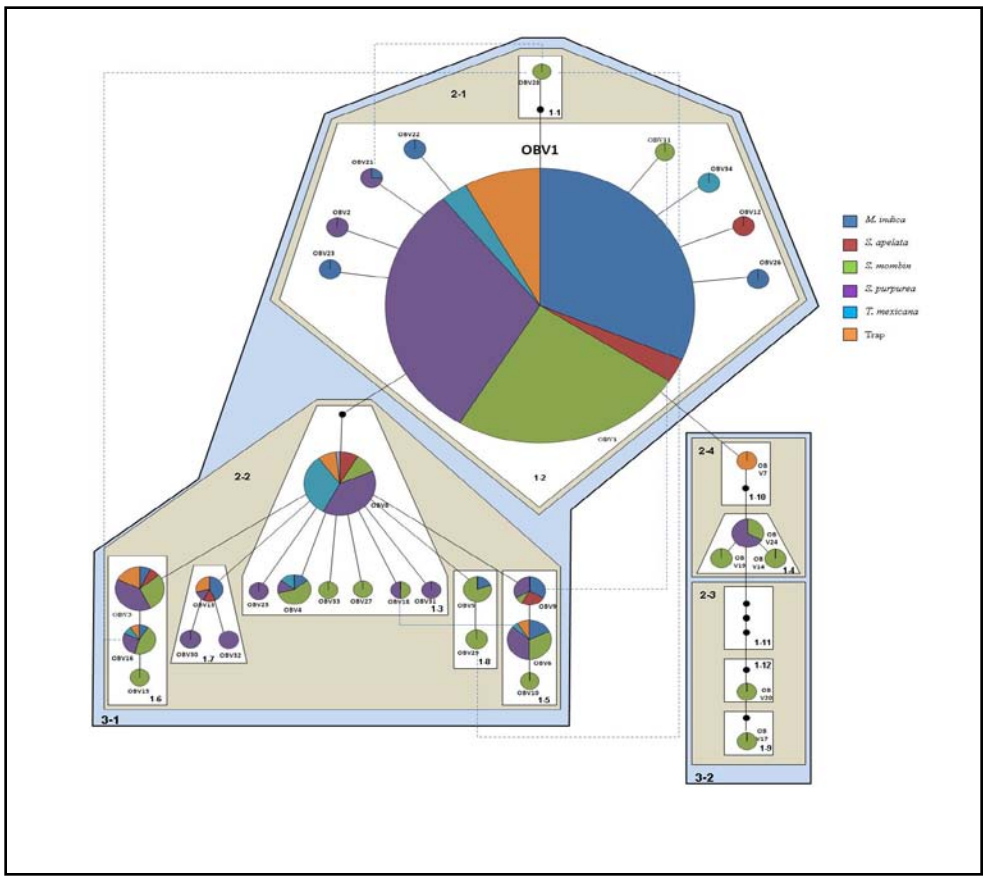


Figure 10. *Anastrepha obliqua* haplotype network with nested clades showing relationships among host collections based on statistical parsimony. Size of the pie charts is proportional to haplotype frequency among host collections (e.g. *M. indica*, see Table 7). Each line between haplotypes indicates one mutational step. Black ovals without letter names are haplotypes not sampled but which are necessary to connect sampled haplotypes.

Significant geographical association at different nesting levels was observed exclusively for haplotypes represented in clade 3-1. Contiguous range expansion was inferred in clade 1-2 and clade 2-2. Range expansion was inferred in clade 1-2 which contains the putative ancestral haplotype OBV01. Haplotype OBV01 was distributed among 48/52 collections and detected in flies gathered from all hosts examined. Within clade 1-2 we see that the eight additional and low frequency haplotypes were closely related to the common haplotype (OBV01) and arranged in a star-

like fashion found within this clade, characteristic for species that have undergone a bottleneck event and subsequent population expansion (Tajima 1989). We see a similar arrangement in clade 2-2 with OBV08 seen as the more common and ancestral haplotype. We investigated this hypothesis by testing for departure from neutrality via Tajima's *D* and Fu's $F_S$ demographic analyses and examining the mismatch distribution of these haplotypes. Tajima's *D* and Fu's $F_S$ values for this clade were -1.92 ($P < 0.001$) and -13.63 ($P < 0.000$), respectively, and differed significantly from the expectation under neutrality supporting NCPA results. A signature of population growth is also evident in the mismatch distribution for haplotypes within clades 1-2 and 2-2 (Fig. 11). A constant-sized population is projected to show multiple peaks while an expanding population shows a unimodal distribution (Slatkin and Hudson 1991) as seen for this clade. Sequences within this clade provided a mismatch distribution that is unimodal and cannot be distinguished from the distribution under a model of sudden population expansion (Slatkin and Hudson 1991, Rogers and Harpending 1992), further supporting NCPA and Tajima's *D* and Fu's $F_S$ demographic results.

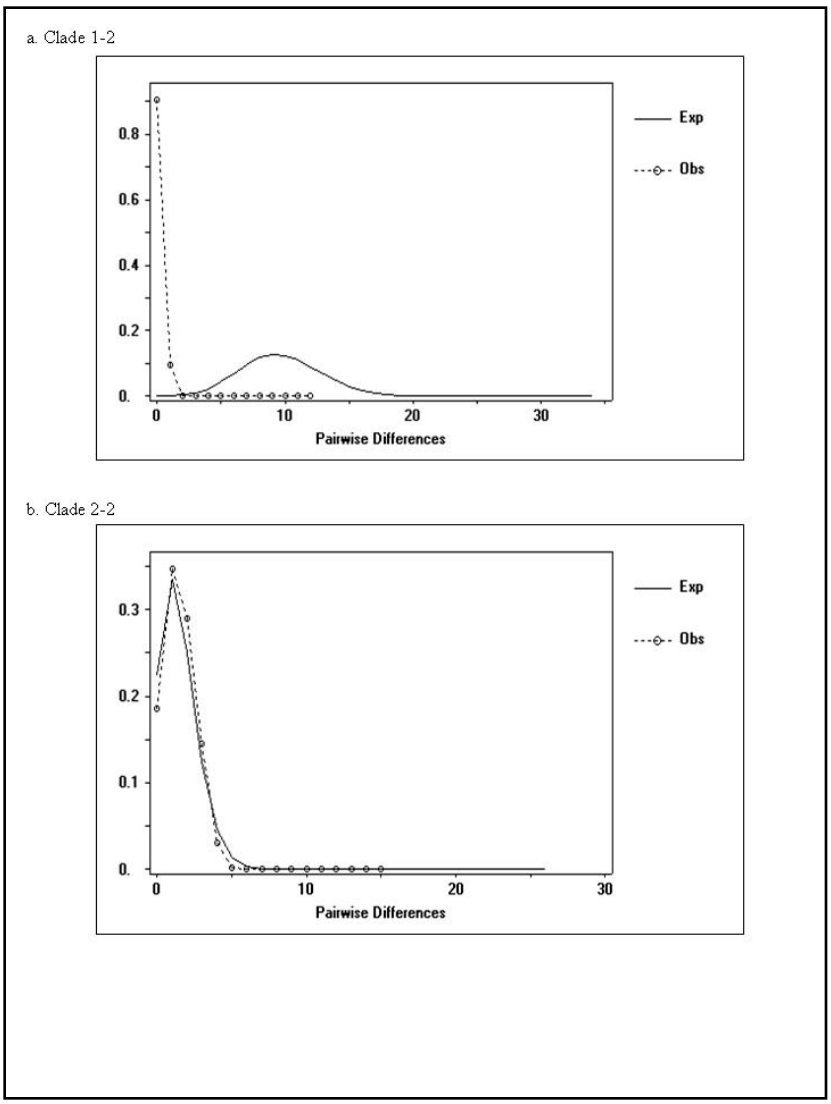


Figure 11. Results of mitochondrial mismatch distributions with data grouped according to nested clade analysis (NCPA) for a.) clade 1-2 and b.) clade 2-2. The expected frequency is based on a population growth-decline model determined using the DNASP v4.0 program (Rozas and Rozas, 1997) and is represented by a continuous line. The observed frequency is represented by a dotted line.

Hierarchical analysis of molecular variance (AMOVA) of populations in Veracruz revealed genetic structure at multiple levels (Table 12). A large percentage of the variation originated within populations. Tests examining structure among groups revealed that separating according to mango vs. all other hosts provided the greatest percent variation (6.92%) among groups.

Table 12. Partitioning of DNA variance at three hierarchical levels as revealed by analysis of molecular variance (AMOVA) for *A. obliqua* COI+NDI mitochondrial DNA sequences sampled in Veracruz, Mx.

| | Within individuals | | Among individuals within groups | | Among groups | |
|---|---|---|---|---|---|---|
| Group division | % var | *p* | % var | *p* | % var | *p* |
| 1. Host (*n* = 9) | 87.72 | <0.000 | 10.14 | <0.000 | 2.14 | 0.128 |
| 2. Host genus (*n* = 4) | 86.41 | <0.000 | 9.54 | <0.000 | 4.05 | 0.0279 |
| 3. Mango vs. others (*n* = 2) | 84.36 | <0.000 | 8.71 | <0.000 | 6.92 | 0.0017 |
| 4. Plum vs. others (*n* = 2) | 86.06 | <0.000 | 9.62 | <0.000 | 4.32 | 0.0096 |
| 5. Locale (*n* = 20) | 87.73 | <0.000 | 7.22 | <0.000 | 5.05 | 0.0312 |

*p*, estimated after 10,000 permutations

Pairwise comparisons among geographic collections provide significant $F_{ST}$ estimates across collections (Appendix B). Differentiation was also seen among individuals within locale. Among Apazapan collections, two collections, from *M. indica* and *S. purpurea* (Table 13), were significantly different. Mantel tests for Apazapan collections revealed no significant correlation between $F_{ST}$ estimates and geographic distance ($r$ = -0.077, $P$ = 0.517). (Accurate mapping coordinates were unavailable for other within-locale comparisons, so additional Mantel tests were not performed.)

Ten collections were obtained from Llano Grande, representing two host species - *S. mombin* and *M. indica* (Table 14). Only one of the six significant $F_{ST}$ values obtained through our analysis was from two collectionsof flies gathered from the same host species, the remaining five pairwise comparisons yielding significant estimates were host-related differences contrasting *S. mombin* vs. *M. indica*. The two collections gathered from *M. indica* and *T. mexicana* in Playa Azul were significantly different according to pairwise estimates (Table 15) as were two collections from *M. indica* in Tejeria (Table 16). At Los Tuxtlas, collection TUX4SM was significantly different from

all other collections sampled in this area. Additionally, one collection (TUX5TM) gathered from *T. mexicana* was significantly different from individuals collected from *S. apetala* (TUX2SA) (Table 17).

Table 13. Population pairwise *F*st values for six Apazapan, Veracruz collections. Population codes and site# in ( ) correspond to Table 7. Values in ***bold*** are significantly different from zero ($P<0.05$).

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1. AP1SP(3) | 0 | | | | | |
| 2. AP3SP(6) | -0.03274 | 0 | | | | |
| 3. AP3MI(7) | -0.036 | 0.00306 | 0 | | | |
| 4. AP4SP(8) | 0.39189 | 0.1875 | ***0.29609*** | 0 | | |
| 5. AP5MI(9) | 0.02032 | -0.05952 | 0.06383 | 0.07216 | 0 | |
| 6. AP6SP(10) | 0.0364 | -0.02091 | 0.02299 | 0.33512 | 0.05381 | 0 |

Table 14. Population pairwise *F*st values for ten Llano, Veracruz collections. Population codes and site# in ( ) correspond to Table 7. Values in ***bold*** are significantly different from zero (*P*<0.05).

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. LLA2MI(26) | 0 | | | | | | | | | |
| 2. LLA1ASM(17) | -0.14718 | 0 | | | | | | | | |
| 3. LLA1BSM(18) | -0.16667 | 0.14718 | 0 | | | | | | | |
| 4. LLA1CSM(19) | -0.16667 | 0.19215 | -0.16667 | 0 | | | | | | |
| 5. LLA1DSM(20) | ***0.67692*** | 0.16392 | ***0.67692*** | 0.7 | 0 | | | | | |
| 6. LLA1ESM(21) | 0.03012 | -0.12981 | 0.03012 | 0 | 0.01316 | 0 | | | | |
| 7. LLA1MI1(22) | -0.12197 | 0.08407 | -0.12197 | -0.0678 | ***0.61643*** | -0.00862 | 0 | | | |
| 8. LLA1MI2(23) | -0.10361 | 0.12629 | -0.10361 | -0.12958 | ***0.59373*** | 0.00798 | - | 0.08612 | 0 | |
| 9. LLA1MI3(24) | -0.04831 | -0.01869 | -0.04831 | -0.09833 | 0.24307 | -0.22203 | -0.05465 | -0.04408 | 0 | |
| 10. LLA1MI4(25) | 0.02041 | ***0.36389*** | 0.02041 | 0 | ***0.87532*** | 0.34247 | 0.10717 | 0.02916 | 0.09707 | 0 |

Table 15. Population pairwise *F*st values for two Playa Azul, Veracruz collections. Population codes and site# in ( ) correspond to Table 7.Values in ***bold*** are significantly different from zero (*P*<0.05).

| | 1 | 2 |
|---|---|---|
| 1. PAZMI(30) | 0 | |
| 2. PAZT(31) | ***0.58333*** | 0 |

Table 16. Population pairwise *F*st values for two Tejeria, Veracruz collections. Population codes and site# in ( ) correspond to Table 7. Values in ***bold*** are significantly different from zero ($P<0.05$).

| | 1 | 2 |
|---|---|---|
| 1. TE1MI1(21) | 0 | |
| 2. TE1MI2(22) | ***0.40459*** | 0 |

Table 17. Population pairwise *F*st values for seven Los Tuxtlas, Veracruz collections. Population codes and site# in ( ) correspond to Table 7. Values in ***bold*** are significantly different from zero ($P<0.05$).

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1. TUX1SM | 0 | | | | | | |
| 2. TUX2SM | -0.125 | 0 | | | | | |
| 3. TUX2SA | -0.01616 | -0.01616 | 0 | | | | |
| 4. TUX3SA | -0.04567 | -0.04567 | -0.021710 | | | | |
| 5. TUX4SM | ***0.14412*** | ***0.14412*** | ***0.10149*** | ***0.08713*** | 0 | | |
| 6. TUX5TM | 0.31646 | 0.31646 | ***0.16224*** | 0.13196 | -0.02042 | 0 | |
| 7. TUX6SM | 0.13847 | 0.13847 | 0.08275 | 0.06447 | -0.06105 | -0.01469 | 0 |

*Microsatellite analyses*

The analysis of 719 flies from 32 geographic collections gathered from five different hosts (Table 8) revealed 247 alleles from 17 loci (14.5 alleles/locus). Polymorphic Information Content (PIC, Botstein *et al*. 1980) values ranged from 0.04 to 0.821 (mean of 0.453) for these 247 alleles, with 41.2% considered highly informative (>0.5) and 47.1% characterized as reasonably informative (0.5>PIC>0.25) for population assignment (MacHugh *et al*. 1998). The observed heterozygosity ($H_O$) ranged from 0.042 to 0.923, with an average of 0.447. Of the 12223 reactions (719 individuals x 17 loci), 692 (5.7%) reactions failed to amplify. We found no evidence of linkage disequilibrium for any pair of loci in any population, and no departures from Hardy-Weinberg expectations for any population or locus.

The STRUCTURE analysis placed individuals into six clusters (Fig. 12) based on the ad hoc statistic Δ*K* (Evanno *et al*. 2005). Preliminary runs based on 10 replicates investigating *K*1-33 showed substantial decrease in probabilities after *K*12 (data not shown). Reducing the final comparisons to *K*1-12 revealed the optimal *K*. Assignment of individuals to the six clusters (Fig. 13) was based on 10 replicate runs and permutation analysis with CLUMPP.

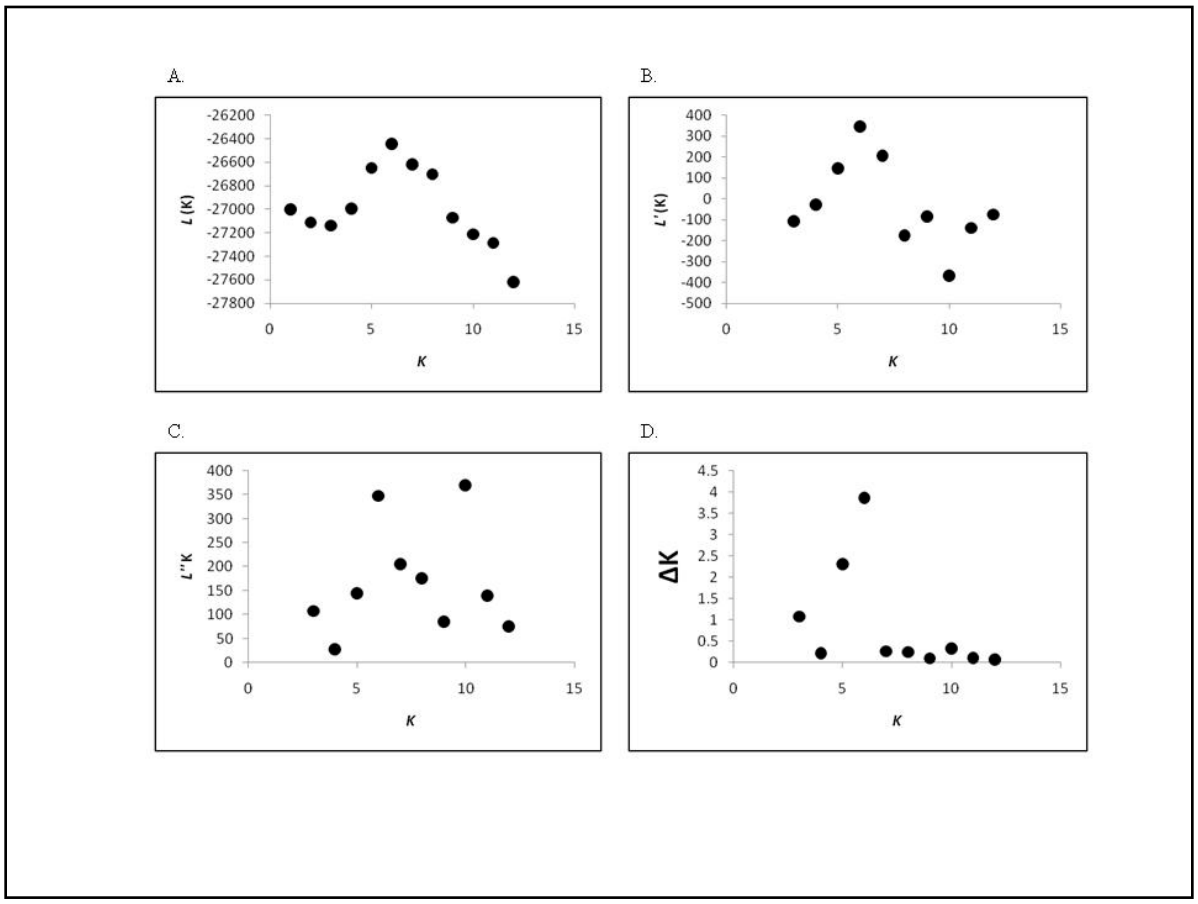


Figure 12. Charts showing the graphical method used in detection of the true number of groups K. (A) mean L(K) over 10 runs for each K value. (B) Rate of change of likelihood distribution, (C) absolute values of second order of rate of change of the likelihood distribution, and (D) ΔK were calculated according to Evanno *et al*. (2005) with the uppermost modal value being represented by six clusters.

The STRUCTURE defined clusters are seen at varying proportions across hosts (Table 18) and geography (Fig. 13 & Table 19). Cluster 1 was seen in all but *S. apetala* host collected material. It was present in 7.7% of *T. mexicana* and ranged to no more than 24.4% of the *S. mombin* samples. Cluster 1 was seen in higher proportions in Tolome (TO, 41.7%) and more moderately present in Llano (LLA, 24.6%), Martinez de la Torre (MAR, 35.7%), Monte Blanco (MB, 30.2%), and Palo Gacho (PG, 25.0%). Cluster 2 was seen in flies collected from all hosts with *S. mombin* having the highest percentage (18.4%) and in Martinez de la Torre (MAR, 28.6%). Cluster 3 was also present in all hosts with a somewhat even distribution in samples collected from *S. apetala* (20.8%), *S. purpurea* (21.3%), and *T. Mexicana* (23.1%) as well as an even distribution in various collection sites. Cluster 4 was present in greater proportions in specimens from *M. indica* (21.6%) and *S. apetala* (29.2%) and in Actopan (AC3, 54.2%) and Paso de Ovejas (PO, 61.5%). We see a higher percentage of cluster 5 in flies collected from *S. apetala* (29.2%) with moderate and somewhat even distribution among other hosts as well as sites with presence seen in no more than 40.0% in two sites, Chichicaxtle (CHI) and Paso San Juan (PSJ). Cluster 6 was also seen as having a moderate distribution among various hosts with a higher percentage from *M. indica* (22.3%) specimens. This cluster had a higher presence in Tejeria (TE, 52.4%) and Rinconada (RI, 44.4%) with low to moderate distribution in all other sites.

Table 18. Genetic clusters as percentage of *A. obliqua* samples from each host. *M. indica* (Mi), *S. apetala* (Sa), *S. mombin* (Sp), *S. purpurea* (Sp), and *T. Mexicana* (Tm).

| | Mi | Sa | Sm | Sp | Tm |
|---|---|---|---|---|---|
| *A. obliqua* (*n*) | 139 | 24 | 266 | 277 | 13 |
| Clusters | Percent of host samples of *A. obliqua* | | | | |
| C1 | 15.1 | - | 24.4 | 18.4 | 7.7 |
| C2 | 10.1 | 4.2 | 18.4 | 11.9 | 15.4 |
| C3 | 12.2 | 20.8 | 15.4 | 21.3 | 23.1 |
| C4 | 21.6 | 29.2 | 10.5 | 15.9 | 15.4 |
| C5 | 18.7 | 29.2 | 18.0 | 17.0 | 23.1 |
| C6 | 22.3 | 16.7 | 13.2 | 15.5 | 15.4 |

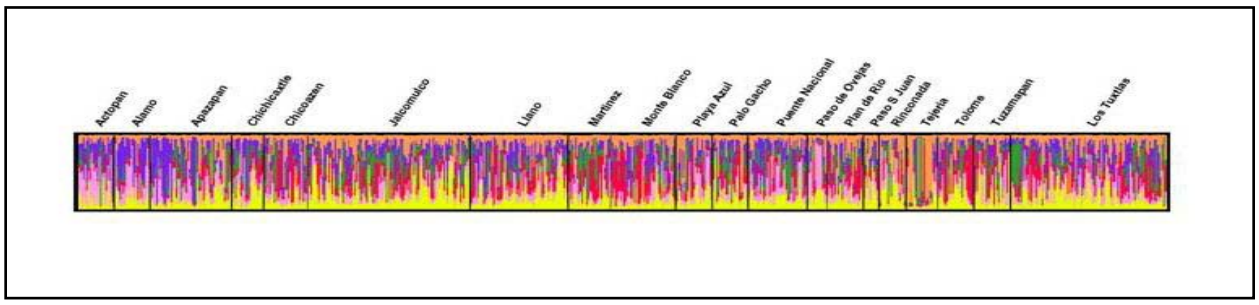


Figure 13. Clustering outcomes for all samples (K = 6) from the twenty geographic sites. Each population is separated from its neighbor by a solid line.

We compared the model of genetic clusters as revealed by STRUCTURE to various alternative a priori hypotheses of population structure. The models were tested using the hierarchical analysis of molecular variance (AMOVA). Five comparisons were conducted and included (1) six major clusters defined through assignment tests by STRUCTURE, (2) host populations, 3) geographic locale, (4) regional areas determined as north (Alamo), south (Los

Tuxtlas), and central (all others), and (5) month when infested host fruit was collected (Table 20). AMOVA partitioned the total variance into covariance components. The evaluation of hypotheses as mentioned above was made by examining the among group results. The division providing the highest percent variation among groups was according to clusters. This percent variation was substantially higher than other comparisons.

Table 19. Genetic clusters as percentage of *A. obliqua* samples from each collection locality. Locality codes are according to Table 8.

| | AC3 | AL01 | AP | CHI | CHI1 | JA | LLA | MAR | MB | PAZ | PG | PN | PO | PR | PSJ | RI | TE | TO | TU | TUX |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *A. obliqua* (*n*) | 24 | 24 | 54 | 20 | 30 | 107 | 65 | 28 | 43 | 24 | 24 | 39 | 13 | 24 | 10 | 18 | 21 | 24 | 24 | 103 |
| Clusters | Percent of site samples of *A. obliqua* | | | | | | | | | | | | | | | | | | | |
| C1 | - | 8.3 | 5.6 | 5.0 | 28.1 | 18.5 | 24.6 | 35.7 | 30.2 | 16.7 | 25.0 | 17.9 | 7.7 | 29.2 | 10.0 | 11.1 | 4.8 | 41.7 | 20.8 | 18.4 |
| C2 | 8.3 | 4.2 | 16.7 | 15.0 | 12.5 | 15.7 | 9.2 | 28.6 | 7.0 | 4.2 | 20.8 | 20.5 | - | 12.5 | 10.0 | 5.6 | 14.3 | 12.5 | 12.5 | 17.5 |
| C3 | 25.0 | 37.5 | 31.5 | 25.0 | 21.9 | 13.0 | 20.0 | 3.6 | 25.6 | 16.7 | 20.8 | 17.9 | 7.7 | 8.3 | 10.0 | - | 9.5 | 8.3 | 4.2 | 16.5 |
| C4 | 54.2 | 41.7 | 18.5 | 10.0 | 9.4 | 8.3 | 10.8 | 10.7 | 9.3 | 20.8 | 4.2 | 28.2 | 61.5 | 4.2 | 10.0 | 11.1 | 9.5 | - | 20.8 | 13.6 |
| C5 | 12.5 | 4.2 | 9.3 | 40.0 | 9.4 | 30.6 | 21.5 | 10.7 | 11.6 | 12.5 | 16.7 | 10.3 | - | 12.5 | 40.0 | 27.8 | 9.5 | 29.2 | 33.3 | 17.5 |
| C6 | - | 4.2 | 18.5 | 5.0 | 18.8 | 13.9 | 13.8 | 10.7 | 16.3 | 29.2 | 12.5 | 5.1 | 23.1 | 33.3 | 20.0 | 44.4 | 52.4 | 8.3 | 8.3 | 16.5 |

Table 20. Partitioning of DNA variance at three hierarchical levels as revealed by analysis of molecular variance (AMOVA) for *A. obliqua* microsatellite DNA sampled in Veracruz, Mx.

| | Within individuals | | Among individuals within groups | | Among groups | |
|---|---|---|---|---|---|---|
| Group division | % var | *p* | % var | *p* | % var | *p* |
| 1. Clusters (*n* = 6) | 95.15 | <0.000 | 1.57 | 0.014 | 3.28 | <0.000 |
| 2. Host (*n* = 5) | 97.44 | <0.000 | 2.45 | <0.000 | 0.10 | 0.205 |
| 3. Locale (*n* = 20) | 97.45 | <0.000 | 2.09 | <0.000 | 0.44 | 0.084 |
| 4. Regional (*n* = 3) | 97.31 | <0.000 | 2.44 | <0.000 | 0.24 | 0.180 |
| 5. Temporal (*n* = 8) | 97.41 | <0.000 | 2.18 | <0.000 | 0.41 | 0.011 |

*p*, estimated after 10,000 permutations

## DISCUSSION

In this study we sampled West Indian fruit fly populations from Veracruz, Mexico in order to investigate the influence of ecological factors on genetic diversity. Sampling was carried out here over a three year period (1993-1995) from a geographic area at locations where multiple host species are co-located. We examine individual collections for diversity and possible source of variation. We test these data in order to determine whether the existing populations sampled are structured as one regional population and individuals migrate from the different host species over time or there are various structured populations existing that pattern their reproductive cycles according to fruit phenology or are structured through other factors. In the course of uncovering biological processes we also investigate historical processes with demographic and nested clade methods using mtDNA sequences and assignment tests and demographic analyses with microsatellite DNA methods.

### *Diversity among collections*

Our results show that *A. obliqua* populations sampled are highly similar yet show low to moderate population structure. Phylogenetic analyses of mtDNA sequences reveal no support for all but one branch also suggesting a high degree of homogeneity among collections. Our sequencing results of mtDNA shows a high frequency (0.53) of the individuals exhibiting a single haplotype (OBV01). Haplotype (0.6846) diversity was moderate and nucleotide (0.0017) diversity was low among the fifty two geographic collections analyzed. This is generally typical of an expanding population or founder effect (von der Heyden *et al*. 2007, Zhang *et al*. 2006) which is associated with little or no population structure (Patarnello *et al*. 2007, and von der Heyden *et al*. 2007). For microsatellites, the mean number of alleles per locus was high (14.5) and PIC (polymorphic information content) mean value (0.447) and number of markers was considered highly informative for the type of

analyses we conduct in this study (Mateescu *et al*. 2005, MacHugh *et al*. 1998).

*Population structure with mtDNA*

We observed population structure (based on $F_{ST}$) from analyses of mtDNA sequences. While the $F_{ST}$ estimates were higher for mtDNA than for microsatellites this is most likely due to the differences in effective population (*N*e) represented by these DNA regions. Nevertheless, these estimates fell within a range that supports the notion that differentiation reflects a minimally restricted geneflow. There were significant differences among pairwise comparisons with sequencing data of samples collected within a locale. Differences were significant among samples collected from different hosts in Apazapan. Mantel tests showed there was no correlation between $F_{ST}$ and distance and thus we cannot rule out that differentiation was due to host association. Analyses showed that there was significant differentiation among collections within other locales as well. In Llano Grande, five of the six pairwise comparisons that yielded significance were of samples from *M. indica* vs. *S. mombin*. Similar results were seen in Playa Azul and Los Tuxtlas with sequencing data. Comparisons of samples collected in Llano Grande, Playa Azul, and Los Tuxtlas suggest that differentiation may have been attributed to host association, however, we were unable to confirm whether differences may have been due to isolation by distance since accurate mapping coordinates were unattainable.

Demographic analyses of sequencing (NCPA/demographic analyses) data suggest that populations have undergone a recent bottleneck event and expansion. The study by Ruiz-Arce *et al*. (in preparation) using sequences of COI/ND6 regions also suggests this species may have experienced a recent bottleneck event with subsequent expansion in this region and may have origins elsewhere and thus have been recently introduced into this and other areas. Introduced individuals

from a more common population cause a bottleneck thereby reducing the genetic diversity in introduced populations (Hartl and Clark 1997). This may help explain the low nucleotide and haplotype diversity estimates. Additionally, AMOVA analyses revealed that no less than 84.36% (mtDNA) of the variation originate from individuals and not groups and only a small percentage (6.92%) was due to comparisons of when data was grouped by *M. indica* vs. all other hosts indicating homogeneity among collections.

*Population structure with microsatellites*

The appropriate choice of the number of subpopulations was straightforward for this data set as optimal likelihood modes were observed at K6. Assignment tests revealed that population differentiation in the West Indian fruit at microsatellite loci is sufficient to cluster individuals however we were unable to associate variation to a particular factor. Our analyses indicate that A. obliqua populations, while not entirely panmictic are certainly highly homogenous within Veracruz. This result is not surprising given the geographical proximity of the subpopulations. Additionally, AMOVA estimates for these five groups also suggest a high degree of homogeneity where low among-group variation was seen. These subpopulations are separated by short distances with low genetic structure ($F_{ST}$ = 0.016–0.139, Table 21) observed among the six groups with microsatellite methods. These results suggest that habitat in Veracruz is much less fragmented and geneflow among collections within the area sampled is high.

Table 21. Estimated of *F* statistics and number of migrants per generation with mean and standard errors for the six groups designated by STRUCTURE for each locus.

| | 3B | A1 | A3 | A7 | A20 | A27 | A28 | Anob 01 | Anob 02 | Anob 03 | Anob 04 | Anob 06 | Anob 12 | Anob 15 | Anob 17 | Anob 18 | Anob 08 | Mean+SE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *F*is | 0.172 | 0.181 | 0.037 | -0.034 | 0.139 | 0.117 | 0.084 | 0.074 | 0.085 | 0.260 | 0.435 | 0.278 | 0.290 | 0.144 | 0.033 | -0.013 | 0.523 | 0.165+0.037 |
| *F*it | 0.211 | 0.209 | 0.056 | 0.003 | 0.175 | 0.138 | 0.102 | 0.105 | 0.132 | 0.284 | 0.450 | 0.379 | 0.319 | 0.178 | 0.048 | 0.018 | 0.583 | 0.199+0.039 |
| *F*st | 0.048 | 0.035 | 0.020 | 0.036 | 0.043 | 0.024 | 0.020 | 0.033 | 0.051 | 0.032 | 0.027 | 0.139 | 0.041 | 0.039 | 0.016 | 0.030 | 0.126 | 0.045+0.008 |
| *N*m | 5.00 | 6.92 | 12.30 | 6.69 | 5.61 | 10.38 | 11.99 | 7.26 | 4.61 | 7.45 | 8.86 | 1.54 | 5.91 | 6.08 | 15.65 | 8.00 | 1.73 | 7.41+0.88 |

From our findings, it appears that the limited genetic diversity revealed by mtDNA may be due to the recent introduction of individuals to this area. Our study revealed a haplotype structure with limited variation where one predominant haplotype occurs throughout the collections analyzed. While pairwise comparisons provide some resolution to geographic structure, sequence variation among haplotypes is minimal with low nucleotide differences among haplotypes suggesting a somewhat unrestricted geneflow among collections. Given the data we collected, there are two possible introduction scenarios. First, introductions into Veracruz may have originated from a small portion of a more diverse population. Second, it is also possible that multiple introductions have occurred from a more diverse portion of the ancestral range and individuals representing unfavorable genetic types were selected against. We are unable to distinguish among these two plausible scenarios with our data at this time. While the diversity may be low with mtDNA methods for collections examined for this study, we still see hints of structure with both mtDNA and microsatellite methods. Diversity estimates were high while structure was considered moderate according to microsatellite analyses. Microsatellite methods reveal limited structuring as well and therefore agree with mtDNA methods.

Various other factors may be contributing to the lack of heterogeneity. Adding to the above biological scenario is human-mediated distribution of individuals contained within infested fruit material from limited gene pools. The amount of individuals distributed through human mediated transportation is unknown but may be substantial given Veracruz is the third most populated state in Mexico (7.1 million as of 2005). Wind assistance and natural flight while at times minimal (Hernandez *et al*. 2007, Baker *et al*. 1986) also contribute to the dispersal of individuals from established populations and may provide a homogenizing effect for this and other congeneric tephritid pest species. Additionally, the subsequent measures for managing this pest practiced by agricultural communities inherently reduce population size and genetic diversity. While it is difficult to quantify the direct influence to the diversity of this particular species with these data, numerous studies have shown that declines in biodiversity coincide with agriculture intensification (Grixti *et al*. 2009, van der Wal *et al*. 2008, Hatfield and LeBuhn 2007).

This study demonstrates the utility of a multi-locus approach for examining and quantifying genetic diversity and geneflow among subpopulations of *A. obliqua*. Here we provide substantial evidence confirming the presence of a relatively homogenous group of populations sampled from Veracruz. This conclusion is especially important in view of the management strategies of *A. obliqua* within this region whereby treatments of less diverse group of pests may be more effective and perhaps more simplified than of a genetically diverse group of individuals. Additionally, the detection of gene flow levels among fly populations in Veracruz adds to a growing body of literature that documents an increasingly fragmented landscape for pest fruit flies in the Americas. This may be of particular importance to action agencies and biodiversity studies. Our data on mtDNA and microsatellites provide several useful perspectives for further understanding the biology and management of *A. obliqua*.

**ACKNOWLEDGEMENTS**

We thank Nicole Botteri, Clinton Cario, Rachelle L. Christie, Eric Deutsche, Samuel Gum, Christopher L. Owen, Sou Young “Sonya” Park, Oralia Ruiz, Hsiao-Chu “Anya” Tu, for technical assistance at The Pennsylvania State University, Deborah S. Grove (Director of the Genomics Core Facility), Ashley Price, and Sheila Plock at The Pennsylvania State University-Huck Institute Nucleic Acid Facility. Funding for this study came from the USDA-APHIS-CPHST, PIC#M6M02 and USDA/Penn State cooperative agreement 07-8100-1051-CA as well as support from the Alfred P. Sloan Foundation. We are grateful to the Mexican Campaña Nacional Contra Moscas de la Fruta (SAGARPA-IICA ), the Consejo Nacional de Ciencia y Tecnología (CONACYT), and the Instituto de Ecología, A.C., Xalapa, Mexico for their financial support that made these collections possible. We are also grateful to Isabel Jácome for the valuable contributions to this study. Part of this work was carried out by using the resources of the Computational Biology Service Unit from Cornell University which is partially funded by Microsoft Corporation and High Performance Computing at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) which belongs to the Alliance of Centers supported by the Consultative Group on International Agricultural Research (CGIAR).

## REFERENCES


Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410.

Aluja M, Celedonio-Hurtado H, Liedo P, Cabrera M, Castillo F, Guill J, Rios E (1996) Seasonal population fluctuations and ecological implications for management of *Anastrepha* fruit flies (Diptera: Tephritidae) in commercial mango orchards in southern Mexico. *Journal of Economic Entomology* **89**: 654-667.

Aluja M (1994) Bionomics and management of *Anastrepha*. *Annual Review of Entomology* **39**: 155-178.

Aluja M, Birke A (1993) Habitat use by adults of *Anastrepha obliqua* (diptera, tephritidae) in a mixed mango and tropical plum orchard. *Annals of the Entomological Society of America* **86**: 799-812.

Aluja M, Diaz-Fleischer F (2006) Foraging behavior of *Anastrepha ludens*, *A obliqua*, and *A serpentina* in response to feces extracts containing host marking pheromone. *Journal of Chemical Ecology* **32**: 367-389.

Baker PS, Chan AST, Zavala MAJ (1986) Dispersal and orientation of sterile *Ceratitis capitata* and *Anastrepha ludens* (tephritidae) in chiapas, mexico. *Journal of Applied Ecology* **23**: 27-38.

Barr NB (2009) Pathway Analysis of Ceratitis capitata (Diptera: Tephritidae) Using Mitochondrial DNA. *Journal of Economic Entomology* **102**: 401-411.

Barr NB, Copeland RS, De Meyer M, Masiga D, Kibogo HG, Billah MK, Osir E, Wharton RA, McPheron BA (2006). Molecular diagnostics of economically important *Ceratitis* fruit fly species (Diptera : Tephritidae) in Africa using PCR and RFLP analyses. *Bulletin of Entomological Research* **96:** 505-521.

Beaumont MA, Panchal M (2008) On the validity of nested clade phylogeographical analysis. *Molecular Ecology* **17**: 2563-2565.

Berlocher SH, Feder JL (2002) Sympatric speciation in phytophagous insects: Moving beyond controversy? *Annual Review of Entomology* **47**: 773-815.

Bitarello BD, Torres TT, Lyra ML, De Azeredo-Espin AML (2009) Development of polymorphic microsatellite markers for the human botfly, *Dermatobia hominis* (Diptera: Oestridae). *Molecular Ecology Resources* **9**: 409-411.

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**: 314– 331.

Bush GL, Feder JL, Berlocher SH, McPheron BA, Smith DC, Chilcote CA (1989) Sympatric origins of *R. pomonella*. *Nature* **339**: 346-346.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657-1659.

Cook LG, Rowell DM (2007) Genetic diversity, host-specificity and unusual phylogeography of a cryptic, host-associated species complex of gall-inducing scale insects. *Ecological Entomology* **32**: 506-515.

Diegisser T, Seitz A, Johannesen J (2006) Phylogeographic patterns of host-race evolution in *Tephritis conura* (Diptera : Tephritidae). *Molecular Ecology* **15**: 681-694.

Epsky ND, Kendra PE, Heath RR (2003) Development of lures for detection and delimitation of invasive *Anastrepha* fruit flies, pp. 84-89 *In* W. Klassen, W. Colon, and W. I. Lugo [eds.], *Proc. of the 39th Annual Meeting of the Caribbean Food Crops Society*, July 2003, Grenada.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611-2620.

Excoffier L, Estoup A, Cornuet JM (2005) Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics* **169**: 1727-1738.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among dna haplotypes - application to human mitochondrial-dna restriction data. *Genetics* **131**: 479-491.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.

Falush D, Wirth T, Linz B, Pritchard JK, Stephens M, Kidd M, Blaser MJ, Graham DY, Vacher S, Perez-Perez GI, Yamaoka Y, Mégraud F, Otto K, Reichard U, Katzowitsch E, Wang X, Achtman M, Suerbaum S (2003a) Traces of human migrations in *Helicobacter pylori* populations. *Science*, **299**: 1582–1585.

Felsenstein J (1985) Phylogenies from gene-frequencies - a statistical problem. *Systematic Zoology* **34**: 300-311.

Foote RH, Blanc FL, Norrbom AL (1993) ‘‘Handbook of the Fruit Flies (Diptera: Tephritidae) of America North of Mexico,’’ Comstock, Ithaca, NY.

Forbes AA, Fisher J, Feder JL (2005) Habitat avoidance: Overlooking an important aspect of host-specific mating and sympatric speciation? *Evolution* **59**: 1552-1559.

Fraga ED, dos Santos JM, Maia JD (2003) Enzymatic variability in *Aedes aegypti* (Diptera : Culicidae) populations from Manaus-AM, Brazil. *Genetics and Molecular Biology* **26**: 181-187.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915-925.

Garcia E (1970) Las climas del estado Veracruz. Anales del Instituto de Biologia, **41**:3-22.

Gasparich GE, Silva JG, Han H-Y, McPheron BA, Steck GJ, Sheppard WS (1997) Population genetic structure of Mediterranean fruit fly (Diptera : Tephritidae) and implications for worldwide colonization patterns. *Annals of the Entomological Society of America* **90**: 790-797.

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. Journal of *Heredity* **86**: 485-486.

Grixti JC, Wong LT, Cameron SA, Favret C (2009) Decline of bumble bees (*Bombus*) in the North American Midwest. Biological Conservation **142**: 75-84.

Guindon S, Lethiec F, DurouxP, Gascuel O (2005) PHYML Online - a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Research* **33**: 557-559.

Hartl DL, Clark AG (1997) Principles of population genetics, 3rd ed. Sinauer Associates, Inc., Sunderland.

Hatfield RG, LeBuhn G (2007) Patch and landscape factors shape community assemblage of bumble bees, *Bombus* spp. (Hymenoptera: Apidae), in montane meadows. *Biological Conservation* **139**: 150-158.

Henning F, Matioli SR (2006) Mating time of the west Indian fruit fly *Anastrepha obliqua* (Macquart) (Diptera : tephritidae) under laboratory conditions. *Neotropical Entomology* **35**: 145-148.

Heppner JB (1991) Larvae of fruit flies 7. *Anastrepha obliqua* (West Indian fruit fly) (Diptera: Tephritidae). Florida Department of Agriculture and Consumer Services, Division of Plant Industry Entomology Circular **339**: 1-2.

Hernandez E, Orozco D, Breceda SF, Dominguez J (2007) Dispersal and longevity of wild and mass-reared *Anastrepha ludens* and *Anastrepha obliqua* (Diptera : Tephritidae). *Florida Entomologist* **90**: 123-135.

Hernandez-Ortiz V, Aluja M (1993) Lista preliminar de species del genero Neotropical *Anastrepha* (Diptera: Tephritidae) con notas sobre su distribucion y plantas hospederas. *Folia Entomol. Mex.* (Mexico) **88**: 89–105.

Hu JJ, Zhang L, Nardi F, Zhang RJ (2008) Population genetic structure of the melon fly, *Bactrocera cucurbitae* (Diptera : Tephritidae), from China and Southeast Asia. *Genetica* **134**: 319-324.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801-1806.

Kimura M (1983) Rare variant alleles in the light of the neutral theory. *Molecular Biology and Evolution* **1**: 84-93.

Knio KM, White IM, Al-Zein MS (2007) Host-race formation in *Chaetostomella cylindrica* (Diptera : Tephritidae): Morphological and morphometric evidence. *Journal of Natural History* **41**: 1697-1715.

Knowles LL (2008) Why does a method that fails continue to be used? *Evolution* **62**: 2713-2717.

MacHugh DE, Loftus RT, Cunningham P, Bradley DG (1998) Genetic structure of seven European cattle breeds assessed using 20 microsatellite markers. *Animal Genetics* **29**: 333–340.

Malavasi A, Morgante JS (1981) Biology of fruit-flies. adult and larval population fluctuation of *Anastrepha fraterculus* (diptera, tephritidae) and its relationship to host availability. *Environmental Entomology* **10**: 275-278.

Mantel N (1967) Detection of disease clustering and a generalized regression approach. *Cancer Research* **27**: 209.

Mateescu B, England P, Halgand F, Yaniv M, Muchardt C (2004) Tethering of HP1 proteins to chromatin is relieved by phosphoacetylation of histone H3. EMBO Rep **5**: 490–496.

McPheron BA, Smith DC, Berlocher SH (1988) Microgeographic genetic-variation in the apple maggot *Rhagoletis pomonella*. *Genetics* **119**: 445-451.

Michel AP, Rull J, Aluja M, Feder JL (2007) The genetic structure of hawthorn-infesting Rhagoletis pomonella populations in Mexico: implications for sympatric host race formation. *Molecular Ecology* **16**: 2867-2878.

Mun JH, Bohonak AJ, Roderick GK (2003) Population structure of the pumpkin fruit fly *Bactrocera depressa* (Tephritidae) in Korea and Japan: Pliocene allopatry or recent invasion? *Molecular Ecology* **12**: 2941–2951.

Norrbom AL (2004) Host plant database for *Anastrepha* and *Toxotrypana* (Diptera: Tephritidae: Toxotrypanini). Diptera Data Dissemination Disk (CDROM) http://www.sel.barc.usda.gov/diptera/Tephritidae/TephHosts/search.html.

Norrbom AL, Kim KC (1988) Revision of the Schausi group of a*nastrepha* schiner (Diptera, Tephritidae), with a discussion of the terminology of the female terminalia in the Tephritoidea. *Annals of the Entomological Society of America* **81**: 164-173.

Novotny V, Clarke AR, Drew RAI, Balagawi S, Clifford B (2005) Host specialization and species richness of fruit flies (Diptera : Tephritidae) in a New Guinea rain forest. *Journal of Tropical Ecology* **21**: 67-77.

Panchal M, Beaumont MA (2007) The automation and evaluation of nested clade phylogeographic analysis. *Evolution* **61**: 1466-1480.

Park SDE, Adomefa K, Dao B, Hanotte O, Kemp SJ, Sow R, Teale AJ, Bradley DG (1999) Application of population genetic analysis of linked, mapped microsatellites to the identification of loci under selection for disease resistance. *Archiv Fur Tierzucht-Archives of Animal Breeding* **42**: 97-99.

Patarnello T, Volckaert FAMJ, Castilho R (2007) Is the Atlantic-Mediterranean transition a phylogeographical break? *Molecular Ecology*, **16**:4426-4444.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**: 288-295.

Petit RJ (2008) The coup de grace for the nested clade phylogeographic analysis? *Molecular Ecology* **17**: 516-518.

Pfenninger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): Fragmentation, corridor migration, and secondary contact. *Evolution* **56**: 1776-1788.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**: 817-818.

Posada D, Crandall KA, Templeton AR (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology* **9**: 487-488.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.

Ravel S, Herve JP, Diarrassouba S, Kone A, Cuny G (2002) Microsatellite markers for population genetic studies in *Aedes aegypti* (Diptera : Culicidae) from Cote d'Ivoire: evidence for a microgeographic genetic differentiation of mosquitoes from Bouake. *Acta Tropica* **82**: 39-49.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**: 552-569.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496-2497.

Sanders WT (1953) The Anthropogeography of Central Veracruz. In *Huastecos, totonacos, y sus vecinos, Revista Mexicana de Estudios Antropologicos,* vol. 13, edited by I. Berna and E. Davalos-Hurtado, 27-78.

Schwarz D, Matta BM, Shakir-Botteri NL, McPheron BA (2005) Host shift to an invasive plant triggers rapid animal hybrid speciation. Nature **436**: 546-549.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**: 555-562.

Shelly TE, Villalobos EM (2004) Host plant influence on the mating success of male Mediterranean fruit flies: variable effects within and between individual plants. *Animal Behaviour* **68**: 417-426.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, 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**: 651–701.

Smith PT, Kambhampati S, Voelkl W, Mackauer M (1999) A phylogeny of aphid parasitoids (Hymenoptera: Braconidae: Aphidiinae) inferred from mitochondrial NADH 1 dehydrogenase gene sequence. *Molecular Phylogenetics and Evolution.* **11**: 236-245.

Steck GJ (2001) Concerning the occurrence of *Anastrepha obliqua* (Diptera : Tephritidae,) in Florida. *Florida Entomologist* **84**: 320-321.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585-595.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**: 1596-1599.

Tan KH, Serit M (1994) Adult population and dynamics of *Bactrocera dorsalis* (Diptera, Tephritidae) In relation to host phenology and weather in 2 villages of Penang Island, Malaysia. *Environmental Entomology* **23**: 267-275.

Teixeira LAF, Polavarapu S (2003) Evolution of phenologically distinct populations of *Rhagoletis mendax* (Diptera : Tephritidae) in highbush blueberry fields. *Annals of the Entomological Society of America* **96**: 818-827.

Templeton AR (2009) Why does a method that fails continue to be used? The answer. *Evolution* **63**: 807-812.

Templeton AR (2008) Nested clade analysis: an extensively validated method for strong phylogeographic inference. *Molecular Ecology* **17**: 1877-1880.

Templeton AR (1998) The complexity of the genotype-phenotype relationship and the limitations of using genetic "markers" at the individual level, pp. 373-389.

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

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619-633.

Thompson JD, Higgins DG, Gibson T (1994) Clustal W - improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680.

Tsutsui ND, Suarez AV, Holway DA, Case TJ (2001) Relationship among native and introduced populations of the Argentine ant (*Linepithema humile*) and the source of introduced populations. *Molecular Ecology*. **10**: 2151–2161.

van der Wal R, Truscott AM, Pearce ISK, Cole L, Harris MP, Wanless S (2008) Multiple anthropogenic changes cause biodiversity loss through plant invasion. *Global Change Biology* **14**: 1428-1436.

von der Heyden S, Lipinski MR, Matthee CA (2007) Mitochondrial DNA analyses of the Cape hakes reveal an expanding population for *Merluccius capensis* and population structuring for mature fish in *Merluccius paradoxus*. *Molecular Phylogenetics and Evolution*, **42**:517-527.

Walker MS, Hartley E, Jones TH (2008) The relative importance of resources and natural enemies in determining herbivore abundance: thistles, tephritids and parasitoids. *Journal of Animal Ecology* **77**: 1063-1071.

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**:438-450.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358-1370.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics* **15**: 323-354.

Zhang J, Cai z, Huang L (2006) Population genetic structure of crimson snapper *Lutjanus erythropterus* in East Asia, revealed by analysis of the mitochondrial control region. *Journal of Marine Science*, **63**:693-704.

## Chapter 4: Phylogeography of West Indian fruit fly, *Anastrepha obliqua*

Raul Ruiz, Bruce A. McPheron, Christopher L. Owen, and Donald B. Thomas

## ABSTRACT

*Anastrepha obliqua* (Macquart) (Diptera: Tephritidae), the West Indian fruit fly, is a frugivorous pest that occasionally finds its way to commercial growing areas outside its native distribution. It inhabits areas in Mexico, Central and South America, and the Caribbean with occasional infestations having occurred in the southern tier states (CA, FL & TX) of the United States. This fly is associated with many plant species and is a major pest of mango and plum. We examine the genetic diversity of the West Indian fruit fly based on mitochondrial COI and ND6 DNA sequences. We use informative molecular regions for delineating populations and bringing to light relationships and historical processes for this species. We analyzed 349 individuals from 54 geographic collections from Mexico, Central America, the Caribbean, and South America. We observed 61 haplotypes that are structured into four phylogenetic clades. The distribution of these clades among populations is associated with geography. Six populations are identified in this analysis: Mesoamerica, C. America, Caribbean, western Mexico, Andean S. America, and eastern Brazil. AMOVA supports this arrangement and we reveal historical processes that may have contributed to the current distribution of populations. We describe those processes and possible scenarios that formed the current set of populations of this fly.

## 1. INTRODUCTION

Fruit flies in the genus *Anastrepha* Schiner are among the most widespread and economically important fruit flies in tropical and subtropical areas in the New World (Aluja, 1994; White & Elson-Harris, 1992). Effective management of these flies requires a thorough knowledge of their taxonomy, geographic distribution, and biology. Previous genetic studies of tephritid fruit flies have addressed aspects of their evolutionary biology and population genetic structure that are important if we are to improve the management of these quarantine pest species. To date, however, few studies, as compared to other tephritid genera such as *Bactrocera* and *Ceratitis*, have addressed the population genetics of *Anastrepha* spp. (Alberti et al., 2008, 2002, 1999; Boykin et al., 2006; Steck, 1991; Vilardi et al., 1994)

*Anastrepha obliqua* (Macquart), the West Indian fruit fly, occurs from northern Mexico to southern Brazil and the Caribbean (Steck, 2001; Aluja, 1994; Foote et al., 1993; Hernández-Ortiz and Aluja, 1993), and is occasionally found in Texas and California (Epsky et al., 2003). This species is a member of the *fraterculus* taxonomic group which contains *A. fraterculus* (S. American fruit fly) considered a cryptic species complex (Hernandez-Ortiz et al., 2004; Lutz & Lima, 1918; Stone, 1942; Zucchi, 1981). It infests at least 60 plant species belonging to 24 plant families in tropical and subtropical areas (Aluja et al., 1996; Hernández-Ortiz and Aluja, 1993; Norrbom and Kim, 1988).

Morphological descriptions used for identifying *A. obliqua* (Foote et al., 1993; White & Elson-Harris, 1992) and other *Anastrepha* species require substantial expertise because characters are not always unambiguous. Morphological keys are ineffective at times when characters such as coloration are subject to interpretation or differences are so subtle that only a few highly trained

taxonomists can discern among individuals to species-level. This is particularly true for immature stages – taxonomic coverage is poor and characters are sometimes difficult to interpret (Steck et al., 1990 Proc Ent Soc Wash reference on larval descriptions). These traditional methods do not reliably distinguish among members belonging to the *fraterculus* group, which includes *A. obliqua* (Norrbom et al., 1999). Names synonymous with *A. obliqua* include: *Anastrepha fraterculus* var. *mombinpraeoptans* (Seín), *Anastrepha mombinpraeoptans* (Seín), *Anastrepha acidusa, Anastrepha trinidadensis* (Greene), *Anastrepha ethalea* (Greene), *Anastrepha fraterculus* var. *ligata* (Costa Lima), *Acrotoxa obliqua* (Macquart), *Tephritis obliqua* (Macquart), and *Trypeta obliqua* (Macquart).

The question of population structure within *A. obliqua* has not been formally addressed. Previous analyses of the cytochrome oxidase subunit 1 (COI) by Smith-Caldas et al. (2001) to analyze the phylogenies and detect cryptic species indicated a high diversity within the species. Smith-Caldas et al. (2001) report the separation of *A. obliqua* into two clades. One clade with strong support contains four *A. obliqua*, two gathered in Mexico (Veracruz), one in Colombia (Sevilla), and one in Brazil (Bahia). The second clade shows *A. obliqua* gathered from four states (ES, SP, RN, MG) in Brazil grouped along with two *A. sororcula* gathered in Brazil (SP and RN) and four *A. fraterculus* also gathered in Brazil (BA, ES, MG, SC). It is possible that *A. obliqua* represents multiple reproductively isolated populations and we test for this hypothesis. Prior work on another *Anastrepha* species, the South American fruit fly, *A. fraterculus*, presented strong evidence of cryptic species (reviewed by Silva and Barr, 2008).

Mitochondrial DNA sequences have been used successfully on tephritids for resolving inter-specific relationships (Barr and McPheron, 2006; Barr et al., 2005; Boykin et al., 2006; McPheron et al., 1999; Smith-Caldas et al., 2001; Han and McPheron, 1999), intra-specific structure (Alberti et

al., 2008; Lanzavecchia et al., 2008; McPheron et al., 1994), and to develop diagnostic methods for species ID (Barr et al., 2006) and pathway analysis of populations (Lanzavecchia et al., 2008; Barr, 2009). Additionally, studies on evolutionary processes that examine nucleic acid variation have expanded our knowledge of phylogenetic relationships among sibling and cryptic species (Xie et al., 2008; Feder et al., 2003) providing insight on diversity and evolutionary pathways. Much has been studied of the distribution and close-knit relationship between tephritid pest species and their hosts (Novonty et al., 2005; Tan & Serit, 1994; Malavasi & Morgante, 1981). Furthermore, substantial behavioral research has been conducted on *Anastrepha* spp. (Hernandez et al., 2007; Aluja and Diaz-Fleischer, 2006).

In this study we use mitochondrial DNA from the COI and the NADH subunit 6 (ND6) genes to infer the population genetic structure of *Anastrepha obliqua* collected in Mexico, Central America, the Caribbean, and South America. We examine the data for evidence of demographic processes that may have contributed to population history by using demographic statistics, phylogeography (Avise, 2000; Knowles and Madison, 2002), and the coalescent theory (Nielsen, 1998; Hudson, 1991; Tavare, 1984). The *A. obliqua* data will also be used to infer the evolutionary history of mtDNA within the species and explore the significance of this genealogical history to species evolution. We propose to use methods here that will provide management programs additional tools for delineating populations and assessing the biodiversity of this species. The phylogeographic approach could reveal information that helps us understand the history of this fly and better predict infestations. We examine variability among *A. obliqua* haplotypes and consider the taxonomic and diagnostic significance of the data set.

## 2. MATERIALS AND METHODS

*Anastrepha obliqua* used in this study were collected in traps (adults) and from host fruit material (larvae) between 1995 and 2006. The 54 sampling sites, with ecological information and geographic distribution, are described in Table 22 and Fig. 14. Specimens were collected from multiple hosts – mango (*Mangifera* sp.), plum (*Spondias* sp.), carambola (*Averrhoa carambola*), and citrus (*Citrus* sp.). Prior to analysis, specimens were captured live, identified by morphology or through association to host by trained taxonomists, then either frozen or placed in ethanol and shipped to the USDA Center for Plant Health Science and Technology, Mission Lab in Edinburg, TX, or Penn State University (PSU), Department of Entomology, University Park, PA. Entire bodies from adult and larvae specimens were maintained at –80$^{o}$C at both locations prior to analysis.

Table 22. Geographical origins and information on collections of *Anastrepha obliqua* and other related species of *Anastrepha* included in this study.

| | | Population | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Site# | Locality | Loc code | Pop | *n* | Latitude | Longitude | Year coll. | Host |
| 1 | Barbados | BAR1 | Car | 2 | 13.193N | -59.543W | 2004 | |
| 2 | Belize, Dangriga | BEL1 | Mes | 2 | 17.183333N | -88.583333W | 2006 | mango |
| 3 | Belize, Stann Creek | BEL2 | Mes | 1 | 16.99N | -88.323W | 2006 | mango |
| 4 | Brazil, Acu, RGN | BRA1 | EBra | 9 | 5.345968S | -36.34938W | 2005 | trap |
| 5 | Brazil, Atibaia, Sao Paulo | BRA2 | EBra | 2 | 23.116096S | -46.552861W | | |
| 6 | Brazil, ES | BRA3 | EBra | 9 | 19.596027S | -40.772472W | 1998 | |
| 7 | Brazil, Janauba, Minas G | BRA4 | EBra | 6 | 15.803693S | -43.317366W | 1997 | trap |
| 8 | Brazil, Patrocinio Paulista, SP | BRA5 | EBra | 1 | 20.639S | -47.282W | | |
| 9 | Brazil, Laranjal Paulista, SP | BRA6 | EBra | 9 | 23.053499S | -47.834049W | 1997 | *Averrhoa carambola* |
| 10 | Brazil, Linhares, ES | BRA7 | EBra | 8 | 19.390915S | -40.071503W | 1998 | trap |
| 11 | Brazil, Mossoro, RGN | BRA8 | EBra | 9 | 5.188036S | -37.344131W | 1999 | trap |
| 12 | Brazil, Presidente Prudente, SP | BRA9 | EBra | 1 | 22.127596S | -51.385616W | | hobo |

Table 22. (cont.)

| Site# | Locality | Loc code | Population: Pop | *n* | Latitude | Longitude | Year coll. | Host |
|---|---|---|---|---|---|---|---|---|
| 13 | Colombia, Bucaramanga | COL1 | And | 7 | 7.105N | -73.116W | 2006 | mango |
| 14 | Colombia, Caldas | COL2 | And | 4 | 5.28265N | -75.302W | 2006 | McPhail trap |
| 15 | Colombia, Giron | COL3 | And | 1 | 7.0718N | -73.175W | 2006 | mango |
| 16 | Colombia, Tolima | COL4 | And | 5 | 4.129N | -75.302W | 2006 | *Spondia* sp. |
| 17 | Costa Rica, Alajuela Province | CR1 | CAm | 4 | 10.45041N | -84.794076W | 2005 | McPhail trap |
| 18 | Costa Rica, Puntarenas Province | CR2 | CAm | 10 | 9.18525N | -84.005369W | 2005 | mango |
| 19 | Dominican Republic, Bonao del Tanque | DR1 | Car | 8 | 18.94N | -70.409W | 2003 | |
| 20 | Dominican Republic, Moca | DR2 | Car | 10 | 19.400086N | -70.51996W | 2004 | hobo |
| 21 | El Salvador, Pipil | ES1 | CAm | 9 | 13.426N | -88.951W | | McPhail trap |
| 22 | Guatemala, Yepocapa | GT1 | CAm | 5 | 14.500008N | -90.950017W | 2001 | trap |
| 23 | Honduras, Comayaqua | HON1 | CAm | 9 | 14.561852N | -87.672394W | 2004 | mango |
| 24 | Honduras, F Morazan | HON2 | CAm | 12 | 15.32395N | -87.599274W | 2004 | mango |
| 25 | Jamaica, Jamaica | JAM1 | Car | 15 | 18.2N | -78.01W | 2004 | mango |
| 26 | Jamaica, Kingston | JAM2 | Car | 9 | 17.989999N | -76.800003W | 1996 | *Spondias purpurea* |
| 27 | Mexico, Campeche, Aguacatal | MX1 | Mes | 10 | 18.624809N | -90.740992W | 2006 | multi-lure trap |
| 28 | Mexico, Campeche, Escarcega | MX2 | Mes | 9 | 18.624809N | -90.740992W | 2001 | mango |
| 29 | Mexico, Morelos, Michipan | MX3 | WMx | 8 | 18.91666667N | -99.06666667W | 1999 | mango |
| 30 | Mexico, Oaxaca, Zacatepec | MX4 | Mes | 5 | 17.191278N | -95.885536W | 2001 | mango |
| 31 | Mexico, Tabasco, Agua Blanca | MX5 | Mes | 4 | 17.944N | -92.583W | 2006 | mango |
| 32 | Mexico, Tabasco, Huemanguillo | MX6 | Mes | 8 | 17.827141N | -93.399399W | | |
| 33 | Mexico, Tamaulipas, Ciudad Mante | MX7 | Mes | 3 | 22.744555N | -98.970428W | | sour orange |
| 34 | Mexico, Veracruz, Los Tuxtlas | MX8 | Mes | 1 | 18.560344N | -95.204088W | 1995 | *Spondias mobin* |
| 35 | Mexico, Colima, Tecoman | MX9 | WMx | 5 | 18.91423N | -103.875198W | | mango |
| 36 | Mexico, Guerrero, Acahuizotla | MX10 | WMx | 9 | 17.405868N | -99.446206W | 2005 | mango |
| 37 | Mexico, Guerrero, La Palma | MX11 | WMx | 4 | 20.2312N | -102.781631W | 2006 | mango |
| 38 | Mexico, Jalisco, Mezquitlan | MX12 | WMx | 2 | 19.817N | -104.34W | 2005 | mango |
| 39 | Mexico, Michoacan, Nueva Italia | MX13 | WMx | 6 | 19.017N | -102.1W | 2005/2006 | mango |
| 40 | Mexico, Michoacan, Sta Amate | MX14 | WMx | 10 | 19.162N | -101.901W | 2000 | mango |
| 41 | Mexico, Morelos, Cuatla | MX15 | WMx | 7 | 18.863498N | -98.86302W | | mango |
| 42 | Mexico, Nayarit, El Jigote | MX16 | WMx | 9 | 21.708699N | -105.064329W | | mango |
| 43 | Mexico, Nayarit, Valle de Lerma | MX17 | WMx | 10 | 21.8154N | -105.084169W | 2000 | mango |
| 44 | Mexico, Sinaloa, Concordia | MX18 | WMx | 6 | 23.278907N | -106.076797W | 2004 | mango |
| 45 | Mexico, Sinaloa, Escuinapa | MX19 | WMx | 10 | 22.84561N | -105.901299W | 2000 | mango |
| 46 | Mexico, Sinaloa, Malpica | MX20 | WMx | 5 | 23.233333N | -106.153643W | 2004 | mango |
| 47 | Panama, Altos de Pacora | PAN1 | CAm | 7 | 9.081758N | -79.290819W | 2006 | McPhail trap |
| 48 | Panama, Cerro Azul | PAN2 | CAm | 6 | 9.166669N | -79.416669W | 2006 | mango |
| 49 | Panama, Chiriqui | PAN3 | CAm | 7 | 8.131187N | -83.348255W | 2006 | mango |
| 50 | Peru, San Martin Province | PER1 | And | 4 | 6.14S | -77.09W | 2005 | *Spondias* sp. |
| 51 | Puerto Rico, Hormiguoros | PR1 | Car | 6 | 18.221315N | -66.430649W | 2006 | trap |
| 52 | Puerto Rico, Maricae | PR2 | Car | 3 | 18.221315N | -66.430649W | 2006 | trap |
| 53 | Puerto Rico, San Juan | PR3 | Car | 8 | 18.46616N | -66.10666W | | McPhail trap |
| 54 | St Vincent, St Vincent | STV1 | Car | 5 | 13.212N | -61.192W | 2004 | unknown |

Table 22. (cont.)

| Site# | Locality | Loc code | Population: Pop | *n* | Latitude | Longitude | Year coll. | Host |
|---|---|---|---|---|---|---|---|---|
| Additional *Anastrepha* spp. included for comparision | | | | | | | | |
| *A. coronilli* | Tachira, Venezuela | | | 1 | 7.997N | -71.918W | 1993 | *Bellacia axinanthera* |
| *A. sororcula* | Bahia, BA, Brazil | | | 1 | -11.409N | -41.281W | 1989 | trap |
| *A. suspensa* | Manabao, Dominican Republic | | | 1 | 19.066N | -70.787W | | Goiaba |
| *A. turpiniae* | Panama City, Panama | | | 1 | 8.994N | -79.519W | | |
| *A. zenildae* | Janauba, MG, Brazil | | | 1 | -15.804N | -43.317W | 1997 | |
| *A. fraterculus* | Espiritu Santo, Brazil | | | 2 | -19.596N | -40.772W | 1997 | |
| *A. amita* | Piracicaba, SP, Brazil | | | 1 | -22.736N | -47.647W | 1998 | |
| *A. distincta* | Balzapote, Ver, Mx | | | 1 | 18.602N | -95.058W | | *Inga jinicuil* |
| *A. bahiensis* | Santa Rosa, Guatemala | | | 3 | 14.139N | -90.312W | | *Brosimum costaricum* |
| *A. pseudoparallela* | Laranjal Paulista, SP, Brazil | | | 1 | -23.054N | -47.834W | 1996 | *Passiflora edulis* |
| *A. alveata* | Llano Grande, Ver, Mx | | | 1 | 19.318N | -96.761W | 1994 | *Ximenia americana* |
| *A.pallens* | Monterrey, NL, Mx | | | 1 | 25.67N | -100.31W | 1997 | |

*n*, number of individual specimens analyzed. Blank fields indicate missing information.


Figure 14. Map showing the distribution of collection sites for *Anastrepha obliqua* specimens used in this study. Geographic populations are delineated according to results of analyses. Collection site numbers correspond to locality in Table 22.

*2.1* DNA extraction, PCR, and sequencing

DNA was isolated using the DNeasy® Blood and Tissue Kit (Qiagen) following standard DNeasy guidelines for animal tissues. Extractions were stored at –20°C for the duration of the study. Whenever possible, DNA from one specimen per collection was isolated via a non-destructive method (Barr et al., 2006) to save the tissue as voucher. Samples were amplified for the mitochondrial cytochrome oxidase subunit I (COI) region using tRNA-cys2 5'-ACTCCTTTAGAATTGCAGTCTAAT-3' and COId-r 5'-GGGCTCATACAATAAATCCTAAT-3' designed from the *Bactrocera abdominalis* (GenBank accession number DQ917577) sequence, generating a fragment of approximately 1050 bp. This fragment included partial coverage of the cysteine tRNA and the entire tyrosine tRNA and partial coverage of the cytochrome oxidase subunit I. The mitochondrial NADH dehydrogenase subunit 6 (ND6) region was amplified using TT-J-9886 5'-TAAAAACATTGGTCTTGTAA-3' (Barr et al., 2006) and ND6r 5'-TTATGATCCAAAATTTCATCA-3' primers and yielded fragments of approximately 800 base pairs. This fragment included a portion of the threonine tRNA, the entire proline tRNA, the entire NADH subunit 6 codon, and partial coverage of the cytochrome B codon. Additionally, we used an alternate reverse primer, COIc-r 5'-GCTATTATTGCATAGATTATTCC-3' designed from *A. ludens* (R. Ruiz-Arce, unpublished) for this work. PCR reactions were performed in 25 μL reactions containing 1 μL of template, 16.875 μL of water, 2.5 μL of buffer (10X conc., Qiagen), 1.5 μL of 25 mM of $MgCl_2$ (Qiagen), 1.25 μL of each primer (5 μM, Operon Technologies), 0.5 μL of 10 mM dNTP mix (Promega), and 0.125 μL of Taq DNA polymerase (Qiagen) to complete the final volume of 25 μL. Amplifications were performed in an Applied Biosystems GeneAmp® PCR System *9700*. Cycling conditions for amplification of COI and ND6 fragments were 3 min at 94°C followed by 39 cycles of 1 min at 94°C, 1 min at 50°C, 1 min at 72°C, and a final extension of 10 min at 72°C. PCR

products were stained with Sybr® Green (Invitrogen) fluorescent dye at 1/10,000X and 3/10,000X and visualized on 1.5% electrophoresis agarose gels. Documentation of these gels was via Gel Doc™ (Bio-Rad) imaging system using Quantity One® software. Amplification products were purified with ExoSAP-IT (USB Corp.) prior to sequencing. PCR products were sequenced asymmetrically using 3' BigDye-labeled dideoxynucleotide triphosphates (v 3.1 dye terminators, Applied Biosystems, Foster City, CA) and run on an ABI 3730XL DNA Analyzer with the ABI Data Collection Program (v 2.0) at the Huck Institute's Nucleic Acid Facility at Penn State University.

Sequences were edited with QScreen (PSU: Huck Institute) designed in collaboration with the McPheron Lab (PSU: Dept of Entomology), and Sequencher 4.8 (Gene Codes Corp.). All sequences were aligned with Mega 4 (Tamura et al., 2007) using ClustalW (Thompson et al., 1994) and trimmed to a length of 812 bp and 736 bp for COI and ND6, respectively. We trimmed approximately 250bp from the five-prime end of the tRNA-cys2-COI PCR fragment to avoid ambiguous alignment of observed insertions/deletions in the tRNA region (Data not shown). Sequences were concatenated using DnaSP 4.10 (Rozas et al., 2003), and subsequent analyses were carried out on the 1548 bp fragment.

*2.2* Phylogenetic analyses and divergence estimates of haplotypes

Phylogenetic tree reconstruction was estimated using Bayesian, maximum parsimony (MP), and maximum-likelihood (ML) approaches. The appropriate model of sequence evolution for Bayesian and ML analyses was determined with hierarchical likelihood tests (hLRTs) using Modeltest 3.7 (Posada and Crandall, 1998). Bayesian analyses were performed using Mr Bayes 3.1 Ronquist and Huelsenbeck, 2003) employing a Monte Carlo Markov Chain (MCMC) method with

four MCMC chains run simultaneously for 1,000,000 generations with an initial 250 generation burn-in. Trees were sampled every 1000th generation. Posterior probabilities were determined from 50% majority-rule consensus trees generated from the remaining saved trees. MP phylogenetic reconstruction was performed by bootstrapping with 500 replications using Mega with Close-neighbor-interchange (CNI) level set at 1 and Random Addition Trees at 10 replications. Maximum parsimony was done on all codon positions ($1^{st}$+$2^{nd}$+$3^{rd}$+Noncoding) with Gaps/Missing Data set at Complete Deletion. ML bootstrap analyses were conducted with 500 replicates (Felsenstein, 1985). The Phyml (Guindon et al., 2005) ML tree search was initiated using a NJ tree that was constructed with Mega incorporating the Modeltest parameters. We rooted these trees with one individual of *Anastrepha ludens.*

Haplotypes were identified from concatenated sequences with DnaSP. Our haplotype file was constructed from all 1548 sites. Since no gaps were recovered from our sequencing they were not considered in the construction. The relative haplotype frequencies within populations were calculated using ARLEQUIN 3.11 (Excoffier et al., 2005).

We estimated divergence times of *A. obliqua* haplotypes from phylogenetic clades using the program MDIV (Nielsen & Wakeley, 2001). MDIV employs a Bayesian approach to approximate the posterior distribution of divergence time ($T = t_{div}/2N_e$), migration rate since divergence ($M = 2N_em$) between haplotypes, and theta ($\theta = 2N_e\mu$, where $\mu$ is the per locus mutation rate between clades). Initial runs using default settings were conducted to estimate time to convergence. These estimates were then used to fine tune and improve the efficiency of the analysis. Analyses were run for 5,000,000 generations MCMC following a burn-in period of 500,000 generations and repeated three times with each run having a unique random number seed. These contrasts used a finite sites

model, and a priori maxima were set for θ, *M*, and *T*. Estimates of *T* were converted to real time (*T*pop) with a range of mtDNA mutation rate estimates for arthropods: $1.0 \times 10^{-8}$, $1.15 \times 10^{-8}$, and $1.2 \times 10^{-8}$ substitutions/site/year (Brower, 1994) using $T\text{pop} = [(T \cdot \theta)/2L]1/\mu$ (Brito, 2005), where L is the sequence length (1548 bp) and μ the mutation rate per site per generation. We used this range of calibration because it is broadly referenced and it is based on several mtDNA regions within the mitochondrial genome. We factor in our calculations that this species, on average, undergoes six generations per year (Weems, 1970).

*2.3* Population structure

Our initial delineation of populations was based on natural geographic barriers which allowed us to define six geographic populations. The Mesoamerica population consisted of those collections made east of the Sierra Madre Oriental and south to the Sierra Madre de Chiapas mountain ranges and included two Central American countries, Belize and Guatemala. The Central America population was made of collections south of the Sierra Madre de Chiapas mountain ranges to the Panama Basin. The Caribbean population consisted of collections made from the Caribbean islands. The western Mexico population consisted of collections west of the Sierra Madre Oriental and along the northern and southern regions of the western coast of Mexico. The Andean population included collections from Colombia and Peru. And, the eastern Brazilian population was made up of collections from eastern Brazil.

ARLEQUIN was used to calculate *F*st (Wright, 1951, Weir and Cockerham, 1984) values and to partition the source of variation among geographic groups, among populations within groups, and within populations using analysis of molecular variance (AMOVA, Excoffier et al., 2005); a permutation method with 10,000 replications was applied to test significance of differences. We

used AMOVA to examine partitioning of mtDNA genetic diversity within and among the geographic collections.

*2.4* Inferred history using summary statistics and phylogeographic analyses

Basic molecular summary statistics, genetic distances, and mismatch distributions were estimated with DnaSP. Mismatch distributions were used to examine the data for evidence of demographic population expansions that can result in unimodal, wave-like patterns (Rogers and Harpening 1992). Demographic parameters Tajima's *D* (Tajima, 1989) and Fu's *Fs* (Fu, 1997) were estimated for the sampled collections (and groups of collections), and their confidence intervals were assessed by 10,000 coalescent simulations to test the hypothesis that all mutations are selectively neutral (Kimura, 1983) with ARLEQUIN. The relationship between mtDNA divergence and geographic distance (isolation by distance) for populations was tested using Mantel tests (Mantel, 1967) with 1000 replications in IBD version 1.52 (Bohonak, 2002).

In order to examine whether associations between genetics and geography for populations were significant and to gain a more thorough understanding of complex biological processes influencing the distribution of *A. obliqua*, we employed a nested clade phylogeographic analysis (NCPA) (Excoffier et al., 1992; Templeton et al., 1995). Although the legitimacy of NCPA methods has been discussed in recent literature (Templeton, 2009 and 2008; Knowles, 2008; Beaumont & Panchal, 2008; Petit, 2008) and claims of incorrect historical processes inferred in over 75% of simulated datasets (Panchal & Beaumont, 2007), we cautiously incorporate the use of this method here along with support from additional statistical phylogeographic analyses to serve in revealing important historical processes and haplotype associations.

A maximum parsimony haplotype network was constructed according to the algorithm described in Templeton et al. (1992) and included in the program TCS 1.18 (Clement et al., 2000). TCS was used to generate schematics representative of parsimonious relationships between and among collections according to geographic association. A nested clade phylogeographic analysis (NCPA) (Templeton et al., 1995) was performed using ANeCA (Panchal, 2007), an automated software tool, to differentiate between population structure and historical events (Templeton, 1998). Nested clade contingency tests were performed in GeoDis v2.5 (Posada et al., 2000) and 2005 inference key (updated November 11, 2005) was used to determine whether phylogeographical differences were best explained by allopatric fragmentation, recurrent but restricted gene flow, or range expansion, thus providing insight into the evolutionary history and population structure of this species. Frequency and shared locality were used to resolve ambiguities (Pfenninger and Posada 2002).

## 3. RESULTS

### *3.1* Sequence variation

A total of 61 composite (COI+ND6) haplotypes were observed among the 349 individual *A. obliqua* specimens sequenced. There were 82 (5.3%) polymorphic and 57 (3.7%) parsimoniously informative sites observed (Table 23) in the 1548bp concatenated fragment with no gaps or indels observed. The COI (812bp) and ND6 (736bp) fragments contributed 35 and 47 polymorphic sites,

respectively. This resulted in 44 private haplotypes among the 54 geographic collections. The mean number of haplotypes per geographic location was 2.5 and ranged between 1 and 7. The haplotype and nucleotide diversities were 0.9022 and 0.01153, respectively, for the entire West Indian fruit fly dataset.

Table 23. Alignment showing the 61 haplotypes and 82 variable sites for COI and ND6 concatenated sequences observed in *Anastrepha obliqua* collections.

| Positions | COI (35) | ND6 (47) |
|---|---|---|
| | | **1111111111111111111111111111111111111111111** |
| | **1112222333333344555555566666777** | **8999900000001111111111222222222222233333444445555** |
| | **1257235022412347996913457783588878** | **6346622255570011557778001113356799235561257801 3** |
| | **5374908972892542537559199138671271 3** | **6991901301272706671233031291670334955989232317 7** |
| OB01 | GTTTGATTAGTACCGTTACAGTCAGCATCTTCTTA | TTTTTACAAGCAGAGCCTGCTAGACCTTTTCTTACCTATCACCCTGC |
| OB02 | A..... | .....A..... |
| OB03 | .....C.....C.. | ..... |
| OB04 | ..... | .....G..... |
| OB05 | A..C.....ACGT.....T.A.T.....T..T | ...C..T.G...A.A.....C.ATTTCCCCTAC....GC..T..A.. |
| OB06 | A..C.....ACGT.....T.A.T.....T..T | ...C..T.G.....A.....C.ATTTCCCCTAC....GC..T..A.. |
| OB07 | ..... | .....A.....T..... |
| OB08 | ..... | .....C.....T....T..... |
| OB09 | ..... | .....C..... |
| OB10 | ..... | .....A..... |
| OB11 | ..... | .....C.....T..... |
| OB12 | .....T..... | .....A..... |
| OB13 | ....A..... | ..... |
| OB14 | .....C..... | .....T..... |
| OB15 | ..... | .....GA..... |
| OB16 | .....C..... | ..... |
| OB17 | .....A..... | .....A..... |
| OB18 | .....C...C..... | ..... |
| OB19 | ..... | .....T..... |
| OB20 | .....A..... | ..... |
| OB21 | ..... | .....C.....T..... |
| OB22 | A..C.....ACGT.....T.A.T.....T... | ...C..T.G...A.A.....C.ATTTCCCCTAC....GC..T..A.. |

Table 23. (cont.)

| Positions | COI (35) | ND6 (47) |
|---|---|---|
| | | 11111111111111111111111111111111111111111 |
| | 11122223333333445555556666677 | 89999000000011111111122222222222233333444445555 |
| | 12572350224123479969134577835888878 | 63466222555700115577780011133567992355612578013 |
| | 53749089728925425375591991386712713 | 69919013012727066712330312916703349559892323177 |
| OB01 | GTTTGATTAGTACCGTTACAGTCAGCATCTTCTTA | TTTTTACAAGCAGAGCCTGCTAGACCTTTTCTTACCTATCACCCTGC |
| OB23 | ..... | .....C.....C..... |
| OB24 | A..C.....A..T......CA.TG.......T.CT | ...C..T.G.....AT..A...ATTT....TAC....GC..T.T... |
| OB25 | A..C.....A..T......CA.TG.......T..T | ...C..T.G.....AT..A...ATTT....TAC....GC..T.T... |
| OB26 | .....T..... | ..... |
| OB27 | A.CC.....A..T......CA.TG.......T.CT | ...C..T.G.....AT..A...ATTT....TAC....GC..T.T... |
| OB28 | A..C.G...A..T......CA.TG.......T.CT | ...C..T.G.....AT..A...ATTT....TAC....GC..T.T... |
| OB29 | A..C.....A..TT.....CA.TG.......T.CT | ...C..T.G.....AT..A...ATTT....TAC....GC..T.T... |
| OB30 | ..... | .....C..... |
| OB31 | A..C.....A..T......CA.TG.......T.CT | ...C..T.......AT..A...ATTT....TAC....GC..T.T... |
| OB32 | A..C.....A..T......CA.TG.......T.CT | ...C..T.G....GAT..A...ATTT....TAC....GC..T.T... |
| OB33 | .....A..... | ..... |
| OB34 | .C..... | ..... |
| OB35 | ..... | .....A. |
| OB36 | ..... | .....C........T..... |
| OB37 | A..C.....ACGT.....T.A.T........T..T | ...C..T.G...A.A..A..C.ATTTCCCCTAC....GC..T..A.. |
| OB38 | A..C.....A..T......CA.TG.......T.CT | ...C..T.G.....AT..AT..ATTT....TAC....GC..T.T... |
| OB39 | A..C.....A..T......CA.TG.......T.CT | ...C..T.G.T...AT..A...ATTT....TAC....GC..T.T... |
| OB40 | A..C.....A..T......CA.TG.......T.CT | ...C..T.G...C.AT..A...ATTT....TAC....GC..T.T... |
| OB41 | .......C.A........T...T...TC.C..... | .C.....G.............A.....C....T...G......... |
| OB42 | A..... | ..... |
| OB43 | ..............C..... | ..... |
| OB44 | A..... | C..... |
| OB45 | .....T......C.. | ..... |
| OB46 | A..C.....A..T......CA.TG.........CT | ...C..T.GA....AT..A...ATTT....TAC....GC..T.T... |
| OB47 | A..C.....A..T......CA.TG.......T.CT | ...C..T.G.....AT..A...ATTT....TAC..T.GC..T.T... |
| OB48 | .............A..... | .C...........A.T..... |
| OB49 | .............A.A..... | .C...........A.T..... |
| OB50 | ..................A..... | ..... |
| OB51 | A..C.....ACGT.....T.A.T........T.CT | ...C..T.G.....A.....C.ATTTCCCCTAC....GC..T..A.. |
| OB52 | A..C.....A..T......CA.TG.......T.CT | ...C..T.G.....A.....C.ATTTCCCCTAC....GC..T..A.. |

Table 23. (cont.)

| Positions | COI (35) | ND6 (47) |
|---|---|---|
| | | 11111111111111111111111111111111111111111 |
| | 11122223333333445555556666677 | 89999000000011111111122222222222233333444445555 |
| | 12572350224123479969134577835888878 | 63466222555700115577780011133567992355612578013 |
| | 53749089728925425375591991386712713 | 69919013012727066712330312916703349559892323177 |
| OB01 | GTTTGATTAGTACCGTTACAGTCAGCATCTTCTTA | TTTTTACAAGCAGAGCCTGCTAGACCTTTTCTTACCTATCACCCTGC |
| OB53 | ..............C.............C.C.. | ..... |
| OB54 | ..............C..... | ..... |
| OB55 | ..... | ..............A.....A.... |
| OB56 | ..... | ..C.C.........A..... |
| OB57 | ..... | .........A...........C..... |
| OB58 | ..........T..... | .....G..... |
| OB59 | .......CGA........T...T...TC.C.T... | .C...T.............A.....C....TT..... |
| OB60 | .......CGA........T...T...TC.C.T... | .C...T.............A.....C....TT..G..... |
| OB61 | ......C.....T..... | .....A |

*3.2* Phylogenetic reconstruction and divergence estimates

The HKY+I+G substitution model was found to be most applicable to the concatenated COI and ND6 sequence based on the Akaike ranking (Akaike, 1974). Modeltest revealed a proportion of invariable sites (I) of 0.7217, transition to transversion ratio of 5.38, and the gamma (Γ) shape distribution shape parameter of 0.3210. Phylogenetic reconstructions with Bayesian, ML, and MP methods provide concordant results that reveal four phylogenetic clades (Fig. 15).

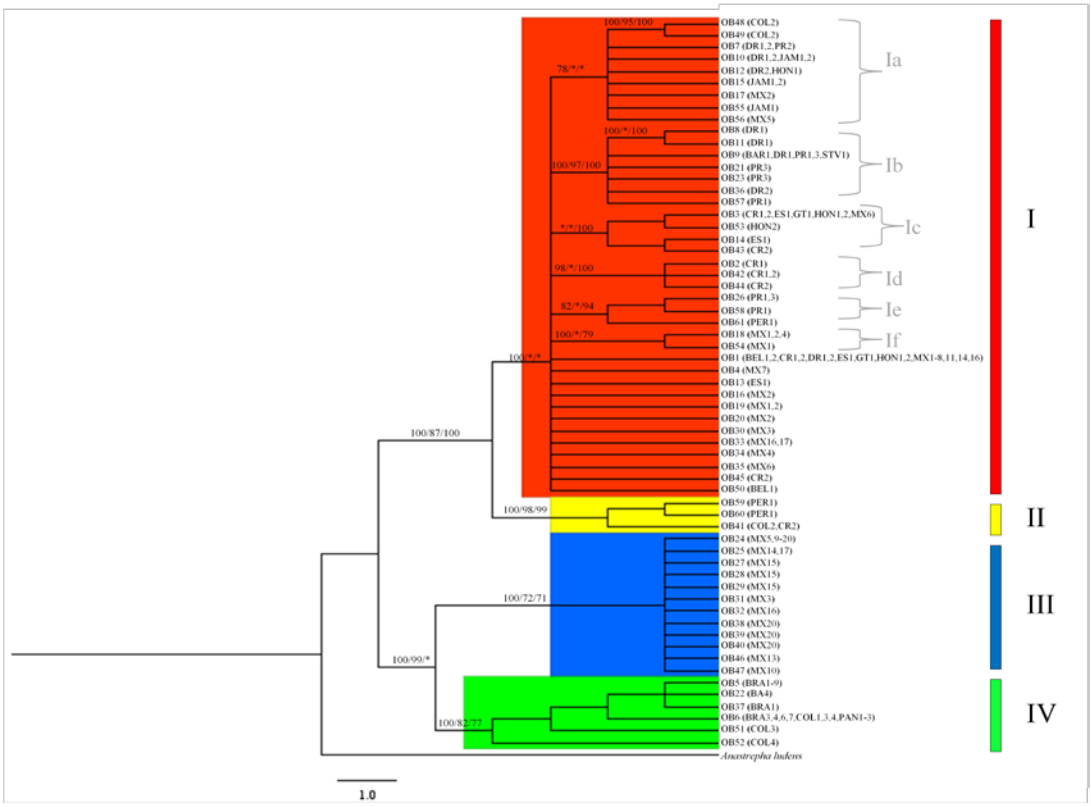


Figure 15 Phylogeny of 61 *Anastrepha obliqua* COI+ND6 haplotypes OB01-OB61, constructed using a Bayesian (B) approach. Four major clades were observed and internal clades within clade I are discussed in text. Numbers above branches show support for Bayesian/maximum parsimony/maximum likelihood. Population codes (e.g. COL2) correspond to Table 22. The tree was rooted with *A. ludens.* Asterisks designate unsupported branches with respective method or methods according to Hillis & Bull (1993).

We see moderate to strong bootstrap support for three clades (II, III, & IV) with all three phylogenetic methods. The average gamma-corrected genetic distance ranged from 0.011 – 0.025 among clades (Table 24) based on the 1548bp fragment. The haplotypes within these four clades show a close association with geographic distribution with little overlap (Fig. 14). Populations with haplotypes represented in multiple clades are reported in Table 25 and are the exceptions. Based on topology and branch support values, clades I and II share a common ancestral sequence and Clades III and IV share a different common ancestral sequence. These two major clades are estimated to have diverged approximately 137Ka (Table 26) with clades I & II and III & IV having a most recent common ancestor approximately 58Ka and 49Ka, respectively.

Table 24. Mean number of differences ±SE and gamma corrected mean pairwise distance ±SE (in parenthesis) among clades observed according to phylogenetic reconstruction for 1548bp COI+ND6 *Anastrepha obliqua* sequences.

| Clade (avg dist±SE) | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1. Clade I (0.002±0.0) | 0 | | | |
| 2. Clade II (0.003±0.001) | 16.13±3.62 (0.011±0.003) | 0 | | |
| 3. Clade III (0.002±0.001) | 32.57±5.24 (0.021±0.004) | 34.0±5.31 (0.025±0.004) | 0 | |
| 4. Clade IV (0.001±0.0) | 29.38±4.90 (0.023±0.004) | 34.61±5.18 (0.024±0.004) | 15.51±3.67 (0.011±0.003) | 0 |

Table 25. Geographic sites that harbor haplotypes from multiple phylogenetic clades (Fig 15) for *Anastrepha obliqua.* Population site codes (e.g. CR2) correspond to Table 22.

| Geographic site | *n* | Clade | Haplotype (# of indiv showing haplotype) |
|---|---|---|---|
| Costa Rica, Alajuela Province (CR2) | 10 | I | OB01 (2), OB03 (1), OB42 (3), OB43 (1), OB44 (1), OB45 (1) |
| | | II | OB41 (1) |
| Colombia, Caldas (COL2) | 4 | I | OB48 (2), OB49 (1) |
| | | II | OB41 (1) |
| Mexico, Morelos, Michipan (MX3) | 8 | I | OB01 (6), OB30 (1) |
| | | III | OB31 (1) |
| Mexico, Tabasco, Agua Blanca (MX5) | 4 | I | OB01 (1), OB56 (2) |
| | | III | OB24 (1) |
| Mexico, Guerrero, La Palma (MX11) | 4 | I | OB01 (1) |
| | | III | OB24 (3) |
| Mexico, Michoacan, Sta Amate (MX14) | 10 | I | OB01 (1) |
| | | III | OB24 (8), OB25 (1) |
| Mexico, Nayarit, El Jigote (MX16) | 9 | I | OB01 (1), OB33 (1) |
| | | III | OB24 (6), OB32 (1) |
| Mexico, Nayarit, Valle de Lerma (MX17) | 10 | I | OB33 (3) |
| | | III | OB24 (6), OB25 (1) |
| Peru, San Martin Province (PER1) | 4 | I | OB61 (1) |
| | | II | OB59 (2), OB60 (1) |

We see that clade I contains flies mainly from collections east of the Sierra Madre Oriental and south of the trans-Mexican volcanic belt, the Central American countries sampled here (excluding Panama), and all Caribbean Islands sampled (Fig. 3). We also see that within subclade Ia (Fig. 2), supported only moderately by Bayesian analysis, two southern Mexico, one Central America, two South America, and eleven Caribbean collections. Subclade Ib (Fig. 15: OB08, 11, 9, 21, 23, 36, & 57), strongly supported by all three phylogenetic methods, contains haplotypes seen only in Caribbean collections.

Table 26. MDIV estimates of divergence time among *Anastrepha obliqua* clades as observed from phylogenetic reconstruction. Divergence time was calibrated according to varying mutational rates.

| | | | | | | Population divergence time ($T_{pop}$) | | |
|---|---|---|---|---|---|---|---|---|
| Clade | vs | Clade | $T_{MRCA}$ | $t_{pop}$ | Θ | *m* | $1.0 \times 10^{-8}$ | $1.15 \times 10^{-8}$ | $1.2 \times 10^{-8}$ |
| I | | II | 1.73 | 1.11 | 11.27 | 0.036 | 67,343 | 58,559 | 56,119 |
| III | | IV | 5.99 | 4.71 | 2.22 | 0.004 | 56,289 | 48,947 | 46,907 |
| Ia | | Ib | 0.952 | 0.67 | 7.61 | 0.063 | 27,448 | 23,868 | 22,873 |
| I/II | | III/IV | 4.02 | 3.04 | 9.63 | 0.002 | 157,666 | 137,101 | 131,888 |

$T_{MRCA}$, most recent common ancestor, and $t_{pop}$ are measured in units of $2N_e♀$, generations; $\theta = 2N_e\mu$, and μ is the mutation rate per sequence per generation; *m*, number of migrants.

We calculated that clades Ia & Ib shared a most recent common ancestor approximately 24Ka (Table 26). Additional subclades reveal some regional specificity – Ic (Fig. 15), which is strongly supported by the ML method only, consists of haplotypes mainly from Central American collections, Id (Fig. 15) contains haplotypes exclusively of specimens from Costa Rica, Ie (Fig. 15) consists of Puerto Rico and Peru collections, and subclade If (Fig. 15) consists of specimens from southern Mexico collections. All of these subclades (Id-f) exhibit moderate to strong support with Bayesian and ML methods.

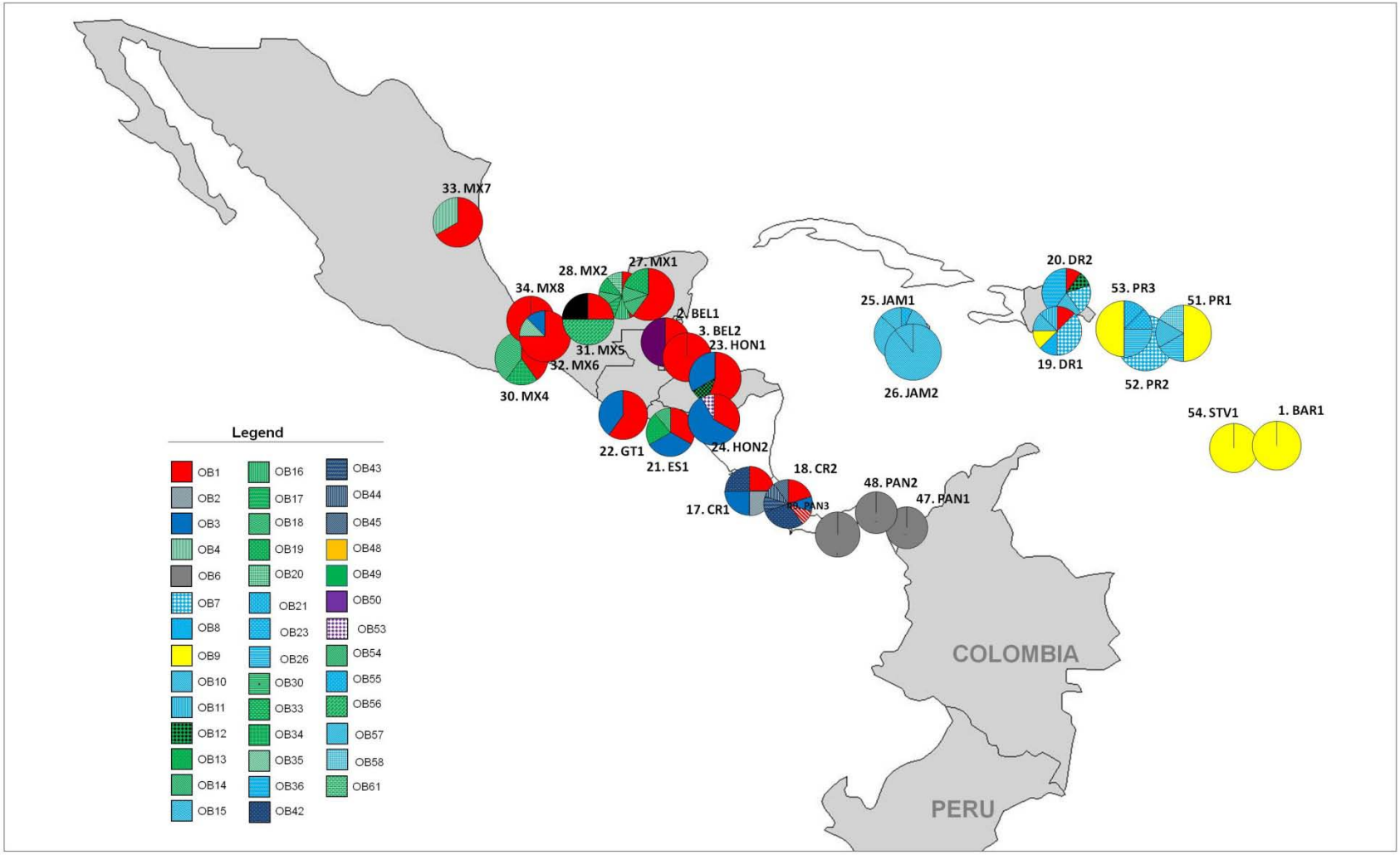


Figure 16. Geographic distribution of *Anastrepha obliqua* haplotypes seen within the Mesoamerica, C. America, and Caribbean populations. Pie charts show haplotype frequency, site number, and population code corresponding to locality in Table 22.

Clade II contains mostly individuals from the Andean region of S. America with the exception of one individual from Costa Rica (CR2) (Figs. 16&17). Clade II shows strong support by all three phylogenetic methods. Clade III contains collections found west of the Sierra Madre Oriental and north of the trans-Mexican volcanic belt (Fig. 18). Clade III is entirely represented by individuals from western Mexico except for one fly from the southeastern Mexican state of Tabasco (MX5) which has the OB24 haplotype, more commonly seen in western Mexico. We see moderate to strong support for this clade with phylogenetic methods. Clade IV was made up of individuals gathered in Brazil, Colombia, and Panama (Figs. 16&17). We also see moderate to strong support for this clade with phylogenetic methods.

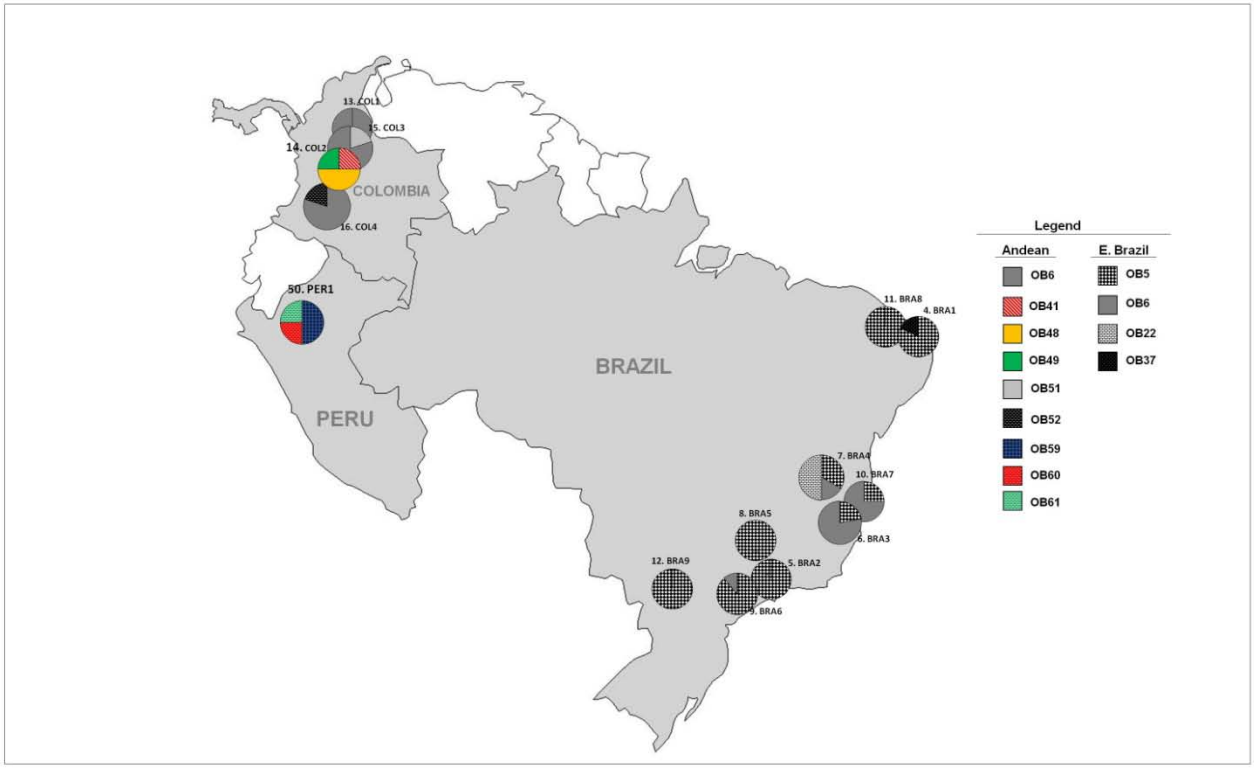


Figure 17. Geographic distribution of *Anastrepha obliqua* haplotypes seen within the Andean & eastern Brazil populations. Pie charts show haplotype frequency, site number, and population code corresponding to locality in Table 22.

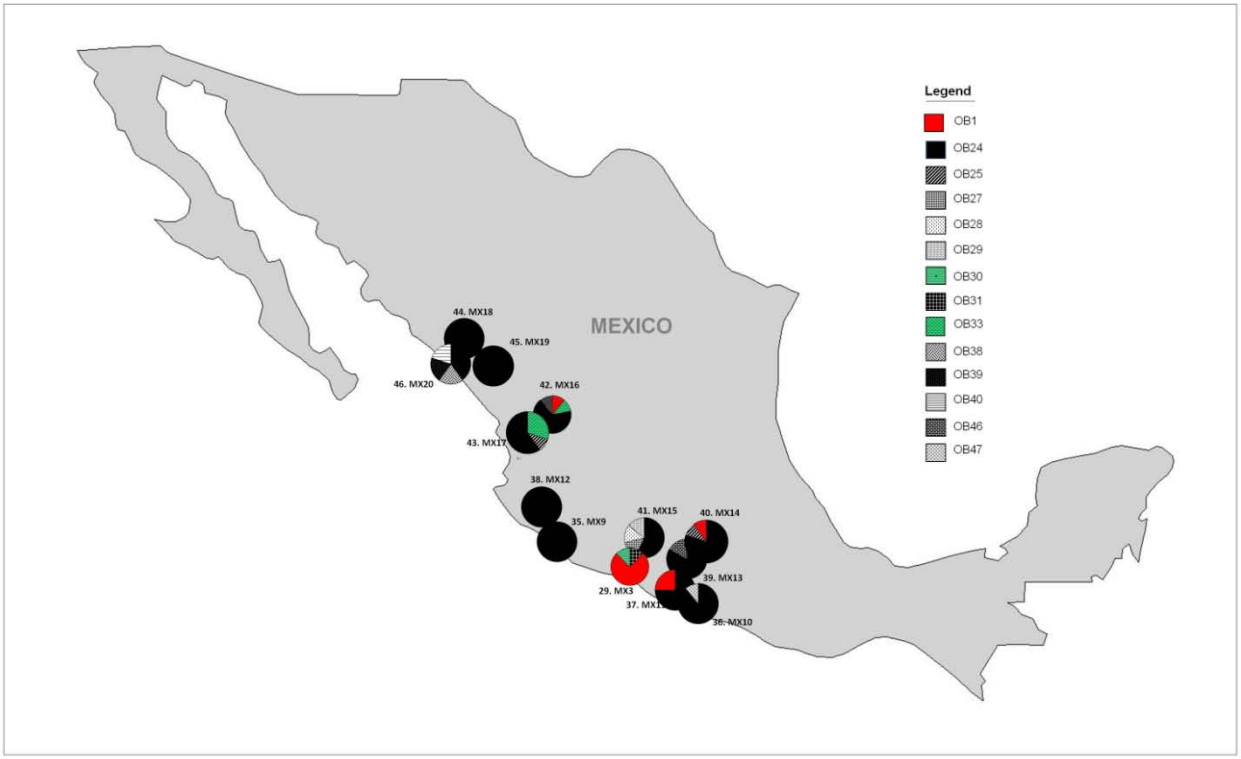


Figure 18. Geographic distribution of haplotypes and respective frequencies observed in *Anastrepha obliqua* specimens gathered in collections from the western Mexico population. Pie charts show haplotype, site number, and population code corresponding to locality in Table 22.

The four major phylogenetic clades mentioned above differ considerably (Table 24). We see that on average haplotypes from clade I vary from haplotypes in clades II, III, and IV by 16.13 bp (1.04%), 29.38 bp (1.90%), and 32.57 bp (2.10%), respectively. Clade II varies from clades III and IV by 34.61 bp (2.24%) and 34.0 bp (2.20%), respectively. And clade III varies from clade IV by 15.51 bp (1.01%). We further examine these data since the initial analyses have revealed an interesting trend that ties geography to the genetic variation seen in COI and ND6. Also, divergence estimates reveal a disparity among the haplogroups that we further address in our results of comparisons among these samples and additional closely-related taxa in Appendix C. We analyze these data for structure with AMOVA and perform pairwise (*F*st ) comparisons in order to further explore our definition of these six populations (Mesoamerica, C. America, Caribbean, western Mexico, Andean, and eastern

Brazil). In addition, we examine the historical perspective and relationships among haplotypes using nested clade analyses.

*3.3* Population structure

In order to examine the geographic delineation of populations, we began by performing a set of analyses on sequences from individuals belonging to the Mesoamerica, Caribbean, and C. America populations. Haplotypes from these populations form a monophyletic group (Fig. 15, clade I) but are geographically separated by either mountain ranges or a large body of water. In order to examine for differences among populations we conducted analyses using AMOVA and pairwise comparisons. AMOVA revealed that a total of 14.40% of the variation in mtDNA was due to differences among these three groups (Table 27). Additionally, estimates for *F*st were significant at the $P < 0.05$ level for these populations (Table 28). The overall level of differentiation for Mesoamerican, Caribbean, and C. American populations ranged from 0.153 to 0.258. We consider those distinctions defined by AMOVA and *F*st estimates along with positioning and separation of inland collections by the Sierra Madre de Chiapas mountain ranges as well as by the separation of mainland and island collections by a large body of water as substantial information to consider these three groups as separate populations.

Analysis of molecular variance (AMOVA) rejected the null hypothesis that the six studied populations are homogenous. The greater majority (63.78%) of the total variation in the mtDNA data was attributed to the differences among the groups (Table 27) when we consider six geographic populations. A significant proportion of variation was also seen in among populations (24.93%) and

within populations (11.30%). We now make our assessments of population variation and structure given our reasonably supported definitions of *A. obliqua* geographic populations.

Table 27. Partitioning of DNA variance at three hierarchical levels as revealed by analysis ofmolecular variance (AMOVA) for *A. obliqua* COI+ND6 mitochondrial DNA sequences.

| | Within populations | | Among populations within groups | | Among groups | |
|---|---|---|---|---|---|---|
| Group division | % var | $F_{ST}$ | % var | $F_{SC}$ | % var | $F_{CT}$ |
| 1. Regions (*n* = 6) | 11.30 | 0.88702 | 24.93 | 0.68812 | 63.78 | 0.63775 |
| 2. Regions (*n* = 5) | 10.90 | 0.89100 | 25.49 | 0.70047 | 63.61 | 0.63610 |
| 3. Mesoamerica vs. C. America vs. Caribbean (*n* = 3) | 68.54 | 0.82939 | 17.06 | 0.80071 | 14.40 | 0.14395 |
| 4. Mesoamerica vs. C. America (*n*=2) | 78.05 | 0.21955 | 9.03 | 0.10367 | 12.93 | 0.12928 |

*p*, estimated after 10,000 permutations; significance $p<0.00000$

Table 28. Pairwise *F*st values observed in *Anastrepha obliqua* C. American (C. Am), Caribbean (Car) and Mesoamerican (Meso) populations. Values in ***bold*** are significantly different from zero ($P<0.05$).

| | C. Am | Car | Meso |
|---|---|---|---|
| 1 | 0.00000 | | |
| 2 | ***0.25803*** | 0.00000 | |
| 3 | ***0.19383*** | ***0.15280*** | 0.00000 |

*3.3.1* Mesoamerica population

Mesoamerica collections possessed 14 haplotypes from the 43 individuals gathered from nine geographic localities (Table 29). We observed that eleven of these haplotypes were unique to a location and three were shared among collection sites (Table 30). We also observed very low *F*st values for a population with fairly wide geographic distribution of haplotypes (Table 31). The lack of structure among collections within this population was most likely due to the large presence of OB01 and other similar haplotypes that differ by an average of 2.4 bases. Pairwise comparisons showed that the moderate structure was mainly seen when collections were compared to the Agua Blanca, Tabasco (MX5) collection. Within this collection, one individual possessed the uniquely different haplotype OB24, which varied by 26 to 30 bases (1.7% - 1.9%) when compared to all other haplotypes sampled from this population. We note that OB24 was seen in higher frequency in collections from western Mexico and clusters among haplotypes belonging to clade III (Fig. 15). As a point of reference, when we discounted haplotype OB24, the remaining haplotypes differed by an average of 1.3 bases. Additional haplotypes that were shared with other populations included OB01 and OB03. OB01 was seen in the Caribbean (n=2), C. America (n=18), and western Mexico (n=9) populations, and OB03 was present in only one individual within the Mesoamerica population but was seen in 17 flies gathered from the C. America population.

Table 29. Genetic diversity of *Anastrepha obliqua* populations examined in this study. The values in the columns correspond to sample size (*n*), average number of nucleotide differences (k), segregating sites (S), nucleotide diversity ($\pi$), number of haplotypes (h), and haplotype diversity (*H*d).

| Population | *n* | k | S | $\pi$ | h | *H*d |
|---|---|---|---|---|---|---|
| 1. Mesoamerica | 43 | 2.54 | 39 | 0.00164 | 14 | 0.683 |
| 2. Central America | 69 | 14.33 | 46 | 0.00926 | 13 | 0.793 |
| 3. Caribbean | 66 | 1.78 | 13 | 0.00115 | 15 | 0.814 |
| 4. Western Mexico | 91 | 7.72 | 38 | 0.00499 | 15 | 0.482 |
| 5. Andean S. America | 26 | 16.18 | 50 | 0.01045 | 9 | 0.625 |
| 6. Eastern Brazil | 54 | 0.588 | 3 | 0.00038 | 4 | 0.532 |

Table 30. The geographic distribution of 61 haplotypes observed among the *Anastrepha obliqua* populations examined. Data is sorted by population and shows respective clade from phylogenetic reconstruction (Fig. 15). Haplotypes shared by multiple populations are in ***bold***. Population codes (e.g. DR1) correspond to Table 22. GeneBank accession numbers to COI and ND6 sequences that composed the 61 concatenated haplotypes of *A.obliqua* are shown.

| Population(s) | Haplotype | COI GeneBank no. | ND6 GeneBank no. | Clade | Origin of population (n = # of individuals) |
|---|---|---|---|---|---|
| Mesoamerica | | | | | |
| | ***OB01*** | HM545090 | HM564036 | I | BEL1(1), BEL2(1), MX1(6), MX2(4), MX4(2), MX5(1), MX6(6), MX7(2), MX8(1) |
| | ***OB03*** | HM545092 | HM564036 | Ic | MX6(1) |
| | OB04 | HM545090 | HM564055 | I | MX7(1) |
| | OB16 | HM545097 | HM564036 | I | MX2(1) |
| | OB17 | HM545098 | HM564037 | Ia | MX2(1) |
| | OB18 | HM545099 | HM564036 | If | MX1(1), MX2(1), MX4(2) |
| | OB19 | HM545090 | HM564041 | I | MX1(2), MX2(1) |
| | OB20 | HM545100 | HM564036 | I | MX2(1) |
| | ***OB24*** | HM545102 | HM564046 | III | MX5(1) |
| | OB34 | HM545108 | HM564036 | I | MX4(1) |
| | OB35 | HM545090 | HM564050 | I | MX6(1) |
| | OB50 | HM545114 | HM564036 | I | BEL1(1) |
| | OB54 | HM545117 | HM564036 | If | MX1(1) |
| | OB56 | HM545090 | HM564064 | Ia | MX5(2) |
| C. America | | | | | |
| | ***OB01*** | HM545090 | HM564036 | I | CR1(1), CR2(2), ES1(3), GT1(3), HON1(5), HON2(4) |
| | OB02 | HM545091 | HM564042 | Id | CR1(1) |
| | ***OB03*** | HM545092 | HM564036 | Ic | CR1(1), CR2(1), ES1(3), GT1(2), HON1(3), HON2(7) |
| | ***OB06*** | HM545093 | HM564062 | IV | PAN1(7), PAN2(6), PAN3(7) |
| | ***OB12*** | HM545094 | HM564037 | Ia | HON1(1) |
| | OB13 | HM545095 | HM564036 | I | ES1(2) |
| | OB14 | HM545096 | HM564039 | Ic | ES1(1) |
| | ***OB41*** | HM545109 | HM564057 | II | CR2(1) |
| | OB42 | HM545091 | HM564036 | Id | CR1(1), CR2(3) |
| | OB43 | HM545096 | HM564036 | Ic | CR2(1) |
| | OB44 | HM545091 | HM564058 | Id | CR2(1) |
| | OB45 | HM545110 | HM564036 | I | CR2(1) |
| | OB53 | HM545116 | HM564036 | Ic | HON2(1) |
| Caribbean | | | | | |
| | ***OB01*** | HM545090 | HM564036 | I | DR1(1), DR2(1) |
| | OB07 | HM545090 | HM564070 | Ia | DR1(3), DR2(2), PR2(3) |
| | OB08 | HM545090 | HM564071 | Ib | DR1(1) |
| | OB09 | HM545090 | HM564072 | Ib | BAR1(2), DR1(1), PR1(3), PR3(4), STV1(5) |
| | OB10 | HM545090 | HM564037 | Ia | DR1(1), DR2(2), JAM1(12), JAM2(8) |
| | OB11 | HM545090 | HM564038 | Ib | DR1(1) |
| | ***OB12*** | HM545094 | HM564037 | Ia | DR1(1) |
| | OB15 | HM545090 | HM564040 | Ia | JAM1(2), JAM2(1) |
| | OB21 | HM545090 | HM564043 | Ib | PR3(1) |
| | OB23 | HM545090 | HM564045 | Ib | PR3(1) |
| | OB26 | HM545094 | HM564036 | Ie | PR1(1), PR3(2) |
| | OB36 | HM545090 | HM564051 | Ib | DR2(4) |

Table 30. (cont.)

| Population(s) | Haplotype | COI GeneBank no. | ND6 GeneBank no. | Clade | Origin of population (n = # of individuals) |
|---|---|---|---|---|---|
| | OB55 | HM545090 | HM564063 | Ia | JAM1(1) |
| | OB57 | HM545090 | HM564065 | Ib | PR1(1) |
| | OB58 | HM545094 | HM564066 | Ie | PR1(1) |
| Mexico (western) | | | | | |
| | ***OB01*** | HM545090 | HM564036 | I | MX3(6), MX11(1), MX14(1), MX16(1) |
| | ***OB24*** | HM545102 | HM564046 | III | MX9(5), MX10(8), MX11(3), MX12(2), MX13(5), MX14(8), MX15(4), MX16(6), MX17(6), MX18(6), MX19(10), MX20(2) |
| | OB25 | HM545103 | HM564046 | III | MX14(1), MX17(1) |
| | OB27 | HM545104 | HM564046 | III | MX15(1) |
| | OB28 | HM545105 | HM564046 | III | MX15(1) |
| | OB29 | HM545106 | HM564046 | III | MX15(1) |
| | ***OB30*** | HM545090 | HM564047 | I | MX3(1) |
| | OB31 | HM545102 | HM564048 | III | MX3(1) |
| | OB32 | HM545102 | HM564049 | III | MX16(1) |
| | OB33 | HM545107 | HM564036 | I | MX16(1), MX17(3) |
| | OB38 | HM545102 | HM564053 | III | MX20(1) |
| | OB39 | HM545102 | HM564054 | III | MX20(1) |
| | OB40 | HM545102 | HM564056 | III | MX20(1) |
| | OB46 | HM545111 | HM564059 | III | MX13(1) |
| | OB47 | HM545102 | HM564060 | III | MX10(1) |
| S. America (Andean) | | | | | |
| | ***OB06*** | HM545093 | HM564062 | IV | COL1(7), COL3(5), COL4(4) |
| | ***OB41*** | HM545109 | HM564057 | II | COL2(1) |
| | OB51 | HM545115 | HM564062 | IV | COL3(1) |
| | OB48 | HM545112 | HM564061 | Ia | COL2(2) |
| | OB49 | HM545113 | HM564061 | Ia | COL2(1) |
| | OB52 | HM545102 | HM564062 | IV | COL4(1) |
| | OB59 | HM545118 | HM564067 | II | PER1(2) |
| | OB60 | HM545118 | HM564068 | II | PER1(1) |
| | OB61 | HM545119 | HM564069 | Ie | PER1(1) |
| E. Brazil | | | | | |
| | OB05 | HM545093 | HM564044 | IV | BRA1(7), BRA2(2)BRA3(2), BRA4(2), BRA5(1), BRA6(8), BRA7(2), BRA8(9), BRA9(1) |
| | ***OB06*** | HM545093 | HM564062 | IV | BRA3(7), BRA4(1), BRA6(1), BRA7(6) |
| | OB22 | HM545101 | HM564044 | IV | BRA4(3) |
| | OB37 | HM545093 | HM564052 | IV | BRA1(2) |

Table 31. Pairwise *F*st values observed in *Anastrepha obliqua* Mesoamerica population. Collection codes correspond to Table 22. Values in ***bold*** are significantly different from zero ($P<0.05$).

| | BEL1 | BEL2 | MX1 | MX2 | MX4 | MX5 | MX6 | MX7 | MX8 |
|---|---|---|---|---|---|---|---|---|---|
| BEL1 | 0 | | | | | | | | |
| BEL2 | -1 | 0 | | | | | | | |
| MX1 | 0.06504 | -0.82222 | 0 | | | | | | |
| MX2 | -0.08737 | -1 | -0.0546 | 0 | | | | | |
| MX4 | 0.05852 | -0.6 | 0.03567 | 0.01528 | 0 | | | | |
| MX5 | -0.16409 | -0.80392 | ***0.31214*** | ***0.2375*** | ***0.15999*** | 0 | | | |
| MX6 | 0.06323 | -1 | 0.04791 | -0.01712 | 0.18145 | ***0.26375*** | 0 | | |
| MX7 | 0.04545 | -1 | 0.01515 | -0.0856 | 0.08996 | 0.01045 | -0.01336 | 0 | |
| MX8 | -1 | 0 | -0.82222 | -1 | -0.6 | -0.80392 | -1 | -1 | 0 |

*3.3.2* Central America population

There were 13 haplotypes in the 69 individuals sampled from the nine geographic collections (Table 29) in the C. America population. Nine of the haplotypes were site specific and the remaining four were shared among locations. The presence of haplotype OB01 was substantial in this collection as well, occurring in 18 (26%) of the samples examined and was represented in all of the geographic sites sampled from this population except Panama. As previously mentioned, OB01 and OB03 were also present in samples from the Mesoamerica population. Additionally, OB41 seen in one individual from the Puntarenas Province, Costa Rica (CR2) was also present in one individual from the Andean population: Caldas, Colombia (COL2).

Only two sets of comparisons yielded significant *F*st values (Table 32). The Honduras (HON2) vs. Costa Rica (CR2) collections yielded a significant but relatively low *F*st value (0.17025). The *F*st values that were significant and very high (0.91407 - 0.98487 ) were restricted to pairwise comparisons made between the Panamanian and non-Panamanian collections. We

observed that haplotype OB06, characteristic to E. Brazil collections, was also present in the Andean population as well as in all individuals collected from Panama (Table 30). OB06 differed considerably to all other haplotypes seen in the C. America population. When compared to the remaining haplotypes, OB06 differs by 30 (1.9) to 34 (2.2%) bases. On average, haplotypes differ by 2.1 bases when OB06 is not considered. Beyond these comparisons there was no significant indication of limited geneflow among the C. America collections.

Table 32. Pairwise *F*st values observed in *Anastrepha obliqua* C. American population. Collection codes correspond to Table 22. Values in ***bold*** are significantly different from zero ($P<0.05$).

| | CR1 | PAN1 | PAN2 | PAN3 | HON1 | HON2 | ES1 | CR2 | GT1 |
|---|---|---|---|---|---|---|---|---|---|
| CR1 | 0 | | | | | | | | |
| PAN1 | ***0.97633*** | 0 | | | | | | | |
| PAN2 | ***0.97343*** | 0 | 0 | | | | | | |
| PAN3 | ***0.97633*** | 0 | 0 | 0 | | | | | |
| HON1 | 0.03226 | ***0.974*** | ***0.97204*** | ***0.974*** | 0 | | | | |
| HON2 | 0.24503 | ***0.97741*** | ***0.97602*** | ***0.97741*** | 0.09362 | 0 | | | |
| ES1 | 0.04927 | ***0.97042*** | ***0.9682*** | ***0.97042*** | -0.05 | 0.06658 | 0 | | |
| CR2 | -0.14593 | ***0.91948*** | ***0.91407*** | ***0.91948*** | 0.02052 | ***0.17025*** | 0.0409 | 0 | |
| GT1 | -0.00355 | ***0.98487*** | ***0.98321*** | ***0.98487*** | -0.15064 | -0.01249 | -0.12417 | -0.03316 | 0 |

*3.3.3* Caribbean population

Fifteen haplotypes were recovered from the 66 samples gathered from the nine geographic sites in the Caribbean. Within this population we see nine private and six shared haplotypes (Table 30). The two most common haplotypes sampled in this clade were OB09 and OB10 which differ by only a single mutation. These two haplotypes were found solely in the Caribbean. Haplotype OB09 was observed in 15 (23%) of the collections but found exclusively in eastern Antillean sites. Haplotype OB10 was seen in 23 (35%) of the collections within the Caribbean population, but was most common to western Antillean collections. We see moderate to high *F*st values for this

population supporting the notion of limited dispersal of haplotypes collected from these islands (Table 33). Interestingly, two individuals, each from a different site in the Dominican Republic (DR1 & DR2), showed haplotype OB01 which was recovered in more abundance from mainland collections. OB01 is very similar to the remaining haplotypes seen in Caribbean collections and differs by only one to three bases. We discuss further the importance of OB01 to this region in our phylogeographic analyses.

Table 33. Pairwise *F*st values observed in *Anastrepha obliqua* Caribbean population. Collection codes correspond to Table 22. Values in ***bold*** are significantly different from zero ($P<0.05$).

| | BAR1 | DR1 | DR2 | JAM1 | JAM2 | PR1 | PR2 | PR3 | STV1 |
|---|---|---|---|---|---|---|---|---|---|
| BAR1 | 0 | | | | | | | | |
| DR1 | 0.11995 | 0 | | | | | | | |
| DR2 | 0.09634 | 0.02703 | 0 | | | | | | |
| JAM1 | ***0.84385*** | ***0.27302*** | ***0.4098*** | 0 | | | | | |
| JAM2 | ***0.9081*** | ***0.21667*** | ***0.35938*** | -0.07946 | 0 | | | | |
| PR1 | -0.16418 | 0.15793 | 0.16878 | ***0.67742*** | ***0.65299*** | 0 | | | |
| PR2 | 1 | 0.17396 | 0.3645 | ***0.74138*** | ***0.84549*** | ***0.65022*** | 0 | | |
| PR3 | -0.22088 | ***0.20121*** | ***0.19813*** | ***0.69516*** | ***0.67755*** | -0.08615 | ***0.6958*** | 0 | |
| STV1 | 0 | ***0.30245*** | ***0.28264*** | ***0.86999*** | ***0.93069*** | 0.09513 | ***1*** | 0.01809 | 0 |

*3.3.4* Western Mexico population

Our analyses of the western Mexico population revealed 15 haplotypes present in 91 of the specimens examined from 13 geographic sites (Table 29). Haplotype OB24, the most common haplotype, was seen in all thirteen geographic collections and in 65 (72%) of the individuals sampled from this region. OB24 was distributed evenly among northern and southern regions within western Mexico collections. The pairwise comparisons for this population yielded very low *F*st values except for those comparisons made against sequences recovered from Michipan, Morelos, Mx (MX3) in southern Mexico (Table 34). The significant and high estimates are most likely due to the

presence of haplotypes OB01 and OB30 in Michipan which differed to other haplotypes by as much as 27 (1.7%) to 30 (1.9%) bases. Of the eight flies examined from Michipan, six showed the OB01 haplotype and 1 showed OB30. As seen earlier, OB01 is characteristic of haplotypes occurring in a greater abundance in Mesoamerica and C. America. Three additional flies, each from a different locality (MX11, 14, 16) within the western Mexico population also showed OB01. Only one other comparison (MX11 vs. MX19) yielded a significant *F*st value. OB01 was seen in 1 of the 4 individuals gathered in La Palma, Guerrero (MX11) and may have contributed to the significance. While both OB01 and OB30 share a high degree of similarity to each other and to OB33 (present in MX16 & 17) all three differ substantially from the remaining haplotypes recovered in western Mexico. Additionally, these three haplotypes cluster among those pertaining to clade I (Fig. 15) which is populated mostly by sequences from individuals gathered in Mesoamerica, C. America, and Caribbean. The remaining haplotypes for the western Mexico population form the monophyletic clade III. The overall low diversity estimates, low *F*st values for a majority of the collections, and the presence of one abundant haplotype would thus suggest that this population is highly homogeneous.

Table 34. Pairwise Fst values observed in *Anastrepha obliqua* western Mexico population. Collection codes correspond to Table 22. Values in ***bold*** are significantly different from zero ($P<0.05$).

| | MX3 | MX9 | MX10 | MX11 | MX12 | MX13 | MX14 | MX15 | MX16 | MX17 | MX18 | MX19 | MX20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MX03 | 0 | | | | | | | | | | | | |
| MX09 | ***0.82476*** | 0 | | | | | | | | | | | |
| MX10 | ***0.86285*** | -0.07784 | 0 | | | | | | | | | | |
| MX11 | ***0.50734*** | 0.0625 | 0.20777 | 0 | | | | | | | | | |
| MX12 | ***0.76351*** | 0 | -0.3211 | -0.26316 | 0 | | | | | | | | |
| MX13 | ***0.82673*** | -0.03448 | 0.03271 | 0.08725 | -0.30435 | 0 | | | | | | | |
| MX14 | ***0.71917*** | -0.07782 | -0.00329 | -0.10737 | -0.31606 | -0.05149 | 0 | | | | | | |
| MX15 | ***0.83569*** | -0.05528 | 0.02138 | 0.12994 | -0.3125 | -0.00322 | -0.02421 | 0 | | | | | |
| MX16 | ***0.54138*** | 0.0298 | 0.11864 | -0.21465 | -0.19882 | 0.05051 | -0.05755 | 0.08056 | 0 | | | | |
| MX17 | ***0.4322*** | 0.11866 | 0.20825 | -0.19312 | -0.08819 | 0.13803 | 0.02442 | 0.16733 | -0.0943 | 0 | | | |
| MX18 | ***0.83781*** | 0 | -0.05109 | 0.11111 | 0 | 0 | -0.05213 | -0.02439 | 0.05919 | 0.14763 | 0 | | |
| MX19 | ***0.87452*** | 0 | 0.01235 | ***0.24528*** | 0 | 0.09091 | 0.00766 | 0.05405 | 0.13669 | 0.22727 | 0 | 0 | |
| MX20 | ***0.81113*** | 0 | 0.07909 | 0.05186 | -0.29032 | 0.01081 | -0.054 | 0.01072 | 0.03516 | 0.11985 | 0.04 | 0.14894 | 0 |

*3.3.5* Andean S. America population

Twenty-six specimens from five geographic localities were analyzed for this population. We recovered nine haplotypes from these individuals (Table 29). Eight of these nine haplotypes were site specific. Haplotype OB06 was seen in 16 (62%) specimens from three collection sites in Colombia. The *F*st values were very high (Table 35) and significance was most likely due to the disparity of haplotypes seen in the collections. The three haplogroups represented within the samples examined here differ by 15 (0.9%) to 35 (2.3%) bases. We also note that the nine haplotypes segregate into three of the four major clades seen in our phylogenetic reconstruction (Fig. 15). Haplotypes OB48, OB49, [co-occurring at one site in Colombia (COL2)] and OB61

[originating in Peru (PER1)] all cluster among haplotypes belonging to clade I. Haplotype OB06, sharing a high degree of similarity to OB51 and OB52, was present at high frequencies in Panama and E. Brazil populations and clusters among haplotypes within clade IV. Only one other haplotype was shared among different populations. Haplotype OB41, showing a similarity to OB59 and OB60, was present in samples from Colombia and C. America as previously mentioned. These last three haplotypes form clade II.

Table 35. Pairwise *F*st values observed in *Anastrepha obliqua* Andean S. America population. Collection codes correspond to Table 1. Values in ***bold*** are significantly different from zero ($P<0.05$).

| | COL1 | COL2 | COL3 | COL4 | PER1 |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | 0.00000 | | | | |
| 2 | ***0.91704*** | 0.00000 | | | |
| 3 | 0.02778 | ***0.90149*** | 0.00000 | | |
| 4 | 0.07285 | ***0.85393*** | -0.02321 | 0.00000 | |
| 5 | ***0.90326*** | 0.34579 | ***0.88642*** | ***0.83871*** | 0.00000 |

*3.3.6* Eastern Brazil population

Among this population we observed four haplotypes from the 54 individuals examined. Only two haplotypes (OB22 & OB37) were site specific. Haplotype OB05 was seen in all collections from E. Brazil, in 34 (63%) of the specimens. OB06 was found in 15 (28%) of the flies examined in E. Brazil. This haplotype was also seen in the Andean and Central America populations as mentioned previously. No other haplotype from this population was shared. Pairwise comparisons

revealed moderate to high *F*st values for the nine sampling sites (Table 36). A total of eleven significant values were observed.

Table 36. Pairwise *F*st values observed in *Anastrepha obliqua* E. Brazil population. Collection codes correspond to Table 22. Values in ***bold*** are significantly different from zero ($P<0.05$).

| | BRA1 | BRA2 | BRA3 | BRA4 | BRA5 | BRA6 | BRA7 | BRA8 | BRA9 |
|---|---|---|---|---|---|---|---|---|---|
| BRA1 | 0 | | | | | | | | |
| BRA2 | -0.19431 | 0 | | | | | | | |
| BRA3 | ***0.61111*** | 0.59091 | 0 | | | | | | |
| BRA4 | ***0.28974*** | 0.02326 | ***0.47456*** | 0 | | | | | |
| BRA5 | -0.75 | 0 | 0.5 | -0.4 | 0 | | | | |
| BRA6 | 0.08333 | -0.3211 | ***0.57328*** | ***0.27095*** | -1 | 0 | | | |
| BRA7 | ***0.58051*** | 0.54286 | -0.13143 | ***0.4376*** | 0.42857 | ***0.53777*** | 0 | | |
| BRA8 | 0.125 | 0 | ***0.75*** | ***0.38835*** | 0 | 0 | ***0.72932*** | 0 | |
| BRA9 | -0.75 | 0 | 0.5 | -0.4 | 0 | -1 | 0.42857 | 0 | 0 |

*3.4* Inferred history using summary statistics and phylogeographic analyses

In order to examine the evolutionary history of the haplotypes and explore possible processes for structure, we utilize demographic statistics and NCPA. The hypothesis of random distribution of haplotypes was tested and the sources for geographic patterns seen were inferred through the inference key in ANeCA (Table 37). We were able to evaluate the evolutionary forces that shaped the wide-scale geographic patterns seen among the regions inhabited by the six populations whose haplotypes cluster within the four observed phylogenetic clades (Fig. 15). The overall patterns seen in the parsimony network are concordant with phylogenetic results and as expected, the haplotypes group into two networks separated by 17 mutational steps. Network NC-A and NC-B are presented in Figs. 19 & 20, respectively. The NC-A groups appeared as sister groups as did the NC-B groups as suggested by phylogenetic analyses.

Table 37. Significant demographic inferences for Mesoamerica, C. America, and Caribbean nested clade NC-A (Fig. 19) based upon the results from nested clade phylogeographic analysis performed with GeoDis V. 2.5 and the 2005 inference key.

| Clade | Inference Chain | Inferred Pattern |
|---|---|---|
| Clade 1-5 | 1-2-11-12 NO | Contiguous range expansion. |
| Clade 2-3 | 1-2-3-4 NO | Restricted gene flow with isolation by distance |
| Clade 4-2 | 1-2-11-17-4 NO | Restricted gene flow with isolation by distance |

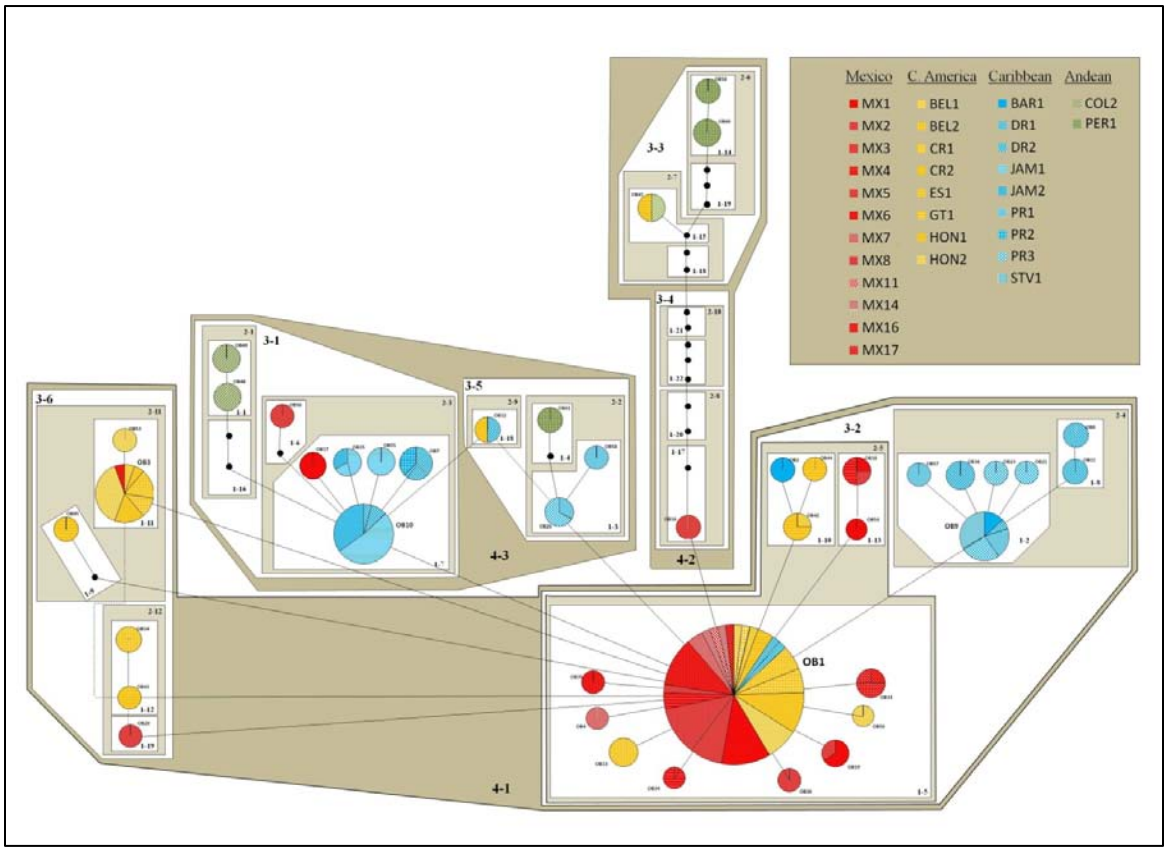


Figure 19. Cladogram NC-A. *Anastrepha obliqua* haplotype network with nested clades showing relationships among haplotypes gathered from Mexico, Central America, Caribbean, and Andean S. America regions based on statistical parsimony. Size of the pie charts is proportional to haplotype frequency in each population (e.g. MX1, see Table 22). Color scheme is associated with geographic region. Each line between haplotypes indicates one mutational step. Black ovals without letter names are haplotypes not sampled but which are necessary to connect sampled haplotypes. Dashed lines connecting haplotypes represent broken connections and is further explained in text.

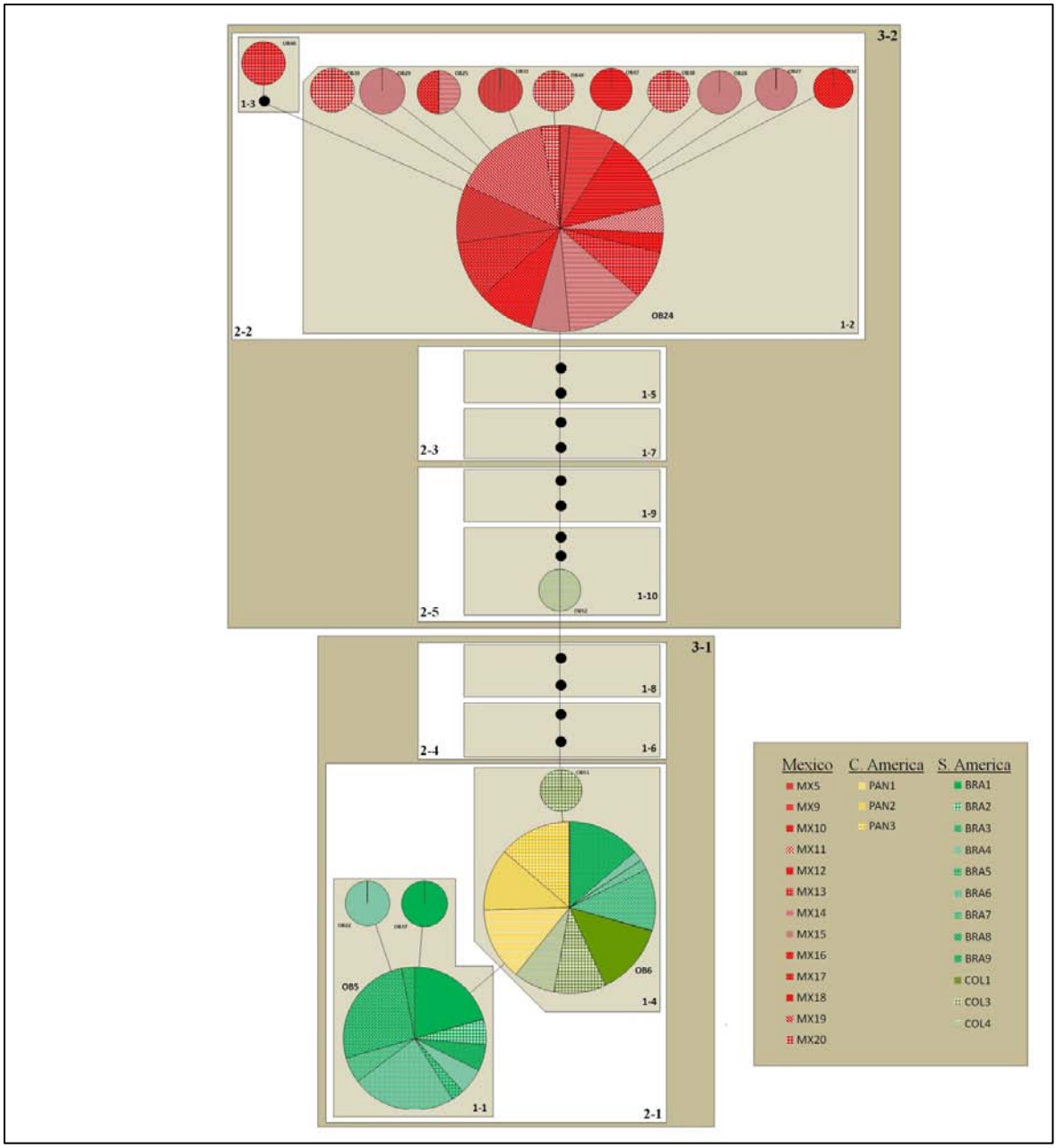


Figure 20. Cladogram NC-B. *Anastrepha obliqua* haplotype network with nested clades showing relationships among haplotypes collected in Mexico, Central America, and South America regions based on statistical parsimony. Size of the pie charts is proportional to haplotype frequency in each collection (e.g. MX1, see Table 22). Color scheme is associated with geographic region. Each line between haplotypes indicates one mutational step. Black ovals without letter names are haplotypes not sampled but which are necessary to connect sampled haplotypes.

*3.4.1* Population summary statistics

We estimated the haplotype and nucleotide diversity for collections in Mesoamerica to be 0.683 and 0.00164 (Table 29), respectively. Sequences for individuals from the 14 collections showed 39 polymorphic sites. Within Mesoamerica we see an abundance of low frequency haplotypes (Table 30) that are closely related (Fig. 15) to the common haplotype (OB01). These characteristic are typical for species that have undergone a bottleneck event and subsequent expansion. We investigated this hypothesis by testing for departure from neutrality via demographic analyses and examining the mismatch distribution of these haplotypes. Demographic analyses for this population revealed significant values for Tajima's $D$ (-2.52, $P < 0.001$), Fu and Li's $D$ (-4.47, $P < 0.02$), and Fu and Li's $F$ (-4.51, $P < 0.02$) tests. Sequences within this population provided a mismatch distribution that is for the most part unimodal (Fig. 21) and would be consistent with the distribution under a model of sudden population expansion (Slatkin and Hudson, 1991), further supporting demographic results. One small peak at approximately 30 bases was the mismatch produced by the very different haplotype (OB24) discussed earlier.

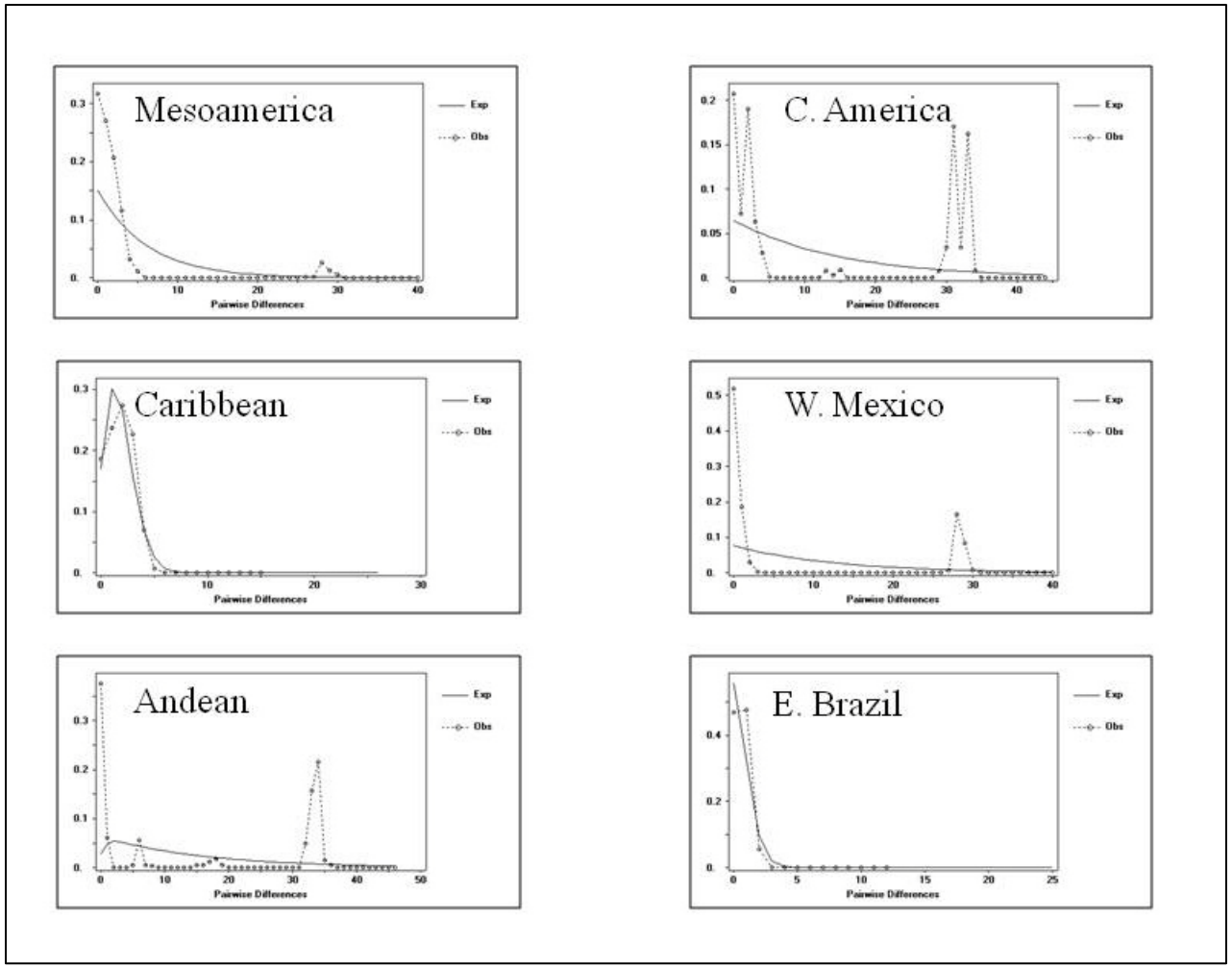


Figure 21 Results of mitochondrial mismatch distributions of the six geographic populations of *Anastrepha obliqua* examined for this study. The expected frequency is based on a population growth-decline model determined using the DNASP v4.0 program (Rozas and Rozas, 1997) and is represented by a continuous line. The observed frequency is represented by a dotted line.

In C. America, the haplotype and nucleotide diversities we observed were 0.793 and 0.00926, respectively (Table 29). Here, however, the results of the demographic analyses were not significant at $P$ =0.05 for any test conducted. The mismatch distribution for this population shows multiple peaks and was thus not consistent with a distribution that represents a population experiencing expansion thereby in agreement with the demographic analyses. The mismatch distribution reveals major peaks along those comparisons made with the very different haplotypes observed in Panama, previously discussed.

The Caribbean population enjoys the highest haplotype diversity among the six populations examined for the West Indian fruit fly. Here we see that the haplotype and nucleotide diversities were 0.814 and 0.00115, respectively (Table 29). All demographic statistics resulted in estimates that were not signification at the $P$ =0.05 level and thus do not support a recent population expansion. We did, however, observe a unimodal mismatch distribution which would be consistent with a population having experienced a recent bottleneck event.

The haplotype and nucleotide diversities were estimated as 0.482 and 0.00499, respectively in western Mexico. The sequences yielded 38 polymorphic sites. The demographic analyses we conducted were not significant at $P$ =0.05 for any test. Additionally, we see multiple peaks in the mismatch distributions for this population that would not be consistent with a population that has experienced a recent bottleneck event. These peaks show the presence of substantially different haplotypes as previously discussed.

Collections within the Andean population showed that haplotype and nucleotide diversities were 0.625 and 0.01045, respectively (Table 29). Here, sequences showed 50 polymorphic sites detected. The values obtained through demographic analyses were not significant at $P$ =0.05 for any

test conducted and thus do not support models of population experiencing a recent expansion. The mismatch distributions concur showing multiple peaks.

The eastern Brazil population showed that haplotype and nucleotide diversities were 0.532 and 0.00038 (Table 29), respectively. The sequences yielded 3 polymorphic sites. Here, the demographic analyses were not significant at $P$ =0.05 for any test conducted and thus do not support a recent bottleneck event and population expansion. The mismatch distribution for this population shows a single peak and would be consistent with the distribution under a model of sudden population expansion.

*3.4.2* Phylogeographic analyses

NCPA methods revealed three loops for haplotypes from Mesoamerica, C. America, and the Caribbean (Fig. 19). Significant geographical association at different nesting levels was observed exclusively for haplotypes represented in cladogram NC-A (Fig. 19). Contiguous range expansion in clade 1-5 and restricted gene flow with isolation by distance (clades 2-3 and 4-2) were inferred at different nesting levels within this cladogram.

Range expansion was inferred in clade 1-5, which contains the putative ancestral haplotype OB01. Haplotype OB01 was determined as the ancestral haplotype based upon network root probabilities estimated by TCS whereby each haplotype is given a probability of being the ancestral haplotype. This haplotype was distributed among 16/16 mainland collections with two specimens showing this haplotype in the Caribbean (DR1, 2). Within clade 1-5 we see that the eight additional and low frequency haplotypes were closely related to the common haplotype (OB01) and arranged in

a star-like fashion found within clade 1-5, typical of species that have undergone a bottleneck event and subsequent expansion.

NCPA methods inferred a restricted gene flow and isolation by distance pattern for clade 2-3 (Fig. 19). This clade is composed of haplotypes that are predominantly found in the Caribbean. In order to further investigate the inferences by NCPA, we used the Mantel permutation procedure to test for genetic isolation by distance (Mantel, 1967). Mantel test comparisons were statistically significant for clade 2-3. The Mantel test showed a significant positive correlation between geographical and genetic distance for this clade ($r = 0.290453$, $p < 0.019$).

Eighteen haplotypes and three nesting levels were observed in cladogram NC-B (Fig. 20). NCPA showed no significant association within any of the clades in NC-B. We did notice that the topological arrangement of haplotypes surrounding a more ancestral haplotype (OB24) occurring in this region support the occurrence of a recent bottleneck event and subsequent population expansion as mentioned earlier.

## 4. DISCUSSION

This study provides the first substantial look into the analysis of West Indian fruit fly mtDNA diversity and reveals information on the genetic structure of populations throughout its geographic range. The two mtDNA regions used in this study provide sufficient data to distinguish among geographic clades and suggest historical processes. Our interpretation of the evolutionary history presented here is based upon rigorous analyses of informative mitochondrial regions (Simon et al., 2006, 1994; Folmer et al., 1994) that permit robust conclusions. Our results highlight the importance of historical biogeography in forming the current geographic distribution of these six

populations that may ultimately lead to speciation (Avise, 2000, 1994; Hewitt, 1996). We discuss our findings and present our interpretation of the current status of *Anastrepha obliqua*.

The phylogenies from the combined datasets suggest four distinct clades within *A. obliqua*. These four clades can be summarized as containing collections from Mesoamerica, C. America, and Caribbean (clade I), Andean S. America (clade II), western Mexico (clade III), and South America and Panama (clade IV) regions. The four lineages are not entirely allopatric since we see overlap by populations represented in these clades. Overlap is most likely due to either the natural distribution of a population or from introductions among populations recovered through our sampling. This is a pest insect infesting cultivated fruits, so human-mediated dispersal is certainly possible.

The magnitude of the genetic divergence among *A. obliqua* clades was unexpected based upon previous studies of the genus (McPheron et al., 1999; Smith-Caldas et al., 2001) using other mtDNA markers. To examine the relatively large disparity among clades we sequenced a number of individuals from this genus with the COI+ND6 markers used in this study. Included in our comparison are various species from within the *fraterculus* taxonomic group, which contains *A. obliqua*. Five *Anastrepha* spp. cluster among the four phylogenetic clades. We describe our investigation of the substantial variation in relation to other species within the *fraterculus* in Appendix C. We do mention that by and large mtDNA regions are not useful in discriminate among closely-related taxa so some caution in these results is warranted.

Our investigation of population structure within *A. obliqua* supports six distinct populations: Mesoamerica, C. America, Caribbean, western Mexico, Andean S. America, and S. America. Geographically defined populations and a large proportion of the haplotypes is regionally and/or sample site specific. We must, however, mention that our delineation of populations according to

broad geography is merely one view and other spatial, ecological, and temporal characters could be important for defining populations It may be that population or taxonomic boundaries are restricted to interactions between flies with closely related genotypes (for example clusters I-IV) that perhaps represent a more biologically compatible set of flies. Further work is needed to evaluate the level of interaction among the genetic types seen in distinct haplogroups that vary considerably and are either in close proximity or co-located.

Our use of parsimony analysis clusters individuals from Mesoamerica, C. America, and Caribbean populations within close proximity and so discussion here will also reflect this association. Within the Mesoamerica and C. America populations, we see the predominance of one haplotype, OB01 that was recovered from 83% of the sites sampled. One possible explanation for this high frequency is that this haplotype represents a more invasive lineage of *A. obliqua*. OB01 having the largest presence would suggest this is the more ancestral haplotype (supported by NCPA) and has had more time to infiltrate favorable environments in this region. Alternatively, the predominance of OB01 could represent effects of genetic drift or population expansion of a lineage of related flies with a shared mtDNA haplotype into previously unoccupied niches. If the movement of flies with the OB01 haplotype is the result of human activity, then it is possible that higher propagule pressure (resulting from multiple introductions from the same source population) has increased its prevalence in the region. Further work is needed to arrive at a more informed explanation to these hypotheses.

Also, among those haplotypes characteristic to C. America we see the large presence of OB06 in Panama. This considerably different haplotype (to others recovered in C. America) was found in collections that were made in Panama during the 2006 season. Additional samplings are

warranted and may help reveal the movement of this and other haplotypes over time in this area. This haplotype was also seen in Colombia and found in greater abundance in eastern Brazil. It is possible that the OB06 flies in Panama represent evidence of an invasive population from S. America that has widened its distribution to C. America.

We were able to obtain further insights into the phylogeographic patterns in the Mesoamerica and C. America populations with nested clade analysis. First, a contiguous range expansion at low nesting levels for haplotypes seen in mostly southern Mexico and Central American collections (NCPA A, 1-5, Fig. 19) represents an event in the evolutionary history of this *A. obliqua* population that may have direct ties to human intervention. This was supported for Mesoamerica by population demographic analyses, unimodal mismatch distributions (when we consider that a second peak is represented by individuals from the western Mexico population), and the general topological nature in the phylogenetic reconstruction of haplotypes. The demographic analyses of the C. America population, however, did not support the NCPA inferences of a bottleneck event with population expansion.

Additional haplotypes occurring nearby in the Caribbean differ only slightly to haplotypes from Mesoamerica and C. America populations. Also, the genetic diversity among island populations is slightly higher than in mainland, suggesting that the origin of this fly may be centered in the Caribbean, most likely in the Greater Antilles based on haplotype diversity estimates. This scenario is not, however, supported by NCPA as is further explained below. Additional sampling is needed in the Lesser Antilles and Hispaniola and throughout mainland populations to test this hypothesis. We calculate that individual fruit fly populations established within the Caribbean did so approximately 24Ka based on coalescent analysis, too early for Paleo-Indians to have aided

(Dillehay, 1997) in transport. Over-water colonization by animals and plants to the Caribbean has occurred (Ricklefs & Birmingham, 2001; Hedges et al., 1992), and dispersal in form of a flotsam as discharge by outflow of continental rivers (Ricklefs & Birmingham, 2008) may have been responsible for a large percentage of ancestral migrants to this region. It may be that the more recent introductions characteristic of current haplotypes seen in the Caribbean were introduced via water currents as mentioned above or through the establishment of human cultures in this region, which dates as a far back as 6Ka (Wilson, 2001, 1997; Rouse, 1992, 1989), and the more recent arrival of Europeans at the end of the 15$^{th}$ century. Again, differences among the island haplotypes are low but distribution of the two more common haplotypes (OB09 and OB10) is polarized suggesting either two different introductions into the island regions followed by structure due to limited gene flow or perhaps that our sampling does not permit us to see a more homogeneous distribution.

NCPA suggests restricted gene flow with isolation by distance of higher nesting levels shown by individuals from southern Mexico and islands of the Greater Antilles (NCPA A, 2-3, Fig. 19) and southern Mexico, Central America, and Andean South America (NCPA A, 4-2, Fig. 19). NCPA inferences for clade 2-3 were supported by Mantel test results. The NCPA inferences for clade 4-2 were not supported by the Mantel tests most likely due to the small sample size. The information gathered from NCPA nevertheless provides one possible interpretation of the historical distribution and diversification of this group inhabiting the Caribbean where haplotypes OB09 and OB10 may have arisen from the more ancestral haplotype (OB01). The region may have experienced two separate introductions by OB01 and then disseminated separately among the Lesser and Greater Antilles as suggested by NCPA. NCPA also suggests that haplotype OB10 is an ancestral haplotype to various mainland collections. Haplotypes OB56 and OB17, occurring in southern Mexico collections (MX2 and MX5), and haplotypes OB49 and OB48, found in Colombian collections, may

have originated from ancestral Caribbean haplotypes as a reintroduction to mainland areas. This pattern was seen again within clade 2-2 where haplotype OB61 from Peru may have evolved from Caribbean haplotype OB26.

Haplotypes recovered from the Andean S. America population cluster among three different phylogenetic clades and represent substantially different genetic types, some co-occurring among sites sampled. Thus, diversity estimates and population structure were justifiably high. This area may represent a hub or intersection of distinct haplotypes and lineages. Haplotypes having a high degree of similarity to Mesoamerica and the Caribbean and shared among C. America and eastern Brazil populations would support this scenario. We have insufficient information here to properly determine whether an individual should be considered as an introduction and may be part of the natural distribution for this or other populations, further studies are needed.

NCPA (NC-A, Fig. 19), suggests that haplotypes within clade 3-3 belonging to collections in Costa Rica and Peru evolved from the ancestral haplotype OB01 via haplotype OB16 occurring in collection MX2 from the southern Mexican state of Campeche. This would suggest that the expansion of populations led to a diversification in a southerly direction towards the northern areas of S. America. Supplemental sampling within this region may provide additional information regarding the historical perspective, diversity, and distribution of haplotypes for this and other populations. We see from coalescent analysis that the estimated time of divergence of clade I and clade II (Fig. 15) was approximately 58Ka, however these estimates should be considered with caution given the relatively small sample size for Andean collections represented in clade II.

Collections from the western Mexico population yielded a considerably lower haplotype and nucleotide diversity when compared to other populations examined in this study. The presence of

one abundant haplotype (OB24) results in a high degree of homogeneity among collections. The remaining low frequency, closely-related haplotypes occurred in both of the northern and southern regions of western Mexico. Demographic statistics and mismatch distribution do not support a recent bottleneck event in the history of this fly yet the arrangement of haplotypes sampled from this region suggests a recent population expansion. If so then this may have resulted from the introduction of mango, a primary host for this fly, to the region (Baker et al., 1944). It may be that members from this population were persisting in low frequencies and began to flourish after additional host material was introduced into the region and habitats in western Mexico became more suitable for expansion. We do note that this population was quite distinct from others identified in this study. Mesoamerica, C. America, and Caribbean populations when compared to western Mexico differed by an average of approximately 33 bases. Andean and western Mexico populations differed by an average of 34 bases. Differences were less between the eastern Brazil and western Mexico populations where average difference among these two was approximately 15 bases (Table 24). A few haplotypes were site specific (which could be a sample size phenomenon) but we did not see any distribution pattern or significant association with geography for this population. Coalescent estimates suggest that clades III (western Mexico) and IV (E. Brazil/Panama) may have diverged approximately 49Ka. The lack of genetic diversity shown by individuals within the western Mexico population may be due to the isolation provided by the physical barriers imposed to the east by the Sierra Madre Occidental and Oriental, trans-Mexican volcanic belt as well as the shifts in habitat due to recent climate change (Axelrod 1979; Graham 1999). Recent ecological changes in this region are due to expansion of Sonoran vegetation that has resulted in fragmentation to woodlands and dry forests over the last 50,000 years (Axelrod 1979) characteristically inhabited by hosts of this species. The habitats have fragmented thus limiting gene flow among populations, reducing their diversities.

The diversity estimates were also low for collections from eastern Brazil. Only four haplotypes were seen in these samples and two haplotypes (OB05 & OB06) were seen in 91% of the flies. We did see evidence of population structure however, the estimates were moderate. We observed that haplotype OB05 occurs only in Brazilian collections, and haplotype OB06 was shared among the Andean and C. America population. We do note that our sampling represents collections made in the late 1990s. Additional sampling in these areas would provide a view of a more current distribution and diversity of this fly in eastern Brazil.

Results from NCPA (Fig. 20) suggest that the ancestral haplotype is OB05. NCPA results also suggest that additional haplotypes occurring within the vicinity of OB05 such as OB06 and haplotypes seen in the western Mexico population evolved from OB05. It is difficult to accurately determine the origin of OB06 given the lack of sampling in some areas within S. America. With caution, we provide three possible scenarios given the current sampling. Haplotype OB06, differing by one base from OB05 may have evolved near the eastern Brazil region then migrated recently to northern S. America and Panama or OB06 came to being recently in northern areas of S. America and Central America from introduced OB05 haplotypes or OB06 originated in other areas that have not been sampled yet. Given the absence of OB05 in northern S. America, the first explanation seems the most parsimonious. NCPA also suggests that collections within clade III and IV are closely related and have historical ties; haplotypes in western Mexico may have evolved in S. America or Panama then established in western Mexico. The lack of long distance dispersal, characteristic of this fly, would hamper gene flow over great distances and elevations allowing for genetic divergence by separated populations.

Of particular importance to management of this species and to regulatory agencies is the

level of genetic variation among populations this study revealed. The considerable genetic difference observed may translate to variation in pestiferous behavior. The disparity among populations may even translate or have a correlation to variation in characteristics key to managing this pest such as host preference and mating behavior. Aluja and Birke (1993) reported that mating time was during the morning hours for this species in C. America while A. Malavasi (unpublished) has determined that the mating activities from this species in Brazil occurs in the afternoon. The current management strategies for this fly include the sterile insect technique (SIT) which relies greatly on the competitiveness of the mass-reared flies (Hernandez et al., 2007). Variation in mating behavior and genotype could translate into a mass-reared fly whose mating schedule is not synchronized or fit to compete with the feral population. It then becomes critically important to match according to genetic types and thus more information is needed to assess the variation seen among the populations we identified.

The volume of geographic sites sampled and sequences yielded from this study are substantial. The potential for this information having a major impact in the development of diagnostic tools for this and other closely related species is considerably high given the 44 site-specific haplotypes and numerous haplotypes that are unique to geographic regions. Use of informative mtDNA regions such as these have been the foundation for various studies that result in the development of genetic markers that can be used to discriminate among geographic collections (Bonizzoni et al., 2000; Gasparich et al., 1994; McPheron et al., 1994). Additional work is needed in order to properly assess the biology behind these lineages. At this time there are no morphological analyses that can correlate the distinctions seen in this study. An in-depth comprehensive review of discriminating taxonomic characters from these groups plus additional un-linked loci is needed and may provide further evidence as to the identity of individuals within these six populations. This

information would test the hypotheses of long-term isolation of these groups within the regions and may provide further clarification into the history of this fly.

**ACKNOWLEDGEMENTS**

We thank Roxanne E. Garza-Farris, Rosita Deleon, Juan D. Vasquez, Lisa A. Ledezma, and Oscar Obregon for providing DNA extracts from the USDA-APHIS-CPHST Pest Detection, Diagnostic, and Management Laboratory in Edinburg, TX, Samuel Gum and Oralia Ruiz for technical assistance at Penn State University, Deborah S. Grove (Director of the Genomics Core Facility), Ashley Price, and Sheila Plock at The Pennsylvania State University-Huck Institute Nucleic Acid Facility. Funding for this study came from the USDA-APHIS-CPHST, PIC#M6M02, USDA/Penn State cooperative agreement 07-8100-1051-CA, and the Alfred P. Sloan Foundation.

## REFERENCES


Akaike, H., 1974. New look at statistical-model identification. IEEE Transactions on Automatic Control AC19, 716-723.

Alberti, A.C., Confalonieri, V.A., Zandomeni, R.O., Vilardi, J.C., 2008. Phylogeographic studies on natural populations of the South American fruit fly, Anastrepha fraterculus (Diptera: Tephritidae). Genetica 132, 1-8.

Alberti, A.C., Rodriguero, M.S., Gomez, C. P, Saidman, B.O., Vilardi, J.C., 2002. Evidence indicating that a Argentine populations of *Anastrepha fraterculus* (Diptera: Tephritidae) belong to a single biological species. Ann Entomol Soc Am 95(4), 505–512.

Alberti, A.C., Calcagno, G., Saidman, B.O., Vilardi, J.C., 1999. Analysis of the genetic structure of a natural population of *Anastrepha fraterculus* (Diptera: Tephritidae). Ann Entomol Soc Am 92(5), 731–736.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403-410.

Aluja, M., Diaz-Fleischer, F., 2006. Foraging behavior of *Anastrepha ludens*, *A. obliqua*, and *A. serpentina* in response to feces extracts containing host marking pheromone. Journal of Chemical Ecology 32, 367-389.

Aluja, M., Celedonio-Hurtado, H., Liedo, P., Cabrera, M., Castillo, F., Guill, J., Rios, E., 1996. Seasonal population fluctuations and ecological implications for management of Anastrepha fruit flies (Diptera: Tephritidae) in commercial mango orchards in southern Mexico. Journal of Economic Entomology 89, 654-667.

Aluja, M., 1994. Bionomics and management of *Anastrepha*. Annual Review of Entomology 39, 155-178.

Aluja, M., Birke, B., 1993. Habitat use by adults of *Anastrepha obliqua* (Diptera: Tephritidae) in a mixed mango and tropical plum orchard. Ann. Entomol. Soc. Am. 86: 799-812.
Avise, J.C., 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, Massachusetts.

Avise, J.C., 1994. Molecular Markers, Natural History and Evolution. Chapman and Hall, New York.

Axelrod, D.I., 1979. Age and origin of Sonoran Desert vegetation. Occasional Papers of the California Academy of Sciences, 132, 1–74.

Baker, A.C., Stone, W.E., Plummer, C.C., McPhail, M., 1944. A review of studies on the Mexican fruitfly and related Mexican species. U.S. Department of Agriculture Miscellaneous Publication 531, 1-155.

Barr, N.B., Cook, A., Elder, P., Molongoski, J., Prasher, D., Robinson, D.G., 2009. Application of a DNA barcode using the 16S rRNA gene to diagnose pest *Arion* species in the USA. J of Molluscan Studies. 75(2), 187-191.

Barr, N.B., Copeland, R.S., De Meyer, M., Masiga, D., Kibogo, H.G., Billah, M.K., Osir, E., Wharton, R.A., McPheron, B.A., 2006. Molecular diagnostics of economically important *Ceratitis* fruit fly species (Diptera: Tephritidae) in Africa using PCR and RFLP analyses. Bulletin of Entomological Research 96, 505-521.

Barr, N.B., McPheron, B.A., 2006. Molecular phylogenetics of the genus *Ceratitis* (Diptera: Tephritidae). Molecular Phylogenetics and Evolution. 38, 216–230.

Barr, N.B., Cui, L.W., McPheron, B.A., 2005. Molecular systematics of nuclear gene period in genus *Anastrepha* (Tephritidae). Annals of the Entomological Society of America 98, 173-180.

Beaumont, M.A., Panchal, M., 2008. On the validity of nested clade phylogeographical analysis. Molecular Ecology 17, 2563-2565.

Bohonak, A.J., 2002. IBD (isolation by distance): A program for analyses of isolation by distance. Journal of Heredity 93, 153-154.

Bonizzoni, M., Malacrida, A.R., Guglielmino, C.R., Gomulski, L.M., Gasperi, G., and Zheng, L., 2000. Microsatellite polymorphism in the Mediterranean fruit fly, *Ceratitis capitata.* Insect Mol. Biol. 9, 251-261.

Boykin, L.M., Shatters, R.G., Hall, D.G., Burns, R.E., Franqui, R.A., 2006. Analysis of host preference and geographical distribution of *Anastrepha suspensa* (Diptera: Tephritidae) using phylogenetic analyses of mitochondrial cytochrome oxidase I DNA sequence data. Bulletin of Entomological Research 96, 457-469.

Brito, P.H., 2005. The influence of Pleistocene glacial refugia on tawny owl genetic diversity and phylogeography in western Europe. Molecular Ecology 14, 3077-3094.

Brower, A.V.Z., 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial-DNA evolution. Proceedings of the National Academy of Sciences of the United States of America 91, 6491-6495.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9, 1657-1659.

Dillehay, T., 1997. Monte Verde: A Late Pleistocene Settlement in Chile, Volume 2: The Archaeological Context and Interpretation. Washington, D.C., Smithsonian Institution Press, 1071.

Epsky, N.D., Kendra, P.E., Heath, R.R., 2003. Development of lures for detection and delimitation of invasive *Anastrepha* fruit flies, pp. 84-89 *In* W. Klassen, W. Colon, and W. I. Lugo [eds.], Proc. of the 39th Annual Meeting of the Caribbean Food Crops Society, July 2003, Grenada.

Excoffier, L., Estoup, A., Cornuet, J.M., 2005. Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. Genetics 169, 1727-1738.

Excoffier, L., Smouse, P.E., Quattro, J.M.., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial-DNA restriction data. Genetics 131, 479-491.

Feder, J.L., Berlocher, S.H., Roethele, J.B., Dambrowski, H., Smith, J.J., Perry, W.L., Gavrilovic, V., Filchak, K.E., Rull, J., Aluja, M., 2003. Allopatric genetic origins for sympatric host-plant shifts and race formation in Rhagoletis. Proceedings of the National Academy of Sciences of the United States of America 100, 10314-10319.

Felsenstein, J., 1985. Phylogenies from gene-frequencies - a statistical problem. Systematic Zoology 34, 300-311.

Folmer, O., Back, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–299.

Foote, R.H., Blanc, F.L., Norrbom, A.L., 1993. ‘‘Handbook of the Fruit Flies (Diptera: Tephritidae) of America North of Mexico,’’ Comstock, Ithaca, NY.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915-925.

Gasparich, G.E., Sheppard, W.S., Han, H-Y., McPheron, B.A., Steck, G.J., 1994. Analysis of mitochondrial DNA and development of PCR-based diagnostic markers for Mediterranean fruit fly (*Ceratitis capitata*) populations. Insect Mol. Biol. 4, 61-67.

Graham, A., 1999. The tertiary history of the northern temperate element in the northern Latin American biota. American Journal of Botany 86, 32-38.

Guindon, S., Lethiec, F., Duroux, P., Gascuel, O., 2005. PHYML Online - a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Research 33, W557-W559.

Han, H-Y., McPheron, B.A., 1999. Molecular data as a tool to test phylogenetic relationships among higher groups of Tephritidae: A case study using mitochondrial ribosomal DNA sequences. Pp. 115-132 in Fruit Flies (Tephritidae): Phylogeny and Evolution of Behavior, M. Aluja and A. L. Norrbom, eds., CRC Press.

Hedges, S.B., Hass, C.A., Maxson, L.R., 1992. Caribbean biogeography - molecular evidence for dispersal in west-indian terrestrial vertebrates. Proceedings of the National Academy of Sciences of the United States of America 89, 1909-1913.

Hernandez, E., Orozco, D., Breceda, S.F., Dominguez, J., 2007. Dispersal and longevity of wild and mass-reared *Anastrepha ludens* and *Anastrepha obliqua* (Diptera: Tephritidae). Florida Entomologist 90, 123-135.

Hernandez-Ortiz, V., Gomez-Anaya, J.A., Sanchez, A., McPheron, B.A., Aluja, M., 2004. Morphometric analysis of Mexican and South American populations of the *Anastrepha fraterculus* complex (Diptera: Tephritidae) and recognition of a distinct Mexican morphotype. Bulletin of Entomological Research 94, 487-499.

Hernández-Ortiz, V., Aluja, M., 1993. Lista preliminar de especies del genero Neotropical *Anastrepha* (Diptera: Tephritidae) con notas sobre su distribución y plantas hospederas. Folia Entomol. Mex. (Mexico) 88, 89–105.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnean Society, 58, 247–276.

Hudson, R.R., 1991. Gene genealogies and the coalescent process, pp. 1–44 in Oxford Surveys in Evolutionary Biology, edited by D. Futuyama and J. Antonovics. Oxford University Press, Oxford.

Kimura, M., 1983. Rare variant alleles in the light of the neutral theory. Molecular Biology and Evolution 1, 84-93.

Knowles, L.L., 2008. Why does a method that fails continue to be used? Evolution 62, 2713-2717.

Knowles, L.L., Maddison, W.P., 2002. Statistical phylogeography. Molecular Ecology 11, 2623-2635.

Lanzavecchia, S.B., Cladera, J.L., Faccio, P., Petit, M. N., Vilardi, J.C., Zandomeni, R.O., 2008. Origin and distribution of *Ceratitis capitata* mitochondrial DNA haplotypes in Argentina. Annals of the Entomological Society of America 101, 627-638.

Lutz, A., Lima da Costa, A.M., 1918. Contribuçao para o estudo das Tripaneidas (moscas de frutas) brazileiras. Memorias do Instituto Oswaldo Cruz 10, 5–15.

Malavasi, A., Morgante, J.S., 1981. Biology of fruit-flies .3. adult and larval population fluctuation of *Anastrepha fraterculus* (Diptera, Tephritidae) and its relationship to host availability. Environmental Entomology 10, 275-278.

Mantel, N., 1967. Detection of disease clustering and a generalized regression approach. Cancer Research 27, 209-220.

McPheron, B.A., Han, H-Y., Silva, J.G., Norrbom, A.L., 1999. Phylogeny of Anastrepha and Toxotrypana based upon 16S rRNA mitochondrial DNA sequences. Pp. 343-361 in Fruit Flies (Tephritidae): Phylogeny and Evolution of Behavior, M. Aluja and A. L. Norrbom, eds., CRC Press.

McPheron, B.A., Gasparich, G.E., Han, HY., Steck, G.J., Sheppard, W.S., 1994. Mitochondrial-DNA restriction map for the Mediterranean fruit fly, *Ceratitis capitata*. Biochemical Genetics 32, 25-33.

Nielsen, R., Wakeley, J., 2001. Distinguishing migration from isolation: A Markov chain Monte Carlo approach. Genetics 158, 885-896.

Nielsen, R., 1998. Maximum likelihood estimation of population divergence times and population phylogenies under the infinite sites model. Theoretical Population Biology 53, 143-151.

Norrbom, A.L., Kim, K.C., 1988. Revision of the *schausi* group of *Anastrepha* Schiner (Diptera, Tephritidae), with a discussion of the terminology of the female terminalia in the Tephritoidea. Annals of the Entomological Society of America 81, 164-173.

Norrbom, A.L., Zucchi, R.A., Hernández-Ortiz, V., 1999. Phylogeny of the genera *Anastrepha* and *Toxotrypana* (Trypetinae: Toxotrypanini) based on morphology, p. 299-342. In M. Aluja A.L. Norrbom (eds.), Fruit flies (Tephritidae): phylogeny and evolution of behavior. Boca Raton, Florida, CRC Press, 944p.

Novotny, V., Clarke, A.R., Drew, R.A.I., Balagawi, S., Clifford, B., 2005. Host specialization and species richness of fruit flies (Diptera: Tephritidae) in a New Guinea rain forest. Journal of Tropical Ecology 21, 67-77.

Panchal, M., 2007. The automation nested clade phylogeographic analysis. Bioinformatics 23, 509-510.

Panchal, M., Beaumont, M.A., 2007. The automation and evaluation of nested clade phylogeographic analysis. Evolution 61, 1466-1480.

Petit, R.J., 2008. The coup de grace for the nested clade phylogeographic analysis? Molecular Ecology 17, 516-518.

Pfenninger, M., Posada, D., 2002. Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. Evolution Int. J. Org. Evolution, 56, 1776–1788.

Posada, D., Crandall, K.A., Templeton, A.R., 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Molecular Ecology 9, 487-488.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817-818.

Ricklefs, R.E., Bermingham, E., 2001. Nonequilibrium diversity dynamics of the Lesser Antillean avifauna. Science 294, 1522-1524.

Ricklefs, R.E., Bermingham, E., 2008. Likely human introduction of the Red-Legged Thrush (Turdus plumbeus) to Dominica, West Indies. Auk 125, 299-303.

Rogers, A.R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9,552–569.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Rouse, I., 1992. The Tainos. Rise and decline of the people who greeted Columbus. New Haven, CT; London, UK: Yale University Press. 562.

Rouse, I., 1989. In Biogeography of the West Indies. Past, present, and future (ed. C. A. Woods). Gainesville, FL: Sandhill Crane.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496-2497.

Sliva, J.G., Barr, N.B., 2008. Recent advances in molecular systematics of *Anastrepha* (Schiner). Fruit Flies of Economic Importance: From Basic to Applied Knowledge. Proceedings of the 7th International Symposium on Fruit Flies of Economic Importance 10-15 September 2006, Salvador, Brazil pp. 13-28.

Simon, C., Buckley, T.R., Frati, F., Stewart, J.B., Beckenbach, A.T., 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. Annual Review of Ecology Evolution and Systematics 37, 545-579.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., 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, 651-701.

Slatkin, M., Hudson, R.R., 1991. Pairwise comparisons of mitochondrial-DNA sequences in stable and exponentially growing populations. Genetics 129, 555-562.

Smith-Caldas, M.R.B., McPheron, B.A, Silva, J.G., Zucchi, R.A., 2001. Phylogenetic relationships among species of the *fraterculus* group (*Anastrepha*: Diptera: Tephritidae) inferred from DNA sequences of mitochondrial cytochrome oxidase I. Neotrop. Entomol. 30, 565-573.

Steck, G.J., 2001. Concerning the occurrence of *Anastrepha obliqua* (Diptera: Tephritidae,) in Florida. Florida Entomologist 84, 320-321.

Steck, G.J., 1991. Biochemical systematic and population genetic structure of *Anastrepha fraterculus* and related species (Diptera: Tephritidae). Ann Entomol Soc Am 84, 10–28.

Steck, G.J., Carrol, L.E., Celedonio-Hurtado, H., Guillen-Aguilar, J., 1990. Methods for identification of *Anastrepha* larvae (Diptera: Tephritidae), and key to 13 species. Proc. Entomol. Soc. Washington 92, 333-346.

Stone, A., 1942. The fruit flies of the genus *Anastrepha*. United States Department of Agriculture Miscellaneous Publication 439, 1–112.

Tajima, F., 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585-595.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596-1599.

Tan, K.H., Serit, M., 1994. Adult-population dynamics of *Bactrocera dorsalis* (Diptera, Tephritidae) in relation to host phenology and weather in 2 villages of Penang Island, Malaysia. Environmental Entomology 23, 267-275.

Tavare´, S., 1984. Line-of-descent and genealogical processes, and their applications in population genetics models. Theoretical Population Biology 26, 119–164.

Templeton, A.R., 2009. Why does a method that fails continue to be used? The answer. Evolution 63, 807-812.

Templeton, A.R., 2008. Nested clade analysis: an extensively validated method for strong phylogeographic inference. Molecular Ecology 17, 1877-1880.

Templeton, A.R., 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Molecular Ecology 7, 381-397.

Templeton, A.R, Routman, E., Phillips, C.A., 1995. Separating population-structure from population history - a cladistic-analysis of the geographical-distribution of mitochondrial-DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140, 767-782.

Templeton, A.R, Crandall, K.A., Sing, C.F., 1992. A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data. III. Cladogram estimation. Genetics 132, 619-633.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. Clustal-w - improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673-4680.

Vilardi, J.C., Saidman, B.O., Alberti, A.C., Calcagno, G.E., Basso, A., Manso, F., Cladera, J., 1994. Isoenzymatic and chromosomal analyses in three Argentine populations of *Anastrepha fraterculus*.

In: Proceedings, IV internatl. symposium on fruit flies of economic importance, 5–10 June, VIII-14, Sand Key, Florida.

Weems, HV., 1970. West Indian fruit fly *Anastrepha mombinpraeoptans* Seín. Entomology Circular 101, Florida Dept of Agriculture & Consumer Services, pp 2.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population-structure. Evol. 38, 1358-1370.

White, I.M., Elson-Harris, M., 1992. Fruit flies of economic significance: their identification and bionomics. International Institute of Entomology, London.

Wilson, S.M., 2001. The prehistory and early history of the Caribbean. In Biogeography of the West Indies. Patterns and Perspectives (eds C. A. Woods & F. E. Sergile), pp. 519–527. Boca Raton, FL: CRC Press.

Wilson, S.M, (ed.)., 1997. The indigenous people of the Caribbean. Gainesville, FL: University of Florida Press.

Wright, S., 1951. The genetical structure of populations. Annals of Eugenics 15, 323-354.

Xie, X., Michel, A.P., Schwarz, D., Rull, J., Velez, S., Forbes, A.A., Feder, J.L., 2008. Radiation and divergence in the *Rhagoletis pomonella* species complex: inferences from DNA sequence data. Journal of Evolutionary Biology 21, 900-913.

## Appendix A

Table A1. Population pairwise *F*st values for 67 geographic populations of *Anastrepha ludens*. Population numbers correspond to Table 1. Values in **bold** are significantly different from zero (P<0.05).

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | | | | | | | | | | | | | | | | |
| 2 | -0.3125 | 0 | | | | | | | | | | | | | | | |
| 3 | 0 | 0 | 0 | | | | | | | | | | | | | | |
| 4 | -0.32353 | -0.00743 | -0.0396 | 0 | | | | | | | | | | | | | |
| 5 | -0.3125 | 0 | 0 | -0.00743 | 0 | | | | | | | | | | | | |
| 6 | 0.36842 | 0.3253 | 0.70833 | **0.45022** | 0.4955 | 0 | | | | | | | | | | | |
| 7 | 0 | -0.3125 | 0 | -0.32353 | -0.3125 | 0.36842 | 0 | | | | | | | | | | |
| 8 | 0 | -0.09804 | 0 | -0.1215 | -0.09804 | 0.57895 | 0 | 0 | | | | | | | | | |
| 9 | 0.53973 | **0.78502** | **0.8** | **0.82183** | **0.78502** | 0.63415 | **0.53973** | **0.70213** | 0 | | | | | | | | |
| 10 | -0.2 | 0.27853 | 0.3 | 0.36567 | 0.27853 | 0.05556 | -0.2 | 0.11111 | 0.14286 | 0 | | | | | | | |
| 11 | 0.17241 | **0.3178** | **0.58824** | **0.44247** | **0.44863** | -0.28571 | 0.17241 | 0.42857 | 0.61905 | 0.05263 | 0 | | | | | | |
| 12 | -0.24619 | 0.04371 | 0.01573 | 0.04107 | 0.04371 | 0.54211 | -0.24619 | -0.06221 | **0.85761** | 0.43088 | **0.5316** | 0 | | | | | |
| 13 | 0.46429 | **0.53982** | **0.61957** | **0.51852** | **0.53982** | **0.62135** | 0.46429 | **0.55056** | **0.82687** | **0.43192** | **0.59321** | **0.57151** | 0 | | | | |
| 14 | 0 | 0.05405 | 0 | 0 | 0.05405 | **0.77528** | 0 | 0 | **0.84873** | 0.41176 | **0.6748** | 0.05517 | **0.66667** | 0 | | | |
| 15 | -0.32353 | -0.05205 | -0.0396 | 0 | -0.03209 | -0.00998 | -0.32353 | -0.1215 | **0.68103** | 0.13043 | 0.0283 | 0.03128 | **0.24561** | 0 | 0 | | |
| 16 | 0 | -0.16667 | 0 | -0.18421 | -0.16667 | 0.5 | 0 | 0 | 0.64103 | 0 | 0.33333 | -0.1201 | **0.51613** | 0 | -0.18421 | 0 | |
| 17 | -0.32371 | -0.0601 | -0.06198 | -0.03098 | -0.05253 | -0.0087 | -0.32371 | -0.13187 | **0.71387** | 0.16153 | 0.02948 | -0.01469 | **0.17586** | -0.03345 | -0.03052 | -0.19004 | 0 |

Table A1. (cont.)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 18 | -0.21212 | 0.05923 | 0.06198 | 0.05556 | 0.05923 | 0.48949 | -0.21212 | -0.02564 | **0.82693** | 0.37436 | **0.47368** | 0.10281 | **0.54321** | 0.11111 | 0.01587 | -0.08597 | -0.02516 |
| 19 | -0.31765 | 0.00129 | -0.01818 | -0.00599 | 0.00129 | **0.5269** | -0.31765 | -0.10891 | **0.80533** | 0.31976 | **0.48314** | 0.04255 | **0.55499** | 0.02946 | -0.01996 | -0.17483 | -0.04331 |
| 20 | -0.2375 | 0.04504 | 0.04742 | 0.05492 | 0.04504 | 0.37157 | -0.2375 | -0.04348 | **0.80008** | 0.324 | **0.37295** | 0.10106 | 0.25623 | 0.09836 | 0.01095 | -0.10615 | -0.01335 |
| 21 | -0.32353 | -0.03147 | -0.0396 | -0.01563 | -0.03147 | 0.0671 | -0.32353 | -0.1215 | **0.68223** | 0.1534 | 0.09845 | 0.03158 | **0.25926** | 0 | -0.05505 | -0.18421 | -0.0293 |
| 22 | -0.31765 | -0.01491 | -0.01818 | -0.00442 | -0.01491 | 0.03875 | -0.31765 | -0.10891 | **0.62744** | 0.08451 | 0.06503 | 0.05943 | **0.25323** | 0.02946 | -0.04947 | -0.17483 | -0.02828 |
| 23 | 0 | -0.09804 | 0 | -0.1215 | -0.09804 | 0.57895 | 0 | 0 | **0.70213** | 0.11111 | **0.42857** | -0.06221 | 0.55056 | 0 | -0.1215 | 0 | -0.13187 |
| 24 | 0.08083 | **0.33754** | **0.34071** | **0.39272** | **0.33754** | 0.20027 | 0.08083 | **0.23841** | -0.01104 | 0.02092 | 0.20221 | **0.45259** | **0.42055** | **0.40771** | **0.31483** | 0.18121 | **0.42009** |
| 25 | -0.12474 | **0.17576** | 0.17901 | **0.22894** | 0.17576 | 0.04245 | -0.12474 | 0.06912 | 0.12191 | -0.05579 | 0.04839 | **0.29266** | **0.28455** | **0.24795** | **0.14476** | 0.0025 | **0.23502** |
| 26 | -0.24528 | **0.08444** | **0.08696** | **0.14215** | **0.08444** | -0.05026 | -0.24528 | -0.02724 | **0.2448** | -0.08264 | -0.04216 | **0.20097** | **0.20857** | **0.15601** | 0.05806 | -0.1 | **0.15103** |
| 27 | 0.28612 | **0.51111** | **0.52273** | **0.55975** | **0.51111** | 0.39506 | 0.28612 | 0.42268 | 0.1087 | 0.125 | **0.39222** | **0.61516** | **0.59139** | **0.58783** | **0.39536** | 0.37 | **0.46439** |
| 28 | -0.19786 | 0.08333 | 0.11111 | 0.09305 | 0.08333 | 0.33436 | -0.19786 | 0.00444 | **0.76198** | 0.25509 | 0.32828 | 0.14859 | **0.47118** | **0.17586** | 0.0058 | -0.06329 | -0.02244 |
| 29 | -0.3211 | -0.01304 | -0.03067 | 0.00298 | -0.01304 | 0.33333 | -0.3211 | -0.11628 | **0.7952** | 0.30097 | **0.34146** | -0.03432 | **0.45482** | 0.01235 | -0.00648 | -0.18033 | -0.03759 |
| 30 | -0.32353 | 0.00829 | -0.0396 | 0 | 0.00829 | 0.57961 | -0.32353 | -0.1215 | **0.83616** | 0.38779 | **0.54047** | 0.04572 | **0.58333** | 0 | 0 | -0.18421 | -0.02935 |
| 31 | -0.2819 | 0.01651 | 0.01627 | **0.05781** | 0.01651 | 0.01115 | -0.2819 | -0.07731 | **0.60034** | 0.04892 | 0.03305 | 0.05763 | **0.2302** | **0.06776** | -0.00304 | -0.14286 | 0.00509 |
| 32 | 0 | -0.16667 | 0 | -0.18421 | -0.16667 | 0.5 | 0 | 0 | 0.64103 | 0 | 0.33333 | -0.1201 | 0.51613 | 0 | -0.18421 | 0 | -0.19004 |
| 33 | 0.11111 | 0.45894 | 0.48148 | **0.52525** | 0.45894 | 0.28492 | 0.11111 | 0.33333 | 0.34831 | -0.10714 | 0.2766 | **0.59272** | **0.56666** | 0.56989 | 0.28962 | 0.25 | **0.34455** |
| 34 | 0 | -0.3125 | 0 | -0.32353 | -0.3125 | 0.36842 | 0 | 0 | 0.53973 | -0.2 | 0.17241 | -0.24619 | 0.46429 | 0 | -0.32353 | 0 | -0.32371 |

Table A1. (cont.)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 35 | -0.29032 | 0.06084 | 0.07285 | 0.13773 | 0.06949 | -0.11303 | -0.29032 | -0.05263 | **0.29883** | -0.08954 | -0.10368 | **0.19686** | **0.20518** | 0.14894 | 0.04855 | -0.13208 | 0.13998 |
| 36 | -0.01064 | 0.21836 | 0.23826 | 0.24242 | 0.21836 | **0.31146** | -0.01064 | 0.14703 | **0.77398** | 0.28675 | **0.32241** | **0.19294** | **0.45833** | 0.2963 | 0.12698 | 0.09236 | **0.06733** |
| 37 | 0 | 0.03817 | 0 | -0.01124 | 0.03817 | **0.75676** | 0 | 0 | **0.8354** | 0.37931 | **0.65049** | 0.04383 | **0.65251** | 0 | -0.01124 | 0 | -0.04107 |
| 38 | 0 | -0.02439 | 0 | -0.05882 | -0.02439 | 0.67568 | 0 | 0 | **0.7757** | 0.25 | **0.54717** | -0.00292 | **0.6** | 0 | -0.05882 | 0 | -0.07734 |
| 39 | -0.27634 | 0.01282 | -0.00205 | 0.02946 | 0.01282 | 0.33649 | -0.27634 | -0.0836 | **0.81655** | 0.36074 | 0.35326 | 0.06773 | **0.45278** | 0.0393 | 0.01726 | -0.14416 | -0.01608 |
| 40 | 0 | 0.03817 | 0 | -0.01124 | 0.03817 | **0.75676** | 0 | 0 | **0.8354** | 0.37931 | **0.65049** | 0.04383 | **0.65251** | 0 | -0.01124 | 0 | -0.04107 |
| 41 | 0.6072 | **0.77248** | **0.78356** | **0.80914** | **0.77344** | 0.66022 | 0.6072 | **0.70838** | 0.2544 | 0.18814 | 0.65164 | **0.83818** | **0.8129** | **0.82696** | **0.70574** | 0.66759 | **0.72778** |
| 42 | -0.27762 | -0.00048 | -0.017 | 0.00598 | -0.00048 | 0.24479 | -0.27762 | -0.0911 | **0.82177** | 0.38653 | **0.27973** | 0.05275 | **0.37867** | 0.01761 | 0.03284 | -0.14884 | -0.00437 |
| 43 | -0.27595 | 0.01257 | 0.01176 | 0.05537 | 0.01257 | -0.0502 | -0.27595 | -0.07692 | **0.53634** | 0.07365 | -0.03394 | 0.09178 | **0.15945** | **0.05935** | 0.03621 | -0.14027 | **0.07109** |
| 44 | -0.3253 | 0.01282 | -0.0462 | 0.00365 | 0.01282 | **0.60186** | -0.3253 | -0.12532 | **0.84815** | 0.41644 | **0.56455** | 0.04863 | **0.59639** | -0.00917 | 0.00854 | -0.18705 | -0.02383 |
| 45 | 0.05288 | **0.3061** | **0.31722** | **0.36343** | **0.31226** | 0.16069 | 0.05288 | **0.2144** | 0.09407 | -0.06859 | 0.16491 | **0.42519** | **0.40755** | **0.3843** | **0.24457** | 0.15626 | **0.33267** |
| 46 | -0.25 | 0.03462 | 0.02778 | 0.04444 | 0.03462 | 0.39103 | -0.25 | -0.0582 | **0.81491** | 0.35352 | **0.39474** | 0.08763 | **0.48889** | 0.07407 | 0.01481 | -0.1194 | -0.01126 |
| 47 | 0 | -1 | 0 | -1 | -1 | 0 | 0 | 0 | 0.28205 | -1 | -0.33333 | -0.83333 | 0.33333 | 0 | -1 | 0 | -0.98441 |
| 48 | -0.09091 | 0.2276 | 0.23288 | 0.29675 | 0.2276 | 0.07843 | -0.09091 | 0.11111 | 0.44444 | 0.10469 | 0.0828 | 0.35721 | **0.34093** | 0.31034 | 0.23493 | 0.04 | **0.31258** |
| 49 | **0.76119** | **0.65249** | **0.84741** | **0.69108** | **0.73793** | -0.11418 | 0.76119 | **0.80723** | **0.8171** | **0.43262** | 0.07893 | **0.7382** | **0.76548** | **0.87322** | **0.27243** | **0.78761** | **0.22218** |
| 50 | -0.09873 | 0.1557 | 0.16233 | 0.20915 | 0.16 | 0.05551 | -0.09873 | 0.07164 | **0.31188** | -0.27111 | 0.06603 | **0.23798** | **0.27381** | 0.21769 | 0.11072 | 0.01429 | 0.14331 |
| 51 | 0.19643 | **0.41661** | **0.42149** | **0.46783** | **0.41661** | 0.30241 | 0.19643 | 0.32961 | 0.16279 | -0.09968 | 0.30381 | **0.51449** | **0.49506** | **0.48276** | **0.37449** | 0.28 | **0.43565** |

Table A1. (cont.)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 52 | -0.13402 | 0.05132 | **0.13142** | **0.16667** | **0.12706** | -0.20092 | -0.13402 | 0.04139 | **0.66173** | 0.17454 | -0.14616 | **0.21052** | **0.32** | **0.18519** | 0.02676 | -0.01695 | 0.07704 |
| 53 | -0.3211 | -0.01654 | -0.03067 | 0.00454 | -0.01654 | **0.26609** | -0.3211 | -0.11628 | **0.78416** | 0.29584 | **0.28391** | 0.04323 | **0.32034** | 0.01235 | -0.00529 | -0.18033 | -0.0274 |
| 54 | -0.04238 | **0.20667** | **0.25287** | **0.29248** | **0.24444** | -0.03514 | -0.04238 | 0.14131 | **0.4741** | 0.00047 | -0.0188 | **0.35992** | **0.37378** | **0.32442** | **0.15268** | 0.07645 | **0.20338** |
| 55 | 0 | 0.03817 | 0 | -0.01124 | 0.03817 | 0.75676 | 0 | 0 | **0.8354** | 0.37931 | **0.65049** | 0.04383 | **0.65251** | 0 | -0.01124 | 0 | -0.04107 |
| 56 | **0.76119** | **0.78879** | **0.84741** | **0.7493** | **0.78879** | **0.77527** | **0.76119** | **0.80723** | **0.79732** | **0.49241** | **0.73798** | **0.80733** | **0.80344** | **0.87322** | **0.48456** | **0.78761** | **0.37412** |
| 57 | -0.33043 | -0.01827 | -0.06567 | -0.02911 | -0.01827 | **0.43882** | -0.33043 | -0.13649 | **0.87341** | 0.48942 | **0.46269** | 0.0284 | **0.50818** | -0.03659 | 0.02944 | -0.19531 | -0.00431 |
| 58 | -0.31765 | 0.00129 | -0.01818 | -0.00599 | 0.00129 | **0.5269** | -0.31765 | -0.10891 | **0.80533** | 0.31976 | **0.48314** | 0.04255 | **0.55499** | 0.02946 | -0.01996 | -0.17483 | -0.05007 |
| 59 | 0 | -0.16667 | 0 | -0.18421 | -0.16667 | 0.5 | 0 | 0 | 0.64103 | 0 | 0.33333 | -0.1201 | 0.51613 | 0 | -0.18421 | 0 | -0.19004 |
| 60 | 0 | 0 | 0 | -0.0396 | 0 | 0.70833 | 0 | 0 | **0.8** | 0.3 | **0.58824** | 0.01573 | **0.61957** | 0 | -0.0396 | 0 | -0.06198 |
| 61 | -0.2423 | -0.06587 | 0.00647 | 0.04438 | 0.02277 | 0.05098 | -0.2423 | -0.06479 | **0.8426** | 0.43203 | 0.13195 | **0.07369** | **0.39929** | 0.04006 | 0.02446 | -0.11996 | 0.00961 |
| 62 | 0.40984 | **0.50296** | **0.58824** | **0.48276** | **0.50296** | **0.59091** | 0.40984 | 0.5102 | **0.81261** | 0.39576 | **0.56098** | **0.53997** | **0.64686** | **0.64072** | **0.21215** | 0.47059 | **0.14976** |
| 63 | -0.24631 | 0.03251 | 0.02478 | 0.05162 | 0.03251 | 0.30066 | -0.24631 | -0.05802 | **0.8061** | 0.32765 | **0.3212** | -0.0051 | **0.43119** | 0.06848 | 0.02705 | -0.11782 | -0.00957 |
| 64 | -0.26127 | 0.02593 | 0.02543 | 0.0534 | 0.02593 | -0.01335 | -0.26127 | -0.06415 | **0.474** | -0.05926 | 0.00552 | **0.10615** | **0.19647** | 0.07435 | -0.02288 | -0.12725 | 0.01341 |
| 65 | -0.26866 | 0.01983 | 0.01081 | 0.00383 | 0.01983 | 0.31767 | -0.26866 | -0.07425 | **0.80092** | 0.33417 | 0.33271 | **0.07754** | **0.44444** | 0.05556 | 0.01389 | -0.13586 | -0.02337 |
| 66 | 0 | 0.14865 | 0 | 0.0625 | 0.14865 | **0.85992** | 0 | 0 | **0.90801** | 0.5814 | **0.79009** | 0.12015 | **0.74719** | 0 | 0.0625 | 0 | 0.00203 |
| 67 | -0.32353 | -0.00743 | -0.0396 | 0 | -0.00743 | **0.45022** | -0.32353 | -0.1215 | **0.8239** | 0.36567 | **0.44247** | 0.04107 | **0.51852** | 0 | 0 | -0.18421 | -0.03098 |

Table A1. (cont.)

| | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 18 | 0 | | | | | | | | | | | | | | | | |
| 19 | 0.06305 | 0 | | | | | | | | | | | | | | | |
| 20 | 0.09767 | 0.05367 | 0 | | | | | | | | | | | | | | |
| 21 | 0.01709 | -0.01961 | 0.01253 | 0 | | | | | | | | | | | | | |
| 22 | 0.03707 | 0 | 0.02639 | -0.08572 | 0 | | | | | | | | | | | | |
| 23 | -0.02564 | -0.10891 | -0.04348 | -0.1215 | -0.10891 | 0 | | | | | | | | | | | |
| 24 | **0.39711** | **0.3619** | **0.37618** | **0.30522** | **0.26316** | **0.23841** | 0 | | | | | | | | | | |
| 25 | **0.24188** | **0.20055** | **0.21495** | 0.11731 | 0.06532 | 0.06912 | -0.0207 | 0 | | | | | | | | | |
| 26 | **0.15127** | **0.10922** | **0.12336** | 0.05053 | 0.01876 | -0.02724 | 0.01576 | -0.10363 | 0 | | | | | | | | |
| 27 | **0.56806** | **0.535** | **0.5383** | **0.37476** | **0.33099** | 0.42268 | 0.04722 | -0.0118 | 0.08618 | 0 | | | | | | | |
| 28 | 0.04071 | 0.09751 | 0.10344 | 0.009 | 0.01532 | 0.00444 | **0.33124** | **0.17544** | **0.08672** | **0.49645** | 0 | | | | | | |
| 29 | -0.04235 | -0.00658 | 0.04167 | -0.00614 | 0.00933 | -0.11628 | **0.37395** | **0.21435** | **0.12375** | **0.5335** | 0.02413 | 0 | | | | | |
| 30 | 0.07407 | -0.1232 | 0.07127 | 0 | 0.0251 | -0.1215 | **0.40385** | **0.24311** | **0.15144** | **0.57564** | 0.12461 | 0.0065 | 0 | | | | |
| 31 | 0.02083 | 0.03361 | **0.0548** | -0.01091 | -0.02635 | -0.07731 | **0.27458** | 0.11221 | 0.03314 | 0.35467 | 0.0073 | -0.0075 | **0.0625** | 0 | | | |
| 32 | -0.08597 | -0.17483 | -0.10615 | -0.18421 | -0.17483 | 0 | 0.18121 | 0.0025 | -0.1 | 0.37 | -0.06329 | -0.18033 | -0.18421 | -0.14286 | 0 | | |
| 33 | **0.53245** | 0.49216 | **0.48761** | 0.31562 | 0.26785 | 0.33333 | 0.08388 | 0.07246 | 0.02365 | 0.24251 | **0.43073** | 0.47945 | **0.54682** | **0.23344** | 0.25 | 0 | |
| 34 | -0.21212 | -0.31765 | -0.2375 | -0.32353 | -0.31765 | 0 | 0.08083 | -0.12474 | -0.24528 | 0.28612 | -0.19786 | -0.3211 | -0.32353 | -0.2819 | 0 | 0.11111 | 0 |

Table A1. (cont.)

| | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 35 | **0.14253** | 0.09677 | 0.11013 | 0.0635 | 0.02289 | -0.05263 | 0.05157 | -0.07994 | -0.16486 | 0.16968 | 0.07111 | 0.11159 | 0.1432 | 0.0241 | -0.13208 | 0.0404 | -0.29032 |
| 36 | 0.13712 | 0.21729 | **0.2327** | 0.13333 | 0.13005 | 0.14703 | **0.38756** | **0.2467** | **0.16505** | **0.5367** | 0.15038 | 0.0798 | 0.2517 | -0.01159 | 0.09236 | **0.47549** | -0.01064 |
| 37 | 0.09661 | 0.01562 | 0.08333 | -0.01124 | 0.01563 | 0 | **0.38754** | **0.22724** | **0.13537** | **0.56849** | 0.15649 | 0 | -0.01124 | **0.05263** | 0 | 0.5443 | 0 |
| 38 | 0.04 | -0.04025 | 0.02463 | -0.05882 | -0.04025 | 0 | **0.31263** | 0.14971 | 0.05714 | **0.49499** | 0.08322 | -0.05109 | -0.05882 | -0.00699 | 0 | 0.44186 | 0 |
| 39 | -0.06811 | 0.01968 | 0.06439 | 0.00511 | 0.03472 | -0.0836 | **0.408** | **0.25026** | **0.15968** | **0.56267** | 0.02156 | -0.03418 | 0.03339 | 0.02724 | -0.14416 | **0.51486** | -0.27634 |
| 40 | 0.09661 | 0.01562 | 0.08333 | -0.01124 | 0.01563 | 0 | **0.38754** | **0.22724** | **0.13537** | **0.56849** | 0.15649 | 0 | -0.01124 | **0.05263** | 0 | 0.5443 | 0 |
| 41 | **0.81141** | **0.79058** | **0.78975** | **0.7178** | **0.67763** | 0.70838 | **0.29803** | **0.40272** | **0.42615** | **0.51551** | **0.75795** | **0.78421** | **0.81797** | **0.63974** | 0.66759 | 0.43882 | 0.6072 |
| 42 | -0.01881 | 0.00704 | 0.05165 | 0.02378 | 0.03933 | -0.0911 | **0.45169** | **0.28752** | **0.19902** | **0.581** | 0.02926 | -0.0103 | 0.02049 | **0.05788** | -0.14884 | **0.52028** | -0.27762 |
| 43 | 0.05006 | 0.02941 | 0.03672 | 0.03728 | 0.02624 | -0.07692 | **0.22796** | 0.06834 | -0.03977 | **0.33272** | 0.01736 | 0.03571 | **0.0573** | 0.02798 | -0.14027 | 0.20742 | -0.27595 |
| 44 | 0.08028 | 0.00583 | 0.07989 | 0.00842 | 0.03605 | -0.12532 | **0.42008** | **0.25895** | **0.16535** | **0.59317** | 0.13743 | 0.01284 | 0.00046 | **0.07516** | -0.18705 | 0.56972 | -0.3253 |
| 45 | **0.3753** | **0.33641** | **0.35063** | **0.24091** | 0.20089 | **0.2144** | 0.00101 | 0.01984 | 0.04849 | 0.03749 | **0.30832** | **0.34691** | **0.3781** | **0.22278** | 0.15626 | -0.03829 | 0.05288 |
| 46 | 0.08889 | 0.04128 | 0.07963 | 0.00161 | 0.03278 | -0.0582 | **0.39823** | **0.23729** | **0.14698** | **0.55669** | 0.10462 | 0.03627 | 0.05556 | **0.06433** | -0.1194 | **0.51196** | -0.25 |
| 47 | -0.77778 | -1 | -0.83333 | -1 | -1 | 0 | -0.26923 | -0.64198 | -0.88571 | 0.04545 | -0.77778 | -1 | -1 | -0.92857 | 0 | -0.33333 | 0 |
| 48 | **0.30058** | 0.25555 | 0.2649 | 0.23961 | 0.20655 | 0.11111 | 0.15019 | 0.04756 | -0.06111 | **0.33333** | 0.22409 | 0.26829 | 0.30342 | 0.19601 | 0.04 | 0.22417 | -0.09091 |
| 49 | **0.71359** | **0.75** | **0.63653** | **0.34964** | **0.33333** | **0.80723** | **0.3964** | **0.27339** | **0.20348** | **0.58486** | **0.6227** | **0.61266** | **0.77186** | **0.28571** | **0.78761** | **0.56784** | 0.76119 |
| 50 | 0.21351 | 0.17962 | 0.19501 | 0.1253 | 0.09088 | 0.07164 | **0.18161** | 0.08805 | 0.06445 | **0.20712** | 0.16318 | 0.17985 | 0.21333 | 0.08571 | 0.01429 | 0.03287 | -0.09873 |
| 51 | **0.47411** | **0.43879** | **0.45115** | **0.38311** | **0.34247** | 0.32961 | **0.14055** | **0.19574** | **0.20122** | **0.26449** | **0.41149** | **0.44402** | **0.47714** | **0.32551** | 0.28 | 0.07598 | 0.19643 |

Table A1. (cont.)

| | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 52 | **0.17778** | **0.14472** | **0.1588** | 0.07005 | 0.05962 | 0.04139 | **0.32466** | **0.16375** | 0.08756 | **0.39515** | **0.13828** | **0.14634** | **0.17544** | 0.08046 | -0.01695 | **0.31805** | -0.13402 | |
| 53 | -0.03249 | -0.00808 | -0.02966 | -0.0049 | 0.00778 | -0.11628 | **0.36911** | **0.20699** | **0.11506** | **0.52273** | 0.01617 | -0.03279 | 0.00767 | 0.01638 | -0.18033 | 0.46057 | -0.3211 | |
| 54 | **0.30909** | **0.26968** | **0.28047** | 0.18349 | 0.15488 | 0.14131 | **0.19876** | **0.12665** | 0.0765 | **0.31341** | **0.23958** | **0.27339** | **0.31331** | 0.14781 | 0.07645 | 0.10472 | -0.04238 | |
| 55 | 0.09661 | 0.01562 | 0.08333 | -0.01124 | 0.01563 | 0 | **0.38754** | **0.22724** | **0.13537** | **0.56849** | 0.15649 | 0 | -0.01124 | **0.05263** | 0 | 0.5443 | 0 | |
| 56 | **0.73711** | **0.8** | **0.73193** | **0.47212** | **0.44086** | **0.80723** | **0.34615** | **0.26879** | **0.24653** | **0.58812** | **0.6768** | **0.68895** | **0.81951** | **0.34232** | **0.78761** | **0.60442** | **0.76119** | |
| 57 | 0.04649 | -0.01847 | 0.06145 | 0.02719 | 0.05858 | -0.13649 | **0.50313** | **0.34222** | **0.25132** | **0.64936** | 0.10534 | 0.0123 | -0.01593 | **0.11921** | -0.19531 | **0.61983** | -0.33043 | |
| 58 | -0.10414 | 0 | 0.05367 | -0.01961 | 0 | -0.10891 | **0.35885** | **0.20055** | 0.10922 | **0.535** | 0.03345 | -0.06993 | 0.00298 | 0.00433 | -0.17483 | 0.49216 | -0.31765 | |
| 59 | -0.08597 | -0.17483 | -0.10615 | -0.18421 | -0.17483 | 0 | 0.18121 | 0.0025 | -0.1 | 0.37 | -0.06329 | | -0.18033 | -0.18421 | -0.14286 | 0 | 0.25 | 0 |
| 60 | 0.06198 | -0.01818 | 0.04742 | -0.0396 | -0.01818 | 0 | **0.34071** | 0.17901 | **0.08696** | **0.52273** | 0.11111 | -0.03067 | -0.0396 | 0.01627 | 0 | 0.48148 | 0 | |
| 61 | 0.0148 | 0.02971 | **0.07394** | 0.0528 | **0.07212** | -0.06479 | **0.49294** | **0.33045** | **0.23853** | **0.61437** | 0.05939 | 0.01665 | 0.04245 | **0.08688** | -0.11996 | **0.56483** | -0.2423 | |
| 62 | **0.50939** | **0.51972** | **0.44444** | **0.22509** | **0.21849** | **0.5102** | **0.39883** | **0.25969** | **0.18182** | **0.57055** | **0.43362** | **0.41667** | **0.55085** | **0.19885** | 0.47059 | 0.46335 | 0.40984 | |
| 63 | 0.0317 | 0.04154 | 0.07737 | **0.02812** | 0.04182 | -0.05802 | **0.4081** | **0.24853** | **0.1588** | **0.55546** | 0.06473 | -0.03545 | 0.05863 | 0.01232 | -0.11782 | 0.50236 | -0.24631 | |
| 64 | 0.0614 | 0.04283 | 0.0625 | -0.04287 | -0.06284 | -0.06415 | **0.18027** | 0.02837 | -0.01018 | 0.2123 | 0.03122 | 0.046 | 0.0711 | -0.01721 | -0.12725 | 0.0815 | -0.26127 | |
| 65 | -0.06838 | 0.0282 | 0.06671 | 0.0015 | 0.01435 | -0.07425 | **0.38233** | **0.2236** | **0.13812** | **0.54156** | 0.01666 | -0.03292 | 0.04444 | 0.01623 | -0.13586 | 0.49271 | -0.26866 | |
| 66 | 0.19822 | 0.11033 | 0.18852 | 0.0625 | 0.11033 | 0 | **0.5263** | **0.37207** | **0.28** | **0.69455** | **0.29288** | 0.0828 | 0.0625 | **0.15789** | 0 | **0.70124** | 0 | |
| 67 | -0.04938 | -0.00599 | 0.05492 | 0 | 0.02139 | -0.1215 | **0.39764** | **0.23842** | **0.14706** | **0.5639** | 0.04775 | -0.04091 | 0 | 0.03401 | -0.18421 | 0.52525 | -0.32353 | |

Table A1. (cont.)

| | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 35 | 0 | | | | | | | | | | | | | | | | |
| 36 | 0.15518 | 0 | | | | | | | | | | | | | | | |
| 37 | 0.12621 | 0.27879 | 0 | | | | | | | | | | | | | | |
| 38 | 0.04 | 0.21379 | 0 | 0 | | | | | | | | | | | | | |
| 39 | 0.15048 | 0.12855 | 0.0274 | -0.02159 | 0 | | | | | | | | | | | | |
| 40 | 0.12621 | 0.27879 | 0 | 0 | 0.0274 | 0 | | | | | | | | | | | |
| 41 | **0.43617** | **0.76641** | **0.81462** | **0.76355** | **0.80602** | **0.81462** | 0 | | | | | | | | | | |
| 42 | 0.19482 | **0.14813** | 0.0079 | -0.03415 | -0.00661 | 0.0079 | **0.81722** | 0 | | | | | | | | | |
| 43 | -0.09185 | 0.08708 | 0.04545 | -0.01002 | 0.05712 | 0.04545 | **0.59487** | 0.08726 | 0 | | | | | | | | |
| 44 | 0.15735 | **0.28097** | -0.01957 | -0.06452 | 0.04 | -0.01957 | **0.8291** | 0.02654 | 0.06387 | 0 | | | | | | | |
| 45 | 0.09246 | **0.36947** | **0.36409** | **0.2891** | **0.38103** | **0.36409** | **0.33769** | **0.41451** | **0.24376** | **0.39648** | 0 | | | | | | |
| 46 | 0.1372 | 0.24074 | 0.06054 | 0.00662 | 0.03328 | 0.06054 | **0.80273** | 0.04904 | **0.05945** | 0.06271 | **0.36942** | 0 | | | | | |
| 47 | -1 | -0.40741 | 0 | 0 | -0.9 | 0 | 0.47826 | -0.89286 | -0.90909 | -1 | -0.31616 | -0.85185 | 0 | | | | |
| 48 | -0.13014 | **0.29681** | 0.28713 | 0.2 | 0.3074 | 0.28713 | **0.53751** | **0.35166** | -0.03327 | 0.31777 | **0.26438** | **0.29408** | -0.6 | 0 | | | |
| 49 | **0.15493** | **0.54967** | **0.86567** | **0.83618** | **0.5988** | **0.86567** | **0.79671** | **0.49967** | **0.18375** | **0.78162** | **0.36213** | **0.64549** | 0.71429 | **0.32702** | 0 | | |
| 50 | 0.06311 | 0.19664 | 0.20108 | 0.13858 | **0.21868** | 0.20108 | **0.31885** | **0.25096** | 0.14325 | 0.22827 | 0.07002 | 0.2094 | -0.56463 | 0.22895 | **0.2587** | 0 | |
| 51 | 0.21979 | **0.4492** | **0.46429** | **0.39597** | **0.47966** | **0.46429** | 0.14988 | **0.51406** | **0.34615** | **0.49402** | 0.06451 | **0.46871** | -0.07143 | **0.35311** | **0.47902** | 0.06544 | 0 |

Table A1. (cont.)

| | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 52 | 0.06065 | **0.22222** | **0.16911** | **0.10811** | **0.17125** | **0.16911** | **0.69251** | **0.18323** | **0.0938** | **0.18917** | **0.24496** | **0.16931** | -0.62963 | **0.25397** | -0.00131 | 0.14791 | **0.39237** |
| 53 | 0.10038 | 0.13884 | 0 | -0.05109 | -0.0265 | 0 | **0.77827** | -0.00401 | 0.02888 | 0.01497 | **0.34152** | 0.03283 | -1 | 0.25474 | **0.55783** | 0.18656 | **0.44318** |
| 54 | 0.05578 | **0.30599** | **0.30292** | 0.22267 | **0.30663** | **0.30292** | **0.52664** | **0.32693** | **0.15286** | **0.33259** | 0.08431 | **0.29905** | -0.49425 | **0.22906** | 0.16815 | 0.07409 | 0.15259 |
| 55 | 0.12621 | 0.27879 | 0 | 0 | 0.0274 | 0 | **0.81462** | 0.0079 | 0.04545 | -0.01957 | **0.36409** | 0.06054 | 0 | 0.28713 | **0.86567** | 0.20108 | **0.46429** |
| 56 | **0.27505** | **0.5315** | **0.86567** | **0.83618** | **0.6615** | **0.86567** | **0.80307** | **0.58072** | **0.26417** | **0.82798** | **0.38025** | **0.7412** | 0.71429 | **0.37911** | **0.85714** | **0.35293** | **0.48199** |
| 57 | 0.24899 | **0.2815** | -0.04437 | -0.08127 | 0.04065 | -0.04437 | **0.85986** | **0.03371** | 0.1191 | -0.01406 | **0.46843** | 0.04882 | -1 | 0.41485 | **0.65709** | **0.28954** | **0.56225** |
| 58 | 0.09677 | 0.15709 | 0.01562 | -0.04025 | -0.07431 | 0.01562 | **0.79058** | -0.0384 | 0.02257 | 0.00583 | **0.33641** | 0.04128 | -1 | 0.25555 | **0.75** | 0.17962 | **0.43879** |
| 59 | -0.13208 | 0.09236 | 0 | 0 | -0.14416 | 0 | 0.66759 | -0.14884 | -0.14027 | -0.18705 | 0.15626 | -0.1194 | 0 | 0.04 | **0.78761** | 0.01429 | 0.28 |
| 60 | 0.07285 | 0.23826 | 0 | 0 | -0.00205 | 0 | **0.78356** | -0.017 | 0.01176 | -0.0462 | **0.31722** | 0.02778 | 0 | 0.23288 | **0.84741** | 0.16233 | **0.42149** |
| 61 | 0.21781 | **0.18334** | 0.03061 | -0.01009 | 0.02306 | 0.03061 | **0.83462** | 0.03094 | **0.1129** | 0.04828 | **0.44577** | 0.07071 | -0.81699 | **0.3908** | **0.34839** | **0.2737** | **0.5487** |
| 62 | **0.17582** | **0.42836** | **0.625** | **0.56627** | **0.41662** | **0.625** | **0.80009** | **0.29477** | **0.13462** | **0.56513** | **0.38467** | **0.45307** | 0.25 | **0.31415** | **0.75239** | **0.25128** | **0.475** |
| 63 | 0.14911 | 0.07711 | 0.05577 | 0.00459 | 0.02737 | 0.05577 | **0.79536** | 0.03946 | 0.06499 | 0.06667 | **0.37794** | 0.07526 | -0.84 | **0.30421** | **0.56597** | 0.20136 | **0.46959** |
| 64 | 0.00682 | **0.10519** | 0.06 | 0.00324 | 0.06652 | 0.06 | **0.5643** | 0.08426 | 0.0407 | 0.08341 | 0.08239 | 0.07289 | -0.88 | 0.18802 | **0.23266** | 0.02284 | **0.22453** |
| 65 | 0.13212 | 0.11306 | 0.04255 | -0.0099 | -0.04643 | 0.04255 | **0.79422** | -0.0228 | 0.04675 | 0.05217 | **0.35519** | 0.06349 | -0.88889 | **0.2878** | **0.58891** | 0.20545 | **0.45965** |
| 66 | **0.28429** | **0.40293** | 0 | 0 | 0.10758 | 0 | **0.88681** | **0.07008** | **0.1414** | 0.04716 | **0.50365** | **0.15414** | 0 | 0.44615 | **0.91236** | **0.31908** | **0.59091** |
| 67 | 0.13773 | 0.18301 | -0.01124 | -0.05882 | -0.03669 | -0.01124 | **0.80914** | -0.01141 | 0.04968 | 0.00365 | **0.37209** | 0.04444 | -1 | 0.29675 | **0.69108** | 0.20915 | **0.47164** |

Table A1. (cont.)

| | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 52 | | | | | | | | | | | | | | | | |
| 53 | **0.14015** | 0 | | | | | | | | | | | | | | |
| 54 | 0.08026 | **0.26468** | 0 | | | | | | | | | | | | | |
| 55 | **0.16911** | 0 | **0.30292** | 0 | | | | | | | | | | | | |
| 56 | **0.49077** | **0.64807** | **0.40208** | **0.86567** | 0 | | | | | | | | | | | |
| 57 | **0.2303** | 0.02142 | **0.39064** | -0.04437 | **0.75306** | 0 | | | | | | | | | | |
| 58 | **0.14472** | -0.05824 | **0.26968** | 0.01562 | **0.77358** | -0.01847 | 0 | | | | | | | | | |
| 59 | -0.01695 | -0.18033 | 0.07645 | 0 | **0.78761** | -0.19531 | -0.17483 | 0 | | | | | | | | |
| 60 | **0.13142** | -0.03067 | **0.25287** | 0 | **0.84741** | -0.06567 | -0.01818 | 0 | 0 | | | | | | | |
| 61 | **0.08999** | 0.02221 | **0.30776** | 0.03061 | **0.62869** | 0.06335 | -0.00822 | -0.11996 | 0.00647 | 0 | | | | | | |
| 62 | **0.29149** | **0.375** | **0.34692** | **0.625** | **0.7931** | **0.47504** | **0.51972** | 0.47059 | **0.58824** | **0.36683** | 0 | | | | | |
| 63 | **0.17187** | 0.01908 | **0.3022** | 0.05577 | **0.66272** | **0.06878** | 0.00511 | -0.11782 | 0.02478 | 0.0643 | **0.39553** | 0 | | | | |
| 64 | 0.05797 | 0.0444 | 0.04706 | 0.06 | **0.25778** | 0.11456 | 0.03397 | -0.12725 | 0.02543 | **0.12008** | **0.16935** | 0.07614 | 0 | | | |
| 65 | **0.16162** | -0.0269 | **0.27959** | 0.04255 | **0.60689** | 0.03206 | -0.069 | -0.13586 | 0.01081 | 0.02284 | **0.40775** | 0.02663 | 0.04232 | 0 | | |
| 66 | **0.28313** | **0.0828** | **0.45155** | 0 | **0.91236** | 0 | 0.11033 | 0 | 0 | 0.0915 | **0.72932** | **0.14302** | **0.15967** | **0.13167** | 0 | |
| 67 | **0.16667** | -0.03168 | **0.30286** | -0.01124 | **0.7493** | -0.0024 | -0.08781 | -0.18421 | -0.0396 | 0.01585 | **0.48276** | 0.02492 | 0.0608 | -0.03175 | 0.0625 | 0 |

## Appendix B

Table B1. Population pairwise *F*st values for 52 geographic/host collections as observed in *Anastrepha obliqua*. Population codes and sited# in ( ) correspond to Table 7. Values in **bold** are significantly different from zero ($P<0.05$).

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. AC3SP(1) | 0 | | | | | | | | | | | | | |
| 2. AL1SP(2) | -0.1094 | 0 | | | | | | | | | | | | |
| 3. AP1SP(3) | -0.12676 | -0.14583 | 0 | | | | | | | | | | | |
| 4. AP3SP(6) | -0.1413 | 0.06266 | -0.03274 | 0 | | | | | | | | | | |
| 5. AP3MI(7) | -0.11483 | -0.01356 | -0.036 | 0.00306 | 0 | | | | | | | | | |
| 6. AP4SP(8) | **0.5443** | **0.48774** | 0.39189 | 0.1875 | **0.29609** | 0 | | | | | | | | |
| 7. AP5MI(9) | -0.01024 | 0.13227 | 0.02032 | -0.05952 | 0.06383 | 0.07216 | 0 | | | | | | | |
| 8. AP6SP(10) | -0.07425 | 0.04972 | 0.0364 | -0.02091 | 0.02299 | 0.33512 | 0.05381 | 0 | | | | | | |
| 9. CHI1SP(11) | -0.11385 | 0.02885 | -0.00715 | -0.08823 | -0.00244 | 0.3169 | -0.00294 | -0.11434 | 0 | | | | | |
| 10. CHISP(12) | -0.17246 | -0.05357 | -0.03659 | 0.02229 | -0.01064 | **0.45872** | **0.15112** | 0.03509 | 0.01908 | 0 | | | | |
| 11. JA1SP(13) | 0.0042 | 0.17611 | 0.05005 | -0.07908 | 0.08163 | 0.08114 | -0.09225 | 0.03704 | -0.03078 | 0.15638 | 0 | | | |
| 12. JA2SM(14) | -0.07744 | -0.05541 | -0.08434 | -0.00865 | 0.0056 | **0.31559** | 0.06165 | -0.05505 | -0.08006 | 0.00344 | 0.05498 | 0 | | |
| 13. JA3MI(15) | -0.05263 | 0.10377 | -0.01351 | -0.12945 | 0.01602 | 0.12621 | -0.16683 | -0.00737 | -0.06389 | 0.11067 | -0.16547 | 0.00752 | 0 | |
| 14. JA4SM(16) | -0.13481 | -0.08412 | -0.09284 | 0.02041 | -0.04762 | **0.36667** | 0.10493 | 0.04406 | 0.01839 | -0.05748 | 0.12527 | -0.01572 | 0.06362 | 0 |

Table B1. (cont.)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15. LLA2MI(26) | 0.10123 | 0.23335 | 0.09382 | -0.03747 | 0.11294 | 0.03817 | -0.08198 | 0.15176 | 0.09259 | 0.21485 | -0.06866 | 0.14038 | -0.12979 | 0.15562 |
| 16. LLA1ASM(17) | -0.18269 | -0.01897 | -0.01841 | -0.044 | -0.02357 | **0.38462** | 0.07704 | -0.06018 | -0.07738 | -0.07267 | 0.07039 | -0.03453 | 0.02392 | -0.03618 |
| 17. LLA1BSM(18) | 0.10123 | 0.23335 | 0.09382 | -0.03747 | 0.11294 | 0.03817 | -0.08198 | 0.15176 | 0.09259 | 0.21485 | -0.06866 | 0.14038 | -0.12979 | 0.15562 |
| 18. LLA1CSM(19) | 0.25 | 0.22311 | 0.11765 | -0.00862 | 0.11017 | 0 | -0.15385 | 0.15094 | 0.09055 | 0.28571 | -0.10039 | 0.13043 | -0.13208 | 0.17291 |
| 19. LLA1DSM(20) | 0.25333 | 0.12189 | 0.30988 | 0.36935 | 0.19463 | **0.8882** | 0.43017 | 0.07166 | 0.11176 | 0.27158 | 0.4835 | 0.08618 | 0.45908 | 0.20816 |
| 20. LLA1ESM(21) | -0.16667 | -0.03751 | -0.08629 | -0.1958 | -0.07046 | 0.37931 | -0.13889 | -0.21691 | -0.29528 | 0.01138 | -0.17071 | -0.17579 | -0.2038 | -0.02857 |
| 21. LLA1MI1(22) | -0.01471 | 0.18285 | 0.04215 | -0.07143 | 0.07791 | 0.125 | -0.06264 | 0.11696 | 0.05489 | 0.13926 | -0.05974 | 0.10966 | -0.12477 | 0.10687 |
| 22. LLA1MI2(23) | 0.02941 | 0.19498 | 0.06839 | -0.02971 | 0.11457 | 0.04315 | -0.04988 | 0.14996 | 0.08691 | 0.17928 | -0.03767 | 0.1431 | -0.10134 | **0.14502** |
| 23. LLA1MI3(24) | -0.11235 | 0.00086 | -0.09951 | -0.09922 | 0.00046 | **0.1174** | -0.0799 | -0.01951 | -0.08581 | 0.03604 | -0.08005 | -0.04599 | -0.13539 | 0.00559 |
| 24. LLA1MI4(25) | 0.51515 | **0.4603** | 0.3634 | 0.16837 | **0.2768** | 0 | 0.05138 | 0.31565 | 0.29317 | **0.43961** | 0.06538 | **0.29619** | 0.10112 | 0.34592 |
| 25. MBSM(28) | -0.17615 | -0.06504 | -0.12329 | -0.10278 | -0.07102 | 0.24888 | -0.02856 | -0.06731 | -0.11806 | -0.05037 | -0.03263 | -0.09416 | -0.08714 | -0.0895 |

Table B1. (cont.)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 26. MBTM(29) | -0.05405 | 0.10121 | 0.00374 | -0.08621 | 0.05533 | 0.0625 | -0.08434 | 0.03688 | -0.03591 | 0.10856 | -0.08979 | 0.04449 | -0.14307 | 0.08383 |
| 27. MAR1SM(27) | 0.13772 | -0.03521 | 0.11148 | **0.30682** | 0.1307 | **0.65408** | **0.38379** | **0.22043** | **0.23614** | 0.11631 | **0.41505** | 0.11695 | **0.37312** | 0.03852 |
| 28. PAZMI(30) | 0.5443 | **0.48774** | 0.39189 | 0.1875 | **0.29609** | 0 | **0.07216** | **0.33512** | **0.3169** | **0.45872** | 0.08114 | **0.31559** | 0.12621 | **0.36667** |
| 29. PAZT(31) | 0.0495 | -0.06371 | 0.02777 | **0.22458** | 0.0315 | **0.58333** | **0.31224** | **0.16268** | **0.16993** | 0.04247 | **0.33799** | 0.05866 | **0.29302** | -0.00152 |
| 30. PGSP(32) | -0.02525 | 0.06281 | -0.02203 | 0.00181 | -0.01543 | **0.23362** | -0.00294 | 0.06498 | 0.02222 | 0.10653 | 0.02322 | 0.0373 | -0.0456 | -0.01935 |
| 31. PNSP(33) | 0.19001 | 0.10498 | **0.27472** | **0.36687** | 0.14755 | **0.76778** | **0.46513** | **0.22648** | **0.27451** | 0.14048 | **0.48827** | **0.19656** | **0.47148** | 0.12098 |
| 32. PNMI(34) | 0.05028 | 0.00794 | 0.11017 | 0.17278 | 0.08438 | **0.65451** | 0.27265 | -0.03158 | -0.01746 | 0.08213 | 0.2862 | -0.05263 | 0.25 | 0.06694 |
| 33. POSP(35) | 0.08333 | -0.0754 | 0.02338 | 0.25168 | 0.02778 | **0.70353** | 0.31086 | 0.16589 | 0.16916 | 0.09193 | **0.36639** | 0.05094 | 0.31235 | -0.0741 |
| 34. PRSP(36) | -0.19241 | -0.04348 | -0.08108 | -0.05577 | -0.02246 | **0.35252** | 0.06931 | 0.00239 | -0.04525 | -0.0789 | 0.06156 | -0.02924 | 0.01919 | -0.05685 |
| 35. PSJSP(37) | -0.09589 | 0.08571 | -0.01083 | -0.053 | -0.01571 | 0.14006 | -0.02766 | 0.07407 | 0.01673 | 0.06085 | -0.01218 | 0.07071 | -0.07759 | 0.0515 |
| 36. RISP(38) | -0.10891 | -0.1268 | -0.1772 | 0.02354 | -0.00721 | 0.41168 | 0.10775 | 0.06131 | 0.03284 | -0.04686 | 0.12561 | -0.04979 | 0.07166 | -0.07558 |
| 37. TE1MI1(39) | -0.16667 | -0.0953 | -0.01585 | -0.03323 | -0.07046 | 0.65049 | 0.09559 | -0.21691 | -0.22761 | -0.07253 | 0.10485 | -0.17579 | 0.06327 | -0.06509 |
| 38. TE1MI2(40) | 0.19761 | 0.28636 | 0.11392 | 0.03698 | 0.17603 | -0.01124 | -0.03 | **0.22222** | 0.16903 | 0.28241 | -0.01449 | 0.18478 | -0.06472 | **0.21228** |
| 39. TOSP(41) | 0.01018 | -0.06053 | 0.01492 | 0.14612 | 0.06407 | **0.54676** | **0.24564** | 0.02229 | 0.02595 | 0.02797 | 0.25465 | -0.05753 | 0.21347 | 0.01307 |
| 40. TU1SM(42) | -0.11431 | -0.00955 | -0.03286 | -0.04737 | -0.00601 | 0.31559 | 0.04065 | -0.09428 | -0.12148 | -0.01124 | 0.02299 | -0.08974 | -0.02041 | -0.00532 |

Table B1. (cont.)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 41. TU2SM(44) | -0.18674 | 0.05482 | -0.02689 | -0.08527 | 0 | 0.31559 | 0.01645 | 0.03922 | -0.00997 | -0.02151 | 0.0113 | 0.03704 | -0.04348 | -0.0032 |
| 42. TU3SM(45) | -0.15851 | 0.01562 | -0.05161 | -0.10846 | -0.03782 | 0.19195 | -0.04763 | -0.04251 | -0.10054 | -0.0064 | -0.05429 | -0.0297 | -0.10824 | -0.03144 |
| 43. TUX1SM(46) | -0.2 | -0.03614 | -0.10723 | -0.07143 | -0.02777 | 0.33333 | 0.03116 | 0.02148 | -0.02133 | -0.06823 | 0.03339 | -0.01524 | -0.02273 | -0.05682 |
| 44. TUX2SM(47) | -0.2 | -0.03614 | -0.10723 | -0.07143 | -0.02777 | 0.33333 | 0.03116 | 0.02148 | -0.02133 | -0.06823 | 0.03339 | -0.01524 | -0.02273 | -0.05682 |
| 45. TUX2SA(48) | -0.113 | -0.02041 | -0.00176 | 0.01942 | -0.0066 | **0.44073** | 0.11902 | -0.05735 | -0.04301 | -0.03896 | 0.11765 | -0.05442 | 0.0625 | -0.01655 |
| 46. TUX3SA(49) | -0.14671 | -0.04524 | -0.02967 | 0.03141 | -0.0286 | **0.41279** | 0.13821 | 0.04233 | 0.02401 | -0.06782 | 0.1514 | 0.0135 | 0.0971 | -0.04042 |
| 47. TUX4SM(50) | 0.02911 | -0.00566 | 0.09366 | **0.20446** | **0.10843** | **0.46324** | **0.25717** | **0.13854** | **0.13697** | **0.08646** | **0.29894** | **0.1128** | **0.22506** | 0.05755 |
| 48. TUX5TM(51) | 0.28455 | 0.11714 | 0.33427 | 0.41428 | 0.177 | **0.88736** | **0.53333** | **0.2481** | **0.30951** | 0.14777 | **0.54197** | 0.19742 | **0.57243** | 0.11268 |
| 49. TUX6SM(52) | 0.02826 | -0.02154 | 0.09026 | **0.20483** | 0.11697 | **0.53495** | **0.28712** | **0.11035** | 0.11444 | 0.06295 | **0.31704** | 0.0745 | **0.26** | 0.06251 |
| 50. AP1CHT(4) | -0.01575 | 0.12941 | -0.02041 | -0.07784 | 0.07029 | 0.05026 | -0.08244 | 0.06619 | -0.00858 | 0.13194 | -0.09228 | 0.04353 | -0.14483 | 0.09219 |
| 51. AP1MT(5) | -0.05405 | 0.08268 | -0.05064 | -0.08621 | 0.04541 | 0.0625 | -0.08434 | 0.02619 | -0.03591 | 0.0971 | -0.08979 | 0.01393 | -0.14307 | 0.06135 |
| 52. TU1MI(43) | -0.23077 | -0.06509 | -0.13208 | -0.2233 | -0.1194 | 0.37931 | -0.20652 | -0.19879 | -0.23279 | -0.04369 | -0.2069 | -0.15802 | -0.29921 | -0.0736 |

Table B1. (cont.)

| | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15. LLA2MI(26) | 0 | | | | | | | | | | | | | |
| 16. LLA1ASM(17) | 0.14718 | 0 | | | | | | | | | | | | |
| 17. LLA1BSM(18) | -0.16667 | 0.14718 | 0 | | | | | | | | | | | |
| 18. LLA1CSM(19) | -0.16667 | 0.19215 | -0.16667 | 0 | | | | | | | | | | |
| 19. LLA1DSM(20) | **0.67692** | 0.16392 | **0.67692** | 0.7 | 0 | | | | | | | | | |
| 20. LLA1ESM(21) | 0.03012 | -0.12981 | 0.03012 | 0 | 0.01316 | 0 | | | | | | | | |
| 21. LLA1MI1(22) | -0.12197 | 0.08407 | -0.12197 | -0.0678 | **0.61643** | -0.00862 | 0 | | | | | | | |
| 22. LLA1MI2(23) | -0.10361 | 0.12629 | -0.10361 | -0.12958 | **0.59373** | 0.00798 | -0.08612 | 0 | | | | | | |
| 23. LLA1MI3(24) | -0.04831 | -0.01869 | -0.04831 | -0.09833 | 0.24307 | -0.22203 | -0.05465 | -0.04408 | 0 | | | | | |
| 24. LLA1MI4(25) | 0.02041 | **0.36389** | 0.02041 | 0 | **0.87532** | 0.34247 | 0.10717 | 0.02916 | 0.09707 | 0 | | | | |
| 25. MBSM(28) | 0.02778 | -0.09072 | 0.02778 | 0.02326 | 0.15305 | -0.22378 | -0.00996 | 0.02948 | -0.11003 | 0.22581 | 0 | | | |
| 26. MBTM(29) | -0.08133 | 0.03793 | -0.08133 | -0.125 | 0.40446 | -0.16667 | -0.07609 | -0.06783 | -0.08498 | 0.04636 | -0.05037 | 0 | | |
| 27. MAR1SM(27) | **0.47974** | **0.16164** | **0.47974** | **0.49398** | 0.17825 | **0.23516** | **0.43473** | **0.44102** | **0.23552** | **0.63636** | 0.13199 | **0.3549** | 0 | |
| 28. PAZMI(30) | 0.03817 | **0.38462** | 0.03817 | 0 | **0.8882** | 0.37931 | 0.125 | 0.04315 | 0.1174 | 0 | 0.24888 | 0.0625 | **0.65408** | 0 |

Table B1. (cont.)

| | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 29. PAZT(31) | **0.39715** | 0.09187 | **0.39715** | **0.41935** | 0.18823 | 0.16486 | **0.34763** | **0.37054** | 0.17364 | **0.56495** | 0.07132 | **0.28659** | -0.05953 | **0.58333** |
| 30. PGSP(32) | 0.05314 | 0.0647 | 0.05314 | 0.00858 | **0.33825** | -0.08581 | 0.04421 | 0.06348 | -0.03459 | **0.21076** | -0.08418 | 0.00816 | **0.23614** | **0.23362** |
| 31. PNSP(33) | **0.58824** | 0.15368 | **0.58824** | **0.63287** | 0.16248 | **0.31501** | **0.52709** | **0.52976** | **0.32367** | **0.75334** | 0.2087 | **0.4263** | 0.01711 | **0.76778** |
| 32. PNMI(34) | **0.43432** | 0.01487 | **0.43432** | 0.43128 | -0.2426 | -0.07761 | **0.38418** | **0.39829** | 0.11021 | **0.6319** | 0.01361 | 0.23954 | 0.07196 | **0.65451** |
| 33. POSP(35) | 0.46562 | 0.12049 | **0.46562** | 0.46497 | 0.14894 | 0.12426 | 0.41138 | **0.41482** | 0.15255 | **0.68091** | 0.02145 | **0.2948** | -0.14676 | **0.70353** |
| 34. PRSP(36) | 0.10786 | -0.08756 | 0.10786 | 0.16938 | **0.28816** | -0.08979 | 0.0421 | 0.08717 | -0.03855 | **0.33301** | -0.10059 | 0.02116 | **0.16905** | **0.35252** |
| 35. PSJSP(37) | -0.0345 | 0.02532 | -0.0345 | -0.04491 | **0.40601** | -0.0687 | -0.0508 | -0.01756 | -0.04035 | 0.12291 | -0.03178 | -0.03432 | **0.32022** | 0.14006 |
| 36. RISP(38) | 0.17012 | -0.01203 | 0.17012 | 0.20731 | **0.30639** | -0.00882 | 0.11356 | 0.14589 | -0.02083 | 0.38961 | -0.07521 | 0.08076 | 0.08036 | **0.41168** |
| 37. TE1MI1(39) | 0.3178 | -0.18033 | 0.3178 | 0.33333 | -0.2931 | -0.33333 | 0.23779 | 0.25159 | -0.06973 | **0.62205** | -0.16667 | 0.05263 | 0.06952 | **0.65049** |
| 38. TE1MI2(40) | -0.10889 | **0.22767** | -0.10889 | -0.18421 | **0.71698** | 0.11732 | -0.04957 | -0.04908 | -0.00529 | -0.02418 | 0.08884 | -0.02517 | **0.52096** | -0.01124 |
| 39. TOSP(41) | **0.35311** | 0.01541 | **0.35311** | 0.36122 | -0.04076 | -0.01521 | **0.30729** | **0.33346** | 0.09387 | **0.52632** | 0.00015 | **0.21534** | 0 | **0.54676** |
| 40. TU1SM(42) | 0.12317 | -0.07398 | 0.12317 | 0.13043 | 0.09165 | -0.21774 | 0.08363 | 0.12026 | -0.05115 | 0.29619 | -0.10363 | 0.01573 | 0.17217 | **0.31559** |

Table B1. (cont.)

| | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 41. TU2SM(44) | 0.02991 | -0.04667 | 0.02991 | 0.13043 | **0.46429** | -0.04869 | -0.03911 | 0.01173 | -0.04269 | 0.29619 | -0.0728 | -0.02259 | **0.29898** | 0.31559 |
| 42. TU3SM(45) | -0.01057 | -0.065 | -0.01057 | -0.01606 | 0.23182 | -0.21068 | -0.04061 | -0.01809 | -0.11211 | 0.17143 | -0.12897 | -0.06671 | **0.21515** | 0.19195 |
| 43. TUX1SM(46) | 0.06188 | -0.0678 | 0.06188 | 0.13772 | 0.37616 | -0.06404 | -0.00588 | 0.04124 | -0.05524 | 0.31265 | -0.10006 | -0.00413 | **0.20406** | 0.33333 |
| 44. TUX2SM(47) | 0.06188 | -0.0678 | 0.06188 | 0.13772 | 0.37616 | -0.06404 | -0.00588 | 0.04124 | -0.05524 | 0.31265 | -0.10006 | -0.00413 | **0.20406** | 0.33333 |
| 45. TUX2SA(48) | 0.22292 | -0.06745 | 0.22292 | 0.26554 | 0.14365 | -0.08425 | 0.16297 | 0.19883 | 0.02131 | 0.42147 | -0.06032 | 0.09971 | 0.13425 | **0.44073** |
| 46. TUX3SA(49) | 0.19364 | -0.05207 | 0.19364 | 0.2346 | 0.22026 | -0.00184 | 0.13347 | **0.17303** | 0.03882 | **0.39338** | -0.03717 | 0.10729 | 0.11267 | **0.41279** |
| 47. TUX4SM(50) | **0.31582** | 0.08074 | **0.31582** | **0.27881** | -0.03815 | 0.07831 | **0.29586** | **0.31147** | **0.1546** | **0.44295** | 0.08924 | **0.24756** | -0.03288 | **0.46324** |
| 48. TUX5TM(51) | **0.69933** | 0.18182 | **0.69933** | **0.80263** | 0.30061 | **0.40909** | **0.61281** | **0.58968** | **0.35524** | **0.87885** | 0.22702 | **0.46996** | -0.02128 | **0.88736** |
| 49. TUX6SM(52) | **0.36335** | 0.04694 | **0.36335** | **0.3472** | -0.0522 | 0.07256 | **0.3299** | **0.35572** | **0.16832** | **0.51429** | 0.0828 | **0.2604** | 0.00529 | **0.53495** |
| 50. AP1CHT(4) | -0.09343 | 0.06923 | -0.09343 | -0.12883 | 0.48895 | -0.13316 | -0.08263 | -0.05783 | -0.0945 | 0.03536 | -0.04686 | -0.08844 | **0.38362** | 0.05026 |
| 51. AP1MT(5) | -0.08133 | 0.02448 | -0.08133 | -0.125 | 0.40446 | -0.16667 | -0.07609 | -0.05048 | -0.10426 | 0.04636 | -0.06882 | -0.08871 | **0.33528** | 0.0625 |
| 52. TU1MI(43) | -0.04198 | -0.14596 | -0.04198 | 0 | 0.16 | -0.4 | -0.09814 | -0.06983 | -0.22628 | 0.34247 | -0.2283 | -0.18782 | 0.22581 | 0.37931 |

Table B1. (cont.)

| | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 29. PAZT(31) | 0 | | | | | | | | | | | | |
| 30. PGSP(32) | **0.19629** | 0 | | | | | | | | | | | |
| 31. PNSP(33) | 0.01091 | **0.34596** | 0 | | | | | | | | | | |
| 32. PNMI(34) | 0.0526 | 0.20221 | 0.08711 | 0 | | | | | | | | | |
| 33. POSP(35) | -0.0875 | 0.06735 | 0.0554 | 0.05251 | 0 | | | | | | | | |
| 34. PRSP(36) | | 0.08711 | 0.0521 | **0.2188** | 0.08805 | 0.12471 | 0 | | | | | | |
| 35. PSJSP(37) | **0.21515** | 0.03143 | **0.36027** | 0.24939 | 0.2574 | 0 | 0 | | | | | | |
| 36. RISP(38) | | 0.00327 | 0.05037 | 0.21422 | 0.09982 | 0.02041 | -0.06209 | 0.05133 | 0 | | | | |
| 37. TE1MI1(39) | 0.0197 | 0.04638 | 0.07091 | -0.30334 | 0.00671 | -0.08979 | 0.07285 | -0.00882 | 0 | | | | |
| 38. TE1MI2(40) | **0.44012** | 0.09749 | **0.63065** | **0.49631** | **0.50969** | 0.17535 | 0.03013 | 0.2042 | **0.40459** | 0 | | | |
| 39. TOSP(41) | -0.0262 | 0.17148 | 0.06728 | -0.14615 | -0.01605 | | 0.04142 | **0.20961** | 0.00227 | -0.18663 | **0.40424** | 0 | |
| 40. TU1SM(42) | **0.10853** | 0.03851 | **0.20882** | -0.04348 | 0.11472 | -0.04725 | 0.04444 | -0.00624 | -0.21774 | 0.18651 | -0.01871 | 0 | |
| 41. TU2SM(44) | **0.20795** | 0.04591 | **0.3578** | **0.23688** | 0.26662 | -0.07407 | -0.04377 | 0.00781 | 0.04924 | 0.11348 | **0.17341** | 0 | 0 |

Table B1. (cont.)

| | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 42. TU3SM(45) | 0.16008 | -0.06051 | **0.27477** | 0.09416 | 0.12801 | -0.06606 | -0.04215 | 0 | -0.10913 | 0.06314 | 0.09023 | -0.06481 | -0.07547 |
| 43. TUX1SM(46) | 0.115 | 0.02623 | 0.2766 | 0.15846 | 0.1565 | -0.10392 | -0.03155 | -0.07219 | -0.01408 | 0.12713 | 0.08966 | -0.02878 | -0.10118 |
| 44. TUX2SM(47) | 0.115 | 0.02623 | **0.2766** | 0.15846 | 0.1565 | -0.10392 | -0.03155 | -0.07219 | -0.01408 | 0.12713 | 0.08966 | -0.02878 | -0.10118 |
| 45. TUX2SA(48) | 0.07103 | 0.08733 | **0.15598** | -0.01613 | 0.10264 | -0.02672 | 0.08983 | -0.00392 | -0.17339 | **0.28862** | -0.01808 | -0.07407 | 0.02778 |
| 46. TUX3SA(49) | 0.02627 | 0.10181 | 0.11801 | 0.07407 | 0.07747 | -0.05485 | 0.04589 | -0.03064 | -0.07449 | 0.26152 | 0.03346 | 0.00112 | -0.00309 |
| 47. TUX4SM(50) | 0.03187 | **0.1756** | -0.01409 | 0.01353 | -0.04617 | **0.11875** | **0.2333** | 0.09898 | -0.05517 | **0.38102** | 0.02952 | **0.12348** | **0.19837** |
| 48. TUX5TM(51) | 0.01797 | **0.38001** | -0.10759 | 0.10324 | 0.05703 | 0.23952 | **0.40977** | 0.23582 | 0.13333 | **0.71553** | 0.05315 | **0.21759** | **0.4139** |
| 49. TUX6SM(52) | 0.01283 | **0.22718** | -0.02281 | -0.04426 | 0.00605 | 0.09243 | **0.24282** | 0.08014 | -0.10016 | **0.42666** | -0.03136 | 0.09079 | **0.20591** |
| 50. AP1CHT(4) | **0.30259** | 0.009 | **0.47292** | 0.28714 | **0.32689** | 0.03349 | -0.0274 | 0.06956 | 0.12227 | -0.06173 | 0.23601 | 0.03246 | -0.00753 |
| 51. AP1MT(5) | **0.25761** | -0.00901 | **0.41262** | 0.22585 | 0.26358 | 0.00753 | -0.03432 | 0.03663 | 0.05263 | -0.04681 | 0.18795 | 0.00435 | -0.02259 |
| 52. TU1MI(43) | 0.14471 | -0.13619 | 0.31657 | 0.01544 | 0.12583 | -0.11111 | -0.13297 | -0.0566 | -0.23529 | 0.04987 | 0.02419 | -0.19917 | -0.13586 |

Table B1. (cont.)

| | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 42. TU3SM(45) | 0 | | | | | | | | | | | |
| 43. TUX1SM | (46)-0.07261 | | 0 | | | | | | | | | |
| 44. TUX2SM | (47)-0.07261 | | -0.125 | 0 | | | | | | | | |
| 45. TUX2SA | (48)-0.01232 | | -0.01616 | -0.01616 | 0 | | | | | | | |
| 46. TUX3SA | (49)0.00335 | | -0.04567 | -0.04567 | -0.02171 | 0 | | | | | | |
| 47. TUX4SM | (50)**0.12561** | | **0.14412** | **0.14412** | **0.10149** | **0.08713** | 0 | | | | | |
| 48. TUX5TM(51) | **0.30132** | 0.31646 | 0.31646 | **0.16224** | 0.13196 | -0.02042 | 0 | | | | | |
| 49. TUX6SM | (52)0.14881 | | 0.13847 | 0.13847 | 0.08275 | 0.06447 | -0.06105 | -0.01469 | 0 | | | |
| 50. AP1CHT | (4) | -0.05195 | 0.0001 | 0.0001 | 0.12516 | 0.12963 | **0.2841** | **0.52361** | **0.29743** | 0 | | |
| 51. AP1MT(5) | -0.06671 | -0.02101 | -0.02101 | 0.089 | 0.09704 | **0.23766** | **0.46172** | **0.24008** | -0.10655 | 0 | | |
| 52. TU1MI(43) | -0.22483 | -0.13793 | -0.13793 | -0.15023 | -0.0501 | 0.07923 | **0.43103** | 0.10255 | -0.1629 | -0.18782 | 0 | |

## Appendix C

We sequenced individuals from species within the *fraterculus* group and other *Anastrepha* spp. For this comparison, the original 1548bp fragment was trimmed to 1328bp by trimming 142bp from COI (1-83 & 753-812), and 77bp from ND6 (1-68 and 727-736) to minimize any ambiguity due to aligning difficulties. Modeltest revealed a proportion of invariable sites (I) of 0.4936, and the gamma (Γ) shape distribution shape parameter of 0.2248. The GTR+I+G substitution model was found to be most applicable to the concatenated 1328bp COI and ND6 sequence based on the Akaike ranking (Akaike, 1974). We performed a maximum parsimony phylogenetic reconstruction to graphically investigate associations among sequences used in this study (Fig. 9), before calculating genetic distance estimates with Mega (Table 15).

In order to examine the specific identity of these individuals from sequence data, we compared COI sequences from each of the four phylogenetic clades (Fig. 2) to sequences deposited in GenBank. We were unable to compare ND6 regions from our data since GenBank was ND6 depauperate for this genus. The sequence match for all four lineage groups yielded a 99%-100% and 98%-100% maximum identity match with COI Folmer (Folmer et al., 1994) regions of *A. obliqua* (GenBank Accession #DQ116230) and *A. turpiniae* (GenBank Accession #DQ116229), respectively, with a lower percent identity to *A. sororcula* and *A. fraterculus* sequences. We investigated identity and quantified relative differences among haplotype groups by sequencing additional *Anastrepha* species as a point of reference and comparison.

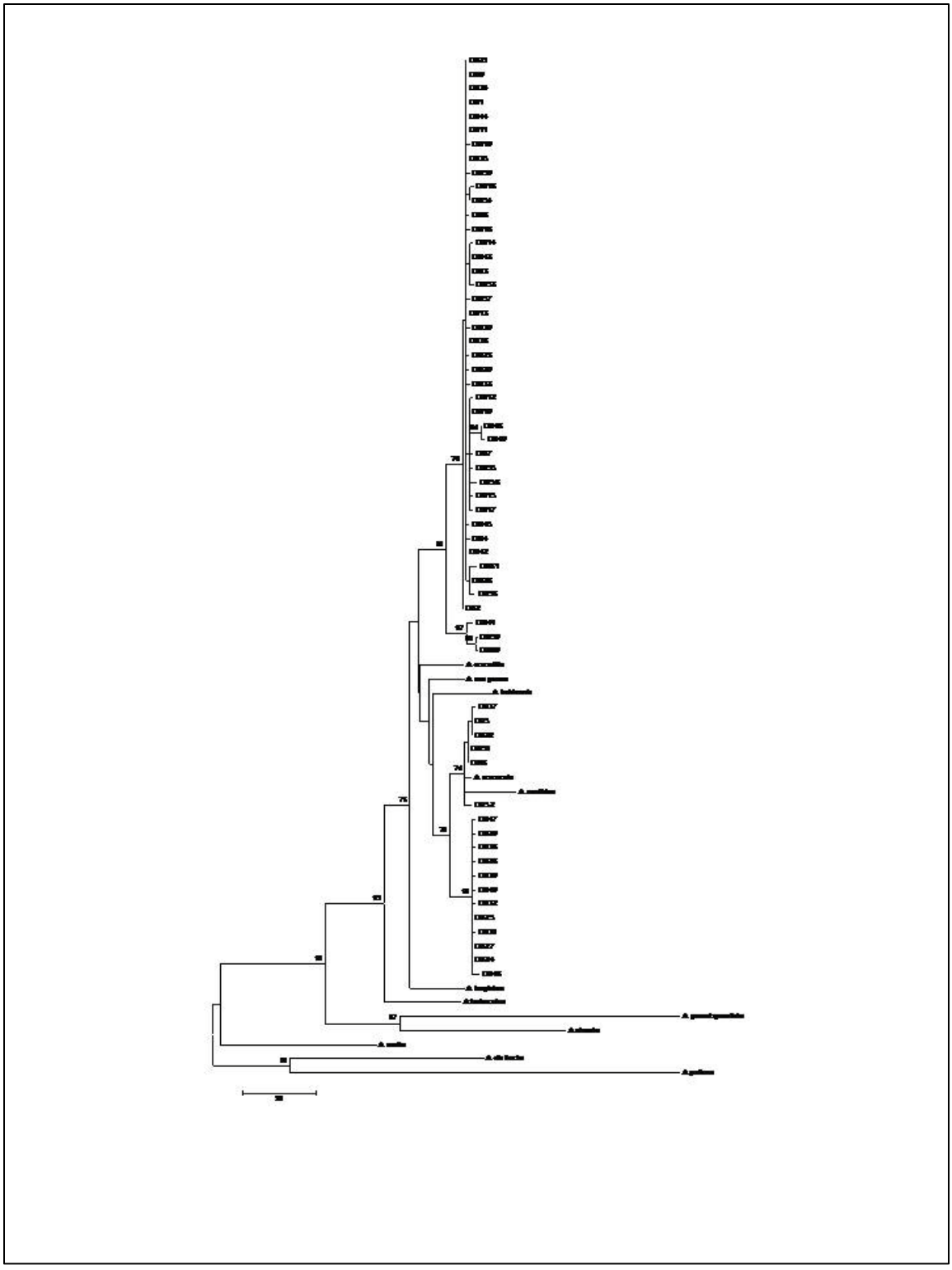

Figure C1. Phylogeny of 61 *Anastrepha obliqua* COI+ND6 haplotypes OB01-OB61 and various *Anastrepha* spp. used as comparison, constructed using a maximum parsimony (MP) approach. Numbers above branches show bootstrap support.

Table C1.Genetic distances for the 1328 bp COI+ND6 of I, II, III, and IV phylogenetic clades (Fig. 15) and selected species of *Anastrepha* examined for comparison. Distances (in lower diagonal) were gamma corrected and include standard error (in upper diagonal) observed.

| Clade/species | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. I | | *0.003* | *0.004* | *0.004* | *0.004* | *0.006* | *0.005* | *0.004* | *0.005* | *0.005* | *0.006* | *0.03* | *0.022* | *0.015* | *0.021* | *0.049* |
| 2. II | 0.011 | | *0.004* | *0.004* | *0.005* | *0.006* | *0.005* | *0.004* | *0.005* | *0.005* | *0.006* | *0.03* | *0.023* | *0.016* | *0.021* | *0.048* |
| 3. III | 0.018 | 0.021 | | *0.003* | *0.003* | *0.005* | *0.005* | *0.004* | *0.005* | *0.005* | *0.007* | *0.028* | *0.02* | *0.014* | *0.019* | *0.049* |
| 4. IV | 0.019 | 0.021 | 0.01 | | *0.001* | *0.003* | *0.004* | *0.004* | *0.005* | *0.005* | *0.006* | *0.028* | *0.022* | *0.016* | *0.02* | *0.05* |
| 5. *A sororcula*[a] | 0.019 | 0.021 | 0.01 | 0.003 | | *0.004* | *0.005* | *0.004* | *0.005* | *0.005* | *0.006* | *0.028* | *0.023* | *0.016* | *0.02* | *0.051* |
| 6. *A zenildae*[a] | 0.031 | 0.031 | 0.022 | 0.013 | 0.013 | | *0.005* | *0.005* | *0.006* | *0.007* | *0.009* | *0.033* | *0.024* | *0.017* | *0.022* | *0.055* |
| 7. *A coronilli*[a] | 0.022 | 0.023 | 0.021 | 0.02 | 0.02 | 0.026 | | *0.005* | *0.006* | *0.005* | *0.007* | *0.029* | *0.019* | *0.016* | *0.023* | *0.049* |
| 8. *A suspensa*[a] | 0.017 | 0.02 | 0.017 | 0.016 | 0.016 | 0.024 | 0.021 | | *0.005* | *0.005* | *0.006* | *0.027* | *0.022* | *0.016* | 0.02 | *0.046* |
| 9. *A bahiensis*[a] | 0.026 | 0.027 | 0.023 | 0.023 | 0.021 | 0.031 | 0.028 | 0.022 | | *0.005* | *0.007* | *0.025* | *0.024* | *0.018* | *0.019* | *0.044* |
| 10. *A turpiniae*[a] | 0.027 | 0.027 | 0.025 | 0.024 | 0.025 | 0.035 | 0.025 | 0.026 | 0.031 | | *0.006* | *0.031* | *0.021* | *0.015* | *0.025* | *0.051* |
| 11. *A fraterculus*[a] | 0.036 | 0.037 | 0.037 | 0.034 | 0.03 | 0.048 | 0.038 | 0.036 | 0.041 | 0.037 | | *0.029* | *0.021* | *0.019* | *0.025* | *0.059* |
| 12. *A pseudoparallela*[b] | 0.178 | 0.181 | 0.17 | 0.17 | 0.17 | 0.194 | 0.173 | 0.163 | 0.161 | 0.185 | 0.171 | | *0.03* | *0.047* | *0.044* | *0.083* |
| 13. *A alveata*[c] | 0.12 | 0.123 | 0.113 | 0.122 | 0.121 | 0.133 | 0.111 | 0.117 | 0.128 | 0.117 | 0.114 | 0.18 | | *0.038* | *0.034* | *0.045* |
| 14. *A amita*[a] | 0.107 | 0.111 | 0.098 | 0.105 | 0.107 | 0.121 | 0.109 | 0.107 | 0.118 | 0.103 | 0.124 | 0.275 | 0.218 | | *0.019* | *0.049* |
| 15. *A distincta*[a] | 0.143 | 0.141 | 0.131 | 0.136 | 0.133 | 0.145 | 0.147 | 0.133 | 0.128 | 0.159 | 0.155 | 0.258 | 0.203 | 0.128 | | *0.055* |
| 16. *A pallens*[d] | 0.275 | 0.272 | 0.27 | 0.272 | 0.274 | 0.295 | 0.269 | 0.267 | 0.262 | 0.28 | 0.305 | 0.427 | 0.265 | 0.258 | 0.276 | |

[a] *fraterculus,* [b] *pseudoparallela,* [c] *spatulata,* [d] *daciformis* , taxonomic groups.

This comparison included 7 species within the *fraterculus* group (including the three species mentioned above) plus 5 species from other taxonomic groups within *Anastrepha*. The genetic distances were gamma corrected for a 1328bp fragment, a shorter portion of the 1548bp fragment used in the larger *A. obliqua* study. The shorter fragment removed ambiguities resulting from alignments of species from different taxonomic groups. Genetic distances based on this shorter fragment among the four *A. obliqua* lineages range from 0.01-0.021 (avg 0.015±0.004), and distances among the 8 members of the *fraterculus* taxonomic group range from 0.013-0.147, with an average of 0.118±0.013. A comparison of the four *A. obliqua* lineages to the other *fraterculus* group species range between

0.003-0.143 with an average genetic distance of 0.020±0.002. *Anastrepha sororcula* and haplotypes from clade IV were the most similar in this comparison. We see from this comparison that genetic distances of haplotypes within the clade IV range from 0.001 to 0.004 (avg 0.0016±0.0011) and differences when compared with *A. sororcula* range from 0.002-0.004 (avg 0.0028±0.0016).

## DISCUSSION

Pairwise sequence comparisons revealed that haplotypes belonging to clade IV share a high degree of similarity with two members of the *fraterculus* taxonomic group. Our phylogenetic analysis of haplotypes from the four clades and *Anastrepha sororcula* and *Anastrepha zenildae* suggest that members from clade IV may be conspecific with these two species.

*Anastrepha sororcula* and *A. zenildae,* from the *fraterculus* group form a monophyletic group with haplotypes from clade IV (Fig. 9). Three additional species from the *fraterculus* group, *A. suspensa, A. coronilli* and *A. bahiensis,* also cluster within the clade containing all *A. obliqua* but are not monophyletic with respect to any single *A. obliqua* lineage. *Anastrepha fraterculus* and *A. turpiniae,* the other two group members included in this analysis, lie outside this assemblage of *A. obliqua* and other species, forming a monophyletic group with respect to five other *Anastrepha* spp. included (Fig. 9). The average genetic distance among *fraterculus* group species was 0.020±0.002 within the range observed by Smith-Caldas et al., (2001) using COI and McPheron et al. (1999) using 16S regions, averaging 0.033 ± 0.006 and 0.018 ± 0.001, respectively. The genetic difference between clades III/IV and I/II examined here averaged 0.020±0.002. We see that *A. sororcula* shared a high degree of similarity with haplotypes from clade IV. Boykin et al., (2006) also saw this relationship among similar geographic groups with COI sequences. Their data, while not as heavily represented by *A. obliqua* collections, are consistent with the paraphyletic arrangement of lineages

represented in this study. The monophyletic association among *A. obliqua* and *A. sororcula* is also seen in the Boykin et al. (2006) study. Our findings suggest that individuals from the SAII lineage and the *A. sororcula* and *A. zenildae* examined may be conspecific, however given the small sampling size of species representing the *fraterculus* taxonomic group and more specifically those species that show close association to the *A. obliqua* haplogroups we have idenitified, further review of these taxa is warranted.

## LITERATURE CITED


Akaike H (1974) New look at statistical-model identification. *IEEE Transactions on Automatic Control* **AC19**, 716-723.

Boykin LM, Shatters RG, Hall DG, Burns RE & Franqui RA (2006) Analysis of host preference and geographical distribution of *Anastrepha suspensa* (Diptera: Tephritidae) using phylogenetic analyses of mitochondrial cytochrome oxidase I DNA sequence data. *Bulletin of Entomological Research* **96**, 457-469.

Folmer O, Back M, Hoeh W, Lutz R, & Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **3**, 294–299.

McPheron BA, Han H-Y, Silva JG, & Norrbom AL (1999) Phylogeny of Anastrepha and Toxotrypana based upon 16S rRNA mitochondrial DNA sequences. Pp. 343-361 in Fruit Flies (Tephritidae): Phylogeny and Evolution of Behavior, M. Aluja and A. L. Norrbom, eds., CRC Press.

Smith-Caldas MRB, McPheron BA, Silva JG, & Zucchi RA (2001) Phylogenetic relationships among species of the *fraterculus* group (*Anastrepha*: Diptera: Tephritidae) inferred from DNA sequences of mitochondrial cytochrome oxidase I. Neotrop. Entomol. **30**, 565-573.

# Curriculum Vitae

## Raul Ruiz

## Education

2010 (Dec.) PhD degree in Entomology, The Pennsylvania State University: University Park, PA.

1999 (Dec.) M.S. degree in Biology, University of Texas-Pan American: Edinburg, Tx.

1987 (May) B.S. degree in Chemistry, University of Texas-Pan American: Edinburg, Tx.

## Thesis

Ruiz, R.A. 2009. Phylogeographic study of Mexican fruit fly (*Anastrepha ludens*) and West Indian fruit fly (*A. obliqua*). PhD Thesis. The Pennsylvania State University.

Ruiz, R.A. 1999. Using RAPD-PCR (Randomly Amplified Polymorphic DNA-Polymerase Chain Reaction) to distinguish U.S. Populations of *Acroptilon repens* L. (DC). M.S. Thesis. University of Texas-Pan American, Edinburg, Texas.

## Experience

01/2001 to present. Geneticist GS440-12, USDA APHIS PPQ Edinburg, Tx.

## Recent Publications/Presentations

Ruiz-Arce, R. Barr, N. B., and DeMeyer, M. Molecular tools for the pathway analysis of *Ceratitis capitata*. 8th International Symposium on Fruit Flies of Economic Importance. Valencia, Spain. November 2010.

Ruiz-Arce, R., B.A. McPheron, H.C. Tu, E. Deutsch, C.L. Owens, M. Aluja, and J. Piñero. Sequenced-based Analysis of Population Structure in *Anastrepha obliqua*. 7th International Symposium on Fruit Flies of the Economic Importance and the 6th Meeting of Working Group on Fruit Flies of the Western Hemisphere. Salvador Bahia, Brazil, September 2006.

Goolsby, J.A., P. J. De Barro, A. A Kirk, R. Sutherst, L. Canas, M. Ciomperlik, P. Ellsworth, J. Gould, K.A. Hoelmer, S. J. Naranjo, M. Rose, W. Roltsch, R. Ruiz, C. Pickett, and D. Vacek. 2005. Post-release evaluation of the biological control of *Bemisia tabaci* biotype "B" in the USA and the development of predictive tools to guide introductions for other countries. Biological Control 32: 70-77.

Ruiz, R.A., D., Schwarz, N.L. Shakir-Botteri, B.A. McPheron. Genetic Diversity in the Lonicera Fly (Diptera: Tephritidae) Population Differentiation of a *Rhagoletis* species by Molecular Analysis. Entomological Society of America Annual Meeting, Ft. Lauderdale, Fl., USA, December 2005.