The CO-I and CO-II Region of Honeybee Mitochondrial DNA: Evidence for Variation in Insect Mitochondrial Evolutionary Rates

R. H. Crozier,* Y. C. Crozier,* and A. G. Mackinlay†
Schools of *Biological Science and †Biochemistry, University of New South Wales

The sequence of a region of honeybee (Apis mellifera ligustica) mitochondrial DNA, which contains the genes for cytochrome c oxidase subunits I and II (CO-I and CO-II) and inferred genes for tRNA^{Asp}, tRNA^{Leu}_{UR}, tRNA^{Lys}, and tRNA^{Trp}, is presented. The region includes the segment previously identified as incurring a length increase in some other bee strains, including Africanized bees. The sequence information of this study and of that by Vlasak et al. shows that several shifts of tRNA genes have occurred between Apis and Drosophila, but shifts of other kinds of genes have yet to be demonstrated. The CO-I and CO-II gene sequences are both more A+T rich than are the corresponding Drosophila genes. Parsimony analyses using the mouse and Xenopus sequences as outgroups show significantly more amino acid substitutions on the branch to Apis (120) than on that to Drosophila (44), indicating a difference in the long-term evolutionary rates of hymenopteran and dipteran mtDNA.

Introduction

The mitochondrial DNA (mtDNA) of animals is a relatively small molecule, usually ~16 kb long (Avise et al. 1987; Moritz et al. 1987), although it ranges from 14 kb to ~39 kb in length (Snyder et al. 1987; Wolstenholme et al. 1987). It generally encodes 13 proteins, two ribosomal RNAs, and 22 tRNAs and has a region controlling replication and devoid of other known functions. One exception is the mtDNA of nematodes, in which the ATPase8 gene has been lost (Wolstenholme et al. 1987). Nematodes also present unusual tRNAs, all of which have been inferred to lack the TvC loop (Wolstenholme et al. 1987).

The mtDNA of one insect, Drosophila yakuba, has been sequenced in full (Clary and Wolstenholme 1985), and partial sequences are known from D. virilis (Clary and Wolstenholme 1987), D. melanogaster (de Bruijn 1983; Satta et al. 1987; Garessi 1988), the mosquito Aedes albopictus (HsuChen et al. 1984a, 1984b), the locust Locusta migratoria (McCranken et al. 1987; Uhlenbusch et al. 1987), and the honeybee Apis mellifera (Vlasak et al. 1987). These studies have indicated that the order of major genes has changed between phyla, that tRNA positions change between members of the same insect order (Drosophila and Aedes), and that insect mtDNAs are very A+T rich.

In vertebrates, mtDNA has long been regarded as evolving much more rapidly than single-copy nuclear DNA (scnDNA), as judged on the basis of mtDNA sequence differences.

1. Key words: Apis, cytochrome c oxidase subunits I and II, evolutionary rates, honeybee, mitochondrial DNA sequences.

Address for correspondence and reprints: Dr. R. H. Crozier, School of Biological Science, University of New South Wales, P.O. Box 1, Kensington, New South Wales 2033, Australia.

© 1989 by The University of Chicago. All rights reserved.
0737-4038/89/0604-007$02.00

399
comparisons and scnDNA heteroduplex formation (Brown et al. 1979; Moritz et al. 1987). By contrast, studies on echinoids (Vawter and Brown 1986) and Drosophila (Powell et al. 1986) indicate roughly similar rates of divergence for scnDNA and mtDNA. Vawter and Brown (1986) and Moritz et al. (1987) attribute these relative differences in evolutionary rates between mtDNA and scnDNA to differences in scnDNA rates, suggesting that the mtDNA rates are relatively invariant. Nuclear genes have indeed been found to vary significantly between various groups in evolutionary rates (Wu and Li 1985; Britten 1986; Li and Tanimura 1987; Lake 1988; also see Ochman and Wilson 1987), but Gillespie (1986) found evidence for some between-lineage variation in evolutionary rates for mitochondrial genes also.

Honeybees have been surveyed for mitochondrial (Moritz et al. 1986; Smith 1988; Smith and Brown 1988) and nuclear (Hall 1986) RFLPs. We here extend the knowledge of honeybee mtDNA sequence, providing information useful for general evolutionary studies on insects, and we show that the rates of mtDNA evolution have differed greatly between these lineages.

Material and Methods

Bees

We used bees from a domestic hive (R. Clark) identified morphometrically by Dr. B. Oldroyd as being “Apis mellifera ligustica of commercial origin.”
Fig. 2.—Sequence of the honeybee mtDNA section reported here, showing the genes identified and, for the protein-encoding genes, the amino acid sequences inferred from the *Drosophila* mitochondrial genetic code.
FIG. 3.—Comparison of the CO-I:CO-II gene regions of *Drosophila yakuba* and *Apis mellifera*. The genes for subunits I and II of cytochrome c oxidase are shown, as are the tRNAs in this region (C = cysteine; D = aspartate; K = lysine; L = leucine; W = tryptophan; Y = tyrosine). Shifts of several tRNAs are evident between these two insect species.

mtDNA Preparation

We used a method based on the methods of Davis et al. (1980), Lansman et al. (1981), and D. R. Smith and W. M. Brown (personal communication). In brief, 9 g worker thoraces were cut in an Ultra Turax® at low speed and were homogenized in a total volume of 300 ml 0.1 M Tris-HCl pH 7.5 buffer containing 0.25 M sucrose, 0.01 M NaCl, and 0.05 M ethylenediaminetetraacetate (EDTA). Mitochondria were collected by differential centrifugation and were resuspended in 12 ml of pH 8.0, 0.2 M Tris-HCl buffer containing 0.1 M EDTA and 1% sodium dodecyl sulfate. Proteinase K was added to a final concentration of 90 µg/ml, and the sample was incubated at 60°C for 2 h and was spun for 10 min at 16,000 g. Two milliliters 5 M potassium acetate pH 6 buffer was added to the supernatant, and the sample was kept on ice for 30 min, then spun at 12,000 g for 10 min. DNA was precipitated from the supernatant by using isopropanol (0.6 vol) and was then subjected to CsCl-ethidium bromide centrifugation in a vertical rotor (Beckman Vti65) at 50,000 x-g for 14 h. Linear and closed circular DNA bands recovered from the gradient were extracted with CsCl-saturated isopropanol to remove the ethidium bromide, were dialyzed against TE buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA), and were precipitated with ethanol. DNA from the lower circular band was used for cloning.

Cloning

*EcoRI* and *BclI* fragments were cloned into pUC8 and pUC18, respectively, using *Escherichia coli* strain JM101 as host. In addition, mtDNA was cut with *AccI*, was end-filled using the Klenow fragment of DNA polymerase I, and then was cut with *BclI* for cloning into pUC18. *HindIII-BclI* fragments were also cloned into pUC18. Plasmids were identified as containing mtDNA inserts through hybridization with labeled honeybee mtDNA and through sizing following gel electrophoresis and restriction digestions. These clones represent all of the honeybee mtDNA except for a 520-bp *AccI* fragment.

Sequencing

The *BclI*, *HindIII-BclI*, and *BclI-AccI* clones used for sequencing are indicated in figure 1. For sequencing by the dideoxy chain-termination method of Sanger et al. (1980), pUC clones were recloned into either m13mp18 or m13mp19 or into both, using *Escherichia coli* JM101 as host. In the case of the *BclI* 1.12 insert, the chain was extended in both directions by using internal oligonucleotide primers.
FIG. 4.—Sequences of four honeybee tRNA genes, folded into the inferred configurations of the corresponding tRNAs. Nucleotide differences between *Apis* and *Drosophila* are indicated by circles; deletions from the honeybee sequence relative to that of *Drosophila* are indicated by arrowheads; and additions to the honeybee sequence are indicated by parentheses.
Analyses

The sequence information was analyzed using the Nucleic Acids Analysis System package (Genesearch, Broadbeach, Queensland) run on an Apple II microcomputer, the DNA Strider 1.0 package (C. Marck, Commissariat a l'Energie Atomique, France) run on a Macintosh SE, and the GENEUS package (Harr et al. 1986) implemented on a Digital Equipment Corporation VAX minicomputer.

Phylogenetic analyses involved either the maximum-likelihood method of Felsenstein (1981), carried out using the PHYLIP 3.01 package (Felsenstein 1985) implemented on an IBM 3090 computer, or parsimony criteria.
CO-I and CO-II Region of Honeybee Mitochondrial DNA

CO-II

**BEE**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISTWFMFMFQESNSYYADLNISFHNMVMMIIIMISTLTVIYILDFLMKNFSLNLFKHHNM</td>
<td>CO-II Sequence for BEE</td>
</tr>
<tr>
<td>M...ANGL.D.A.PLMEQ..F..DHALL.LV..TV.VG.LMFM..F.NVY.R...HGL</td>
<td>CO-II Sequence for DRO</td>
</tr>
<tr>
<td>MAHFSLQG..DAA.PIMEE.LH..DHTL.AVFL..VL..TIMMTT.LT.TN.MDAGE</td>
<td>CO-II Sequence for XEN</td>
</tr>
<tr>
<td>MAYPFQGL.DAT.PIMEE.MN..DHTL..VFL..S.VL..SLMLTT.LHTSTMDAGE</td>
<td>CO-II Sequence for MOU</td>
</tr>
</tbody>
</table>

**DRO**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEIIWITPIIIIIIICFPSLKILYLDIEIVNEFFSIKSIGHQWWSYPEFNNIEFDS</td>
<td>CO-II Sequence for DRO</td>
</tr>
<tr>
<td>M..L.A..F.A...RL..L..NE.SVFL..SD..</td>
<td>CO-II Sequence for XEN</td>
</tr>
</tbody>
</table>

**XEN**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEIIWITPIIIIIIICFPSLKILYLDIEIVNEFFSIKSIGHQWWSYPEFNNIEFDS</td>
<td>CO-II Sequence for DRO</td>
</tr>
<tr>
<td>M..L.A..F.A...RL..L..NE.SVFL..SD..</td>
<td>CO-II Sequence for XEN</td>
</tr>
</tbody>
</table>

**MOU**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEIIWITPIIIIIIICFPSLKILYLDIEIVNEFFSIKSIGHQWWSYPEFNNIEFDS</td>
<td>CO-II Sequence for DRO</td>
</tr>
<tr>
<td>M..L.A..F.A...RL..L..NE.SVFL..SD..</td>
<td>CO-II Sequence for XEN</td>
</tr>
</tbody>
</table>

**FIG. 5.**—Comparisons of the inferred CO-I and CO-II sequences from *Apis*, *Drosophila*, *Xenopus*, and *Mus*. Differences between the *Apis* and other sequences are indicated; otherwise the sequences are the same, except where gaps, indicated by dashes (−), have been introduced to improve the alignment.

**Results and Discussion**

**Genome Organization**

Figure 1 shows a restriction map derived during the course of this work. The restriction-site placements agree with those determined by Smith and Brown (1988). The figure also shows the placement of the sequence reported here, as well as the location of the gene for the large subunit of ribosomal RNA sequenced by Vlasak et al. (1987); this position was identified by (1) the high similarity between the published sequence and the sequence obtained by us (to be reported separately) and (2) the location of EcoRI and PstI sites in the sequence reported by Vlasak et al. Figure 2 gives the sequence determined by us and shows the positions of the inferred genes and, for the protein-encoding genes, the amino acids inferred according to the *Drosophila* mitochondrial code (Clary and Wolstenholme 1983; de Bruijn 1983).

Figure 3 gives a comparison between the gene orders of the segment reported here and the homologous region of the mtDNA from *D. yakuba*. Limited sequence from a protein-encoding gene at the 3′ end of the sequence (fig. 1) enables tentative identification of this gene as that for ATPase8; such a placement would be the same as that seen in *Drosophila* (Clary and Wolstenholme 1985).

There is as yet no evidence for changes in the order of major genes between *Drosophila* and *Apis*, but, of the four honeybee tRNA genes sequenced, thr3c (those for lysine, cysteine, and tyrosine) have different positions compared with those in *Drosophila*.

The region marked “?” in figure 3 includes sequences that can be folded into configurations roughly consistent with tRNA genes; positive identification of these sequences awaits sequencing of the rest of the mtDNA molecule. This region falls in the 1.12-kb BclI fragment and is 270 bp longer in the mtDNA of Africanized bees (*Apis mellifera scutellata*) (Smith and Brown 1988) and *A. m. mellifera* and *A. m. capensis* (Smith 1988) than in races of *A. m. ligustica* surveyed so far with RFLPs.
Table 1
Similarities between the Nucleotide and Inferred Amino Acid Sequences for the CO-I and CO-II Genes of Apis, Drosophila, Xenopus, and Mus

<table>
<thead>
<tr>
<th></th>
<th>Apis</th>
<th>Drosophila</th>
<th>Xenopus</th>
<th>Mus</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO-I:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apis</td>
<td></td>
<td>0.64</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>Drosophila</td>
<td>0.74</td>
<td>0.74</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Xenopus</td>
<td>0.66</td>
<td>0.70</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Mus</td>
<td>0.67</td>
<td>0.72</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>CO-II:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apis</td>
<td></td>
<td>0.60</td>
<td>0.64</td>
<td>0.57</td>
</tr>
<tr>
<td>Drosophila</td>
<td>0.68</td>
<td>0.60</td>
<td>0.57</td>
<td>0.75</td>
</tr>
<tr>
<td>Xenopus</td>
<td>0.60</td>
<td>0.64</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Mus</td>
<td>0.60</td>
<td>0.64</td>
<td>0.74</td>
<td></td>
</tr>
</tbody>
</table>

NOTE.—Amino acid similarities are given in the upper-right halves of the comparisons, and nucleotide similarities are given in the lower-left halves. Similarities were determined for homologous positions only; gaps were ignored.

Genetic Code

Comparisons between the Apis genes known so far and those of Drosophila provide no evidence for any difference in the mitochondrial genetic codes of these insects.

tRNA Genes

The inferred sequences of the tRNA genes for aspartate, leucine (UUR), lysine, and tryptophan are shown in figure 4 in the shapes presumed for the corresponding tRNAs. The extent of the similarity of the Apis genes to those of Drosophila varies, from the tRNA<sup>Asp</sup> gene with one nucleotide deletion, two additions, and five substitutions, to that for tRNA<sup>Trp</sup>, with one deletion, seven additions, and fourteen substitutions.

The anticodon inferred for lysine is UUU as in Xenopus, rather than CUU as in Drosophila.

The CO-I and CO-II Genes

The amino acid sequences of cytochrome c oxidase subunits I and II, inferred from the corresponding nucleotide sequences according to the Drosophila genetic code (Clary and Wolstenholme 1983, 1985; de Bruijn 1983) are shown in figure 5, together with the corresponding sequences from Drosophila, Xenopus, and Mus.

The sequence known so far indicates that the mtDNA of Apis is even more A+T rich than that of Drosophila. The A+T content of the bee CO-I gene is 76% compared with 68% for Drosophila, and the corresponding figures for the CO-II gene are 80% and 75%.

The hydrophobicity profiles of the Apis proteins are very similar to those of their Drosophila and Xenopus counterparts, as determined from using hydropathy plots (Kyte and Doolittle 1982). Conservation of function is therefore indicated.

The similarities of the nucleotide sequences and the inferred amino acid sequences of the two genes are shown in table 1. These data suggest either that the Apis genes have evolved more rapidly than those of the other species or that current views of the
Table 2
Amino Acid Replacements in Cytochrome c Oxidases I and II for the Terminal Branches Leading to *Apis*, *Drosophila*, *Xenopus*, or *Mus*, Inferred According to the Standard Phylogeny

<table>
<thead>
<tr>
<th></th>
<th><em>Apis</em></th>
<th><em>Drosophila</em></th>
<th><em>Xenopus</em></th>
<th><em>Mus</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>CO-I</td>
<td>78</td>
<td>22</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>CO-II</td>
<td>42</td>
<td>22</td>
<td>14</td>
<td>23</td>
</tr>
</tbody>
</table>

relationships of these taxa are in error. However, elementary parsimony analyses using informative amino acid positions give convincing support for the traditional phylogenetic view. Thus, 40 positions support the traditional view (the insects grouped together), as against a total of seven positions supporting the other two possible unrooted trees.

The impression, from the similarity values, of a greater evolutionary rate for honeybee mtDNA than for that of the other taxa can be tested using the inferred amino acid sequences in figure 5. These sequences can be used to infer the amino acid replacements occurring in the terminal branches leading to the individual taxa, and the numbers for the two genes are shown in table 2. For example, given the phylogenetic arrangement [*(Apis, Drosophila) (Xenopus, Mus)*], if a given position has amino acids ABBB or ABBC, a substitution on the *Apis* terminal branch is indicated (and, in the latter case, on that for *Mus* as well).

χ² Tests reject the possibility that *Apis* has had the same number of amino acid replacements since it diverged from *Drosophila as Drosophila* has had (*P < 0.001*). Maximum-likelihood (Felsenstein 1981) estimation of the branch lengths for the same suite of taxa and using the CO-I and CO-II DNA sequences also supports a higher rate of evolution in bees than in flies: for CO-I the length of the *Apis* branch is 0.12 units (estimated substitutions per site) and that for *Drosophila* is 0.07 units, with the corresponding figures for the phylogeny derived from the CO-II sequences being 0.14 and 0.10.

The reason for the greater speed of *Apis* mtDNA evolution vis-à-vis *Drosophila* mtDNA evolution is unclear. Wu and Li (1985) attributed the faster rate of rodent sncDNA evolution to the fact that rodent generation times are shorter than primate generation times. However, modern honeybees have generation times much longer than those of modern fruit flies [although they also have much smaller effective population sizes (Crozier 1979)]. Given that the lineages leading to *Apis* and to *Drosophila* have probably been separated for ~280 Myr (Carpenter and Burnham 1985), it is likely that there have been many changes in life-history parameters during that time. For example, the long-lived *Apis* bees probably arose only about ~100 Mya (Michener and Grimaldi 1988). The role of life-history parameter differences—as against differences in error-correcting capabilities—in mediating long-term evolutionary rates in insects is thus difficult to evaluate.

Sequence Availability

These sequences have been deposited in GenBank under accession number M23409.
Acknowledgments

We thank the Australian Research Council for grants to R.H.C. in support of this work; R. Clark [University of New South Wales (UNSW)] and G. Mulder (Hawkesbury Agricultural College) for donating bees; B. Oldroyd (Victorian Plant Protection Institute) for identifying them; C. Hawkins, I. Alexander, and J. Bonsing (all UNSW) for expert technical advice; and, for helpful suggestions on the manuscript, S. Davis (Texas A&M U.), C. Moritz (University of Queensland), D. R. Wolstenholme (University of Utah), two anonymous reviewers, and M. Nei.

LITERATURE CITED


CO-I and CO-II Region of Honeybee Mitochondrial DNA 411


MASATOSHI NEI, reviewing editor

Received November 23, 1988; revision received February 27, 1989