Hereditas 136: 240–245 (2002)

Genetic divergence within the *Drosophila mayaguana* subcluster, a closely related triad of Caribbean species in the *repleta* species group

P. M. O'GRADY[1], C. M. DURANDO[1,2], W. B. HEED[3], M. WASSERMAN[2], W. ETGES[4] and R. DESALLE[1]

[1]*American Museum of Natural History, Division of Invertebrate Zoology, NY, USA*
[2]*Queens College, Department of Biology*
[3]*University of Arizona, Department of Ecology and Evolutionary Biology, Tucson, AZ, USA*
[4]*University of Arkansas, Department of Biological Sciences, Fayetteville, AR, USA*

O'Grady, P. M., Durando, C. M., Heed, W. B., Wasserman, M., Etges, W. and DeSalle, R. 2002. Genetic divergence within the *Drosophila mayaguana* subcluster, a closely related triad of Caribbean species in the *repleta* species group.—*Hereditas 136*: 240–245. Lund, Sweden. ISSN 0018-0661. Received June 24, 2002. Accepted July 28, 2002

The *mayaguana* triad of the *Drosophila repleta* species group includes *D. mayaguana*, *D. straubae*, and *D. parisiena*, the latter two of which are very similar when examined morphologically. Many morphological characters used to define these taxa are quantitative and overlap substantially among some forms—it is only through suites of such characters that species can be identified. We apply Population Aggregation Analysis and tree building methods to five rapidly evolving gene regions—the mitochondrial AT rich region and the nuclear *acetylcholinesterase*, *hunchback*, *mastermind*, and *vestigial* loci to test the morphological species delineations within the morphocryptic *mayaguana* triad. We find that *D. mayaguana* is diagnosable using DNA sequences, but the other two species form a non-diagnosable paraphyletic assemblage. A single ecological factor, oviposition substrate, is an important diagnostic character distinguishing *D. straubae* from *D. parisiena*, highlighting the importance of examining a diverse array of data (morphological, molecular, ecological, and behavioral) when defining species limits.

R. DeSalle, American Museum of Natural History, Division of Invertebrate Zoology, Central Park West @ 79th St, New York, NY 10024, USA. E-mail: desalle@amnh.org

The *Drosophila repleta* species group contains about one hundred described species and, although some are now cosmopolitan in distribution, all members of this group were originally endemic to the New World (ETGES et al. 2001). The *repleta* species group is divided into the *fasciola*, *hydei*, *mercatorum*, *mulleri*, and *repleta* subgroups based on polytene chromosome banding patterns (reviewed in WASSERMAN 1982, 1992). The *mulleri* subgroup is the largest and contains a number of complexes and clusters. The *mulleri* subgroup also includes the species *mayaguana* subcluster, or triad, which is composed of three very closely related species that are difficult to differentiate from one another using morphological characters. The *mayaguana* triad (HEED and GRIMALDI 1991) is restricted to the Caribbean and includes *D. mayaguana* VILELA (Bahamas, British Virgin Islands, Cuba, Grand Cayman, Hispanola, Jamaica), *D. parisiena* HEED and GRIMALDI (Cuba, Hispanola, Jamaica), and *D. straubae* HEED and GRIMALDI (Cuba, Hispanola, Navassa).

The *mayaguana* triad has been characterized as "morphocryptic," in reference to how difficult it is to separate species in the field. Even so, HEED and GRIMALDI (1991) described two new species, *D. straubae* and *D. mayaguana*, based on larval, pupal, adult, and ecological characters. Specifically, they include the shape and number of prensisetae on the surstyli, length of the aedeagus, size of the distal tooth on the distiphallus, shape of the ovipositor, size of the anterior spiracle on pupae, the depth of the hook on the larval mandible, and larval substrate (HEED and GRIMALDI 1991). Taken together the characters are discrimatory even though there is some degree of morphological overlap between *D. straubae* and *D. parisiena* suggesting little differentiation between these taxa. Hence, the absolute separation in larval substrate becomes highly as an isolating mechanism.

We have analyzed DNA sequences from five rapidly evolving gene regions, the mitochondrial AT rich region and the nuclear *acetylcholinesterase*, *hunchback*, *mastermind*, and *vestigial* loci, to better elucidate patterns of differentiation within the *mayaguana* subcluster. A total of 85 individuals from multiple populations of *D. mayaguana* (n = 6 in 4 populations), *D. parisiena* (n = 61 in 13 populations), and *D. straubae* (n = 13 in 6 populations) were sampled from across the ranges of each species. We compare the patterns of divergence observed at the DNA sequence level with information from previous studies using morphology, mating behavior, chromosomal inversions, allozymes and ecological data for this triad in order to better resolve the genetic affinities of these species.

MATERIALS AND METHODS

Taxon sampling and sequence selection

Table 1 lists both field-collected specimens and lab-cultured flies used in this study. Flies collected in the field were identified by examination of the male genitalia. Single males were used as a source of DNA. In particular, we differentiated between *D. straubae* and *D. parisiena* by inspecting the surstyli. HEED and GRIMALDI (1991) found that the inner margin of the surstylus of *D. straubae* is almost straight while that of *D. parisiena* is markedly crescentic. We chose two Caribbean strains of *D. mulleri*, one from Cayman Brac and another from Haiti, as outgroups to root our trees. Chromosomal data suggest that this species gave rise to *D. mayaguana*, the most basal of the three species in the *D. mayaguana* subcluster (WASSERMAN 1992). Sequences for either one or both of these *D. mulleri* isolates were generated for all gene regions but *mastermind*. DNA was isolated from flies using the methods outlined in DURANDO et al. (2000).

We selected five rapidly evolving gene regions to assay for variability in the strains used (Table 1). Three of these genes, *Acetylcholinesterase* (*Ace*; FOURNIER et al. 1989), *mastermind* (*mam*; BETTLER et al. 1996; NEWFIELD et al. 1991) and *vestigial* (*vg*; WILLIAMS et al. 1991), span introns in nuclear protein coding genes. Another sequenced region is the hunchback gene, which contains a polyglutamine (CAA or CAG) repeat (TREIER et al. 1989). The final region is from the AT rich region of the mitochondrial genome (LEWIS et al. 1994). Primers, length of fragments sequenced and the molecular nature of the variability detected in the five regions are described in Table 2. The polymerase chain reaction (PCR) was used to amplify fragments for sequencing. PCR products were cleaned using Geneclean kits (BIO 101) and sequenced directly using automated sequencing methods with the ABI dye terminator system on an ABI 377 automated sequencer. Sequences were then corrected and edited using the SEQUENCHER software (GENECODES CORPORATION 1995).

Table 1. *Collection information and abbrevations for populations used in this study*

Species	Collection Information	AMNH FTC GenBank
Ingroup		
D. mayaguana	Skyline Drive Restricted Impact Area, Guantanamo Bay Naval Base, Cuba, 15-18.xi.1988, ORV33, Heed & Starmer	102291
	Conception Island, Bahamas, 18.xi.1983, ORV20, Heed & Thomas	102278
	Grand Cayman, British Crown Colony, 28.xi.1983, ORV29, Wasserman & Wheeler	102280
	27 km NE of Barahona, Republica Dominicana, A983	N/A
D. parisiena	Monte Cristi, Republica Dominicana, A950	N/A
	Barahona, Republica Dominicana, A951	N/A
	Guantanamo Bay, Cuba, A961	N/A
	Sigua Beach, Cuba, A961	N/A
	Mirebelais, near Pont Beudet, Haiti, aspirated and reared from *Stenocereus hystrix*, 9.v.1982, ORV3, Starmer	102329
	Monte Cristi, Republica Dominicana, A940	N/A
	Sigua Beach, Cuba, A961	N/A
	Gonaives, Haiti, 9.v.1982, banana bait and reared from *Stenocereus hystrix*, ORV2, Wasserman & Heed WBH	102328
	Port Henderson, Jamaica, banana bait, 23.xi.1983, ORV23, Thomas & Heed	102332
	Airport, Palisades, Jamaica, 22.xi.1983, ORV23, Heed & Wasserman	102331
	Monte Cristi, Republica Dominicana, under *Opunita monoliformis*, A981	N/A
	20 km NW of Barahona, Republica Dominicana, A982	N/A
	27 km NE of Barahona, Republica Dominicana, A983	N/A
D. straubae	Guantanamo Bay, Cuba, A966	N/A
	Sigua Beach, Cuba, A961	N/A
	Navassa Island, west of Hispanola, banana baited and reared from *Opuntia stricta*, 21.xi.1983, ORV22, Johnson & Wasserman	102353
	Fond Parisien, Haiti, banana bait and reared from *Opuntia moniliformis*, 7.v.1982, ORV1, Wasserman & Heed	102352
	Monte Cristi, Republica Dominicana, A981	N/A
	27 km NE of Barahona, Republica Dominicana, A983	N/A
Outgroup		
D. mulleri	Cayman Brac, baited and reared from Opuntia stricta pads, 26.xi.1983, ORV27, Johnston & Benado	N/A
	Gonaives, Haiti, A942	N/A

Table 2. *Gene regions and primer sequences used in this study*

Gene region	Primer sequence	Size (bp)	Nature of variation
Acetylcholinesterase (*Ace*)	BAKER and DESALLE 1997	180	SNPs CT repeats
mastermind (*mam*)	BROWER and DESALLE 1994	190	SNP and CAG indel
vestigial (*vg*)	BROWER and DESALLE 1994	255	SNPs AC, GA, AT repeats three indels
hunchback (*hb*)	BAKER and DESALLE 1997	440	SNPs three CAG repeats
mt AT rich	SIMON et al. 1994	320	SNPs

Sequence alignment and phylogenetic analyses

Multiple sequence alignments, generated in CLUSTAL were manually adjusted and exported to NEXUS file format (MADDISON et al. 1997). Phylogenetic analyses and further sequence manipulation were performed in PAUP*, version 4.0 (SWOFFORD 2002) or MacClade, version 4.0 (MADDISON and MADDISON 2000). The AT rich mtDNA alignments were trivial, as very few indels occurred in this gene region for the individuals we sampled. The four nuclear genes (*Ace*, *hb*, *mam* and *vg*) all contain variable numbers of short tandem repeats (Table 2). However, the regions flanking these repeat segments contained no indels and were trivial. The di- or tri-nucleotide repeats were aligned by adding gaps where necessary to reconcile them with the flanking regions. Indels were then coded as discrete characters and the character state data for the eleven hypervariable regions were added to the aligned sequence matrix. Gaps were treated as missing and were recoded as binary, presence/absence characters appended to the matrix (BROWER and DESALLE 1994).

Tables of diagnostic sites for population aggregation analysis (PAA; DAVIS and NIXON 1992) were compiled using PAUP* (SWOFFORD 2002) and MacClade (MADDISON and MADDISON 2000). Only those sites which diagnosed species or geographic populations within species were included. The match first function was used to indicate where nucleotide substitutions had occurred.

Each of the five gene regions was analyzed separately and in combination in order to investigate the interaction of the various gene regions in the cladistic analyses. Because not all individuals were sequenced for each individual gene region, we selected those which had been sequenced for three or more of the five genes for our simultaneous analysis. Distance matrices were calculated in PAUP* (SWOFFORD 2002) in order to identify taxa with identical sequences. These taxa were removed prior to phylogenetic analysis and subsequently reincluded as polytomies with the taxa to which they were identical. Phylogenetic support was measured by bootstrap proportions (FELSENSTEIN 1985, 1988) and decay indices (BREMER 1988). Decay indices were calculated using AUTODECAY (ERIKSSON 1998).

AMOVA and F_{ST} values

The AMOVA (Analyses of Molecular Variance; EXCOFFIER et al. 1992) feature in Arlequin, version 1.1 (SCHNEIDER et al. 2000), was used in order to determine the percent of the total variability that could be ascribed to among species variability, among populations within species variability and within population variability. Only single nucleotide polymorphisms (SNPs) were used in these comparisons, indels were excluded when calculating these population genetic parameters.

This program was also used to calculate F_{ST} (fixation index) values to ascertain the amount of gene flow occurring between populations. Between population F_{ST} values were considered significant at the 0.05 level. In some cases, laboratory stocks were used to represent given geographic locales. In these cases only a single sequence per gene per lab culture was determined. Such a sampling procedure can produce misleading results with respect to F_{ST}, as there can be no sequence differences "between" one sequence. Therefore, for our discussion of F_{ST} values, we refer only to those locales containing individuals that had been collected in the field (Table 1).

RESULTS

Sequence variability

Table 2 summarizes the nature of the variability in the five loci we examined. Single nucleotide polymorphisms (SNPs) were found in all sequences. Dinucleotide repeats were found in two regions (*vg* and *Ace*). Trinucleotide CAG repeats were found in both the *mam* and *hb* loci. Three of the studied regions (*Ace*, *hb* and *mam*) were only slightly variable (less than 10% of the sites in each gene show variability) in the three species examined. Two regions (AT-rich and *vg*) showed extreme levels of variation (greater than 20% of the sites variable). A comparison of the phylogenetic signal in these hypervariable regions to the signal found in SNPs suggests that there is very little historical information to be found in the hypervariable regions (analysis not shown). A tree generated

from the characters that were recoded from these dinucleotide and trinucelotide repeat regions was fully unresolved. In addition, there are no hypervariable regions that can be used in a PAA to diagnose any of the three species units examined in this study. The diagnosibility of the morphological designations of *D. mayaguana*, *D. parisiena*, and *D. straubae* was tested by examining the sequence data from the five loci generated in this study. Only a single position in all five loci (position 340 in the *hb* gene) diagnoses *D. mayaguana* as distinct from *D. straubae* and *D. parisiena*.

AMOVA and F_{ST}

Figure 2 shows the percent of among species, among population within species and within population variation for the five loci sampled in this study. The *mam* gene reveals the greatest among population within species variation ($>75\%$) with very little between species variation. The AT rich region is similar ($>65\%$ of the differentiation due to among population differences), but with somewhat more among species variation ($>25\%$). The high value in almost certainly due to the large degree of divergence of the Monte Cristi *D. straubae*. Both *vg* and *Ace* show large percentages of between species variation (*vg* $>57\%$ and *Ace* $>45\%$). However, these two regions differ greatly in the amount of variation attributable to within population differences (*vg* $=10\%$, *Ace* $>40\%$). The *hb* gene region reveals the greatest amount of within population variation ($>45\%$) with the variation contributed by among species differences to be $>38\%$. The *hb* and *Ace* gene regions appear to behave in a similar fashion with low variation among populations within a species and much higher values for both among species differences and within population differences. However, comparison of the behavior of the other three genes in terms of the partitioning of variation revealed no similarity.

Individual and combined phylogenetic analyses

Overall, there was very little structure provided from any individual gene. Three of the individual gene trees (*vg*, AT-rich and *hb*) recovered a monophyletic *D. mayaguana*. No individual analysis was able to resolve *D. straubae* or *D. parisiena* as monophyletic, although some *D. straubae* populations (Cuba and Monte Cristi, Republica Dominicana) are monophyletic in some analyses (available online at http//www.amnh.org).

A strict consensus tree was generated from the combined maximum parsimony analysis of all five loci (Fig. 1). As with the individual gene trees, the simultaneous analysis tree was also relatively unresolved. Although *D. parisiena* and *D. straubae* are not

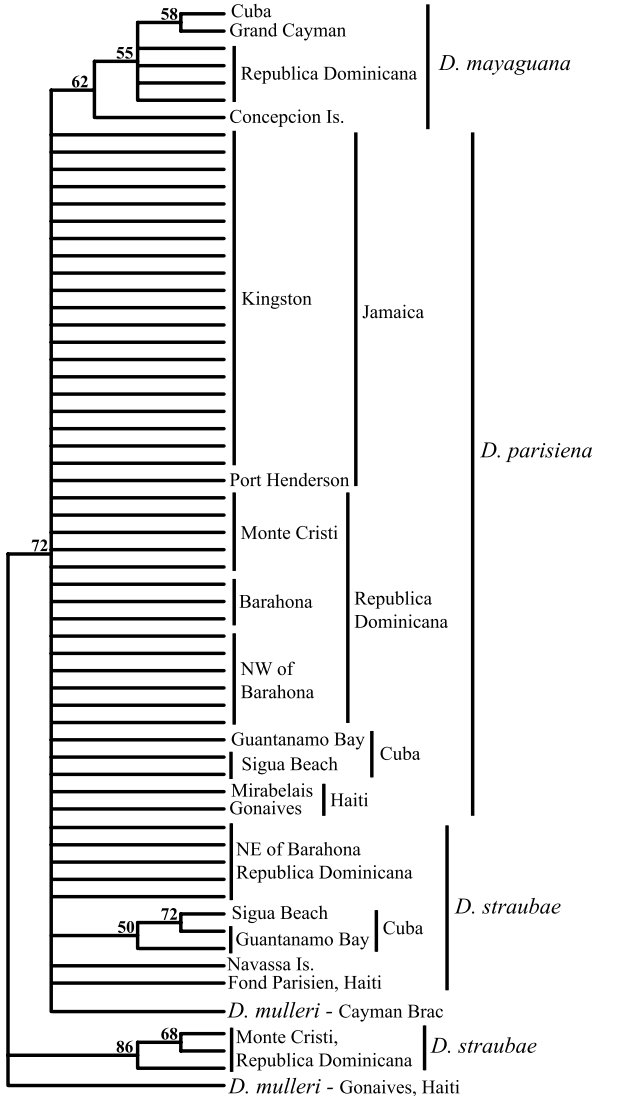

Fig. 1. Strict consensus of the simultaneous analysis. Numbers above the nodes are bootstrap proportions, those below the nodes are decay indices. Populations of each species are labeled at the right.

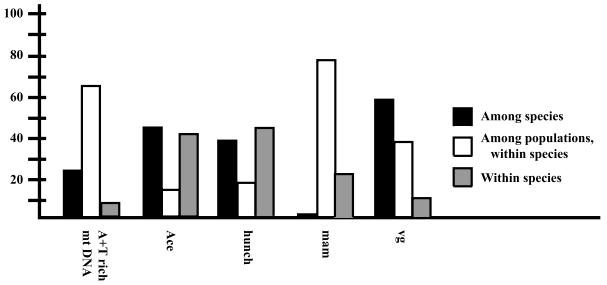

Fig. 2. Results of AMOVA analyses, showing among species, among population (within species) and within population variation.

monophyletic, *D. mayaguana* does form a well-supported clade (DI = 6). Based on these analyses, *D. straubae* and *D. parisiena* are not discernable genetically. In fact, no hierarchical structure between these two closely related taxa is evident. However, two other monophyletic assemblages are of interest (*D. straubae* populations from Cuba and Monte Cristi, Republica Dominicana; Fig. 1). The Monte Cristi population, represented by three individuals, form a strongly supported monophyletic group (DI = 8). The three individuals representing the Cuban populations of *D. straubae* are also monophyletic (DI = 5).

Another surprising result is that the outgroup species, D. *mulleri*, appears to be paraphyletic. Two populations, Cayman Brac and Haiti, were examined. The Cayman Brac individual is the sister taxon of the *D. mayaguana* clade, while the individual from Haiti is the outgroup. This suggests that *D. mulleri*, which is a widespread species both in the Caribbean and on the mainland, may actually comprise a number of different species. This is similar to the situation observed in *D. aldrichi*, another widespread member of the *repleta* group (WASSERMAN 1982). Further study will be required to resolve this issue.

DISCUSSION

Our data, in large part, agree with the previous work on the *mayaguana* triad, and suggests that this is a very recent group that is likely to be in the process of diversifying. The initial descriptions of the three species in the present study rely upon a suite of morphological characters (HEED and GRIMALDI 1991) which were used as guidelines to test the species designations with molecular data. The DNA sequences in the present study clearly diagnose *D. mayaguana*, although they do not corroborate the morphological characters used to designate individuals as *D. parisiena* or *D. straubae* (Fig. 1 and the PAA analysis). In fact, *D. parisiena* and *D. straubae* are paraphyletic with respect to *mayaguana* and highly undifferentiated relative to each other. It is likely that, while *D. mayaguana* has had sufficient time to accumulate unique nucleotide substitutions, *D. parisiena* and *D. straubae* have not. Interspecific matings between populations of *D. parisiena* and *D. straubae* suggest that, while these taxa are slightly reproductively isolated, they can still cross quite readily in the lab (WASSERMAN 1992). Of six crosses attempted, only a single mating (*D. parisiena* Gonaives × *D. straubae* Hispanola) was completely unsuccessful. All other crosses were able to produce viable F1 offspring, many of which were completely fertile.

The lack of DNA sequence characters diagnosing *D. straubae* and *D. parisiena* is not surprising in light of previous genetic and chromosomal studies involving these taxa, including the original descriptions which rely on a number of overlapping, quantitative morphological characters. Studies employing protein electrophoresis on the *mayaguana* subcluster species (HEED et al. 1990) indicated that there is little differentiation between *D. straubae* and *D. parisiena*. These data show that there are no fixed differences in allozyme frequency in either *D. straubae* or *D. parisiena* (HEED et al. 1990). Wagner trees based on allozyme frequencies always recovered a monophyletic *D. mayaguana*, but suggest that *D. straubae* is paraphyletic with respect to *D. parisiena*, regardless of the outgroup employed. Parsimony analyses coding allozyme bands and inversions as present or absent suggested that all three species are monophyletic, although HEED et al. (1990) noted that only a single additional step was required to obtain a paraphyletic *D. straubae*. WASSERMAN (1992) suggested that in most of the species in the *mulleri* subgroup, speciation has occurred without any detectable cytological changes. However, in two *mulleri* complex species, *D. parisiena* and *D. straubae*, a high degree of chromosomal inversion differentiation is exhibited relative to other closely related taxa, such as *D. mayaguana* and *D. mulleri* (WASSERMAN 1992). *Drosophila parisiena* is fixed for several inversions, including 5r, $2n^8$, $2l^8$, and $2m^8$. Its sibling species, *D. straubae*, lacks 5r and $2m^8$ and is polymorphic for $2n^8$ and $2l^8$.

Although *D. parisiena* and *D. straubae*, as currently defined, were not diagnosable with the molecular characters, some populations within each of these species were distinct from one another. Two populations of *D. straubae*, one from Monte Cristi, Republica Dominicana and another from Cuba, each form discrete lineages. This situation may warrant further examination of the ecology, genetics, and morphology of these populations and may even require designating these taxa as separate species.

The inability to diagnose species can be attributed to incomplete lineage sorting of ancestral polymorphisms. This phenomenon becomes more pronounced with larger effective population sizes and shorter time to a most recent common ancestor, such as is the case with the *mayaguana* triad. Lack of characters diagnosing *D. straubae* and *D. parisiena*, even with rapidly evolving gene regions, is due to the short time that these taxa have been isolated from one another and may perhaps be complicated by hybridization in the wild. Sampling additional characters (DAVIS and NIXON 1992), such as those genes or gene regions involved in speciation or ecological and morphological characters, may be the most appropriate tools to discover diagnostic characters. However, focusing on

the genes driving speciation is a daunting task, as they may be located at multiple loci and are likely to different for each speciation event. Furthermore, if hybridization is occurring in the wild, even at a low level, any hierarchical pattern will be obfuscated.

The best character delineating *D. straubae* and *D. parisiena* is likely an ecological one. *Drosophila straubae* is able to use a variety of host plants, including *Opuntia moniliformis*, *O. stricta*, and *Cephalocereus brooksianus* (HEED et al. 1990; HEED and GRIMALDI 1991). Even though other hosts are present in the same area, *D. parisiena* has only been reared from *Stenocereus hystrix*. The highly specific host use by *D. parisiena* is one of the only indicators that this taxon is distinct from *D. straubae* and may point to host specialization as the primary factor leading to the formation of these two species. However, it should be noted that this ecological attribute has not been applied as extensively as the DNA markers or allozyme markers mentioned above and may also eventually not be diagnostic. That the ecological larval substrate characters potentially diagnose the two closely related species in this study as distinct, and the DNA sequence characters do not, reinforces the importance of considering all character information in delineating species (LEGGE et al. 1996) when establishing phylogenetic boundaries for species.

REFERENCES

Baker R and DeSalle R, (1997). Multiple sources of character information and the phylogeny of Hawaiian drosophilids. Syst. Biol. 46: 654–673.

Bettler D, Pearson S and Yedvobnick B, (1996). The nuclear protein encoded by the Drosophila neurogenic gene mastermind is widely expressed and associates with specific chromosomal regions. Genetics 143: 859–875.

Bremer K, (1988). The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42: 795–803.

Brower A and DeSalle R, (1994). Practical and theoretical considerations for choice of DNA sequence region in insect molecular systematics, with a short review of published studies using nuclear gene regions. Annu. Rev. Entomol. 87: 702–716.

Davis JI and Nixon KC, (1992). Populations, genetic variation, and the delineation of phylogenetic species. Syst. Biol. 41: 421–435.

Durando CM, Baker RH, Etges WJ, Heed WB, Wasserman M and DeSalle R, (2000). Phylogenetic analysis of the repleta species group of the genus Drosophila using multiple sources of characters. Mol. Phylogenet. Evol. 16: 296–307.

Eriksson T, (1998). AutoDecay ver. 4.0 (program distributed by the author).

Etges WJ, Armella MA, O'Grady PM and Heed WB, (2001). Two new species of Drosophila (Diptera; Drosophilidae) in the repleta group from Mexico. Ann. Entomol. Soc. Am. 94: 16–20.

Excoffier L, Smouse PE and Quattro JM, (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: applications to human mitochondrial DNA restriction data. Genetics 131: 479–491.

Felsenstein J, (1985). Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

Felsenstein J, (1988). Phylogenies from molecular sequences: inference and reliability. Annu. Rev. Genet. 22: 521–565.

Fournier D, Karch F, Bride JM, Hall LM, Berge JB and Spierer P, (1989). Drosophila melanogaster acetylcholinesterase gene. Structure, evolution and mutations. J. Mol. Biol. 210: 15–22.

Heed WB and Grimaldi DA, (1991). Revision of the morphocryptic, Caribbean mayaguana species subcluster in the Drosophila repleta group (Diptera: Drosophilidae). Am. Mus. Novit. 2999: 1–15.

Heed WB, Sanchez A, Armengol R, and Fontdevila A. (1990). Genetic differentiation among island populations and species of cactophilic Drosophila in the West Indies. In: Ecological and evolutionary genetics of Drosophila. (Barker, Starmer, and MacIntyre, eds.). Plenum Press, New York.

Lewis DL, Farr CL, Farquhar AL and Kaguni LS, (1994). Sequence, organization, and evolution of the A + T region of Drosophila melanogaster mitochondrial DNA. Mol. Biol. Evol. 11: 523–538.

Maddison DR, Swofford DL and Maddison WP, (1997). NEXUS: an extensible file format for systematic information. Syst. Biol. 46: 590–621.

Maddison WP and Maddison DR, (2000). MacClade, Version 4: Analysis of phylogeny and character evolution. Sinauer Associates, Sunderland Massachusetts.

Newfield SJ, Smoller DA and Yedvobnick B, (1991). Interspecific comparison of the unusually repetitive Drosophila locus mastermind. J. Mol. Evol. 32: 415–420.

Schneider S, Roessli D, and Excoffier L, (2000). Arlequin ver. 2.000. A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H and Flook P, (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annu. Rev. Entomol. 87: 651–701.

Swofford DL, (2002). PAUP: Phylogenetic analysis using parsimony, Ver. 4.0 Smithsonian Institution, Washington DC.

Treier M, Pfeifle C and Tautz D, (1989). Comparison of the gap segmentation gene hunchback between Drosophila melanogaster and Drosophila virilis reveals novel modes of evolutionary change. EMBO J. 8: 1517–1525.

Wasserman M, (1982). The repleta species group. In: The genetics and biology of Drosophila, Vol. 3b, Ashburner, Carson and Thompson eds. Academic Press, NY.

Wasserman M, (1992). Cytological evolution of the Drosophila repleta species group, pp. 455–552. In C. B. Krimbas and J. R. Powell [eds.], Drosophila inversion polymorphism. CRC, Boca Raton, FL.

Williams JA, Bell JB and Carroll SB, (1991). Control of Drosophila wing and haltere development by the nuclear vestigial gene product. Genes Dev. 5: 2481–2495.