J Mol Evol (2003) 57:170–180
DOI: 10.1007/s00239-003-2464-z

JOURNAL OF MOLECULAR EVOLUTION

Exploring the Evolutionary History of the Alcohol Dehydrogenase Gene (*Adh*) Duplication in Species of the Family Tephritidae

George N. Goulielmos,[1] Michael Loukas,[1] George Bondinas,[1] Eleftherios Zouros[2,3]

[1] Department of Genetics, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece
[2] Department of Biology, University of Crete, Vassilika Vouton, 714 09 Iraklion, Crete, Greece
[3] Institute of Marine Biology of Crete, Iraklion, Crete, Greece

Received: 3 October 2002 / Accepted: 12 February 2003

Abstract. In the olive fruit fly *Bactrocera oleae* and the med fly *Ceratitis capitata* previous studies have shown the existence of two *Adh* genes in each species. This observation, in combination with the former finding that various *Drosophila* species of *virilis* and *repleta* group encode two isozymes of ADH which are the result of a gene duplication, challenged us to address a scenario dealing with the evolutionary history of the *Adh* gene duplication in Tephritidae. In our lab we proceeded to the cloning and sequence analysis of *Adh* genes from more tephritid species, a prerequisite for further study of this issue. Here we show that phylogenetic trees produced from either the nucleotide or the amino acid sequences of 14 tephritid *Adh* genes consisted of two main clusters, with *Adh* sequences of the same "type" grouping together (i.e., *Adh1* sequences form a cluster and *Adh2* sequences form a second one), as expected if there was one duplication event before speciation within the family Tephritidae. We used the amount of divergence between the two isozymic forms of *Adh* of the species carrying both *Adh1* and *Adh2* genes to obtain an estimate of the age of the duplication event. Interestingly, our data again support the hypothesis that the duplication of an ancestral *Adh* single gene in the family Tephritidae occurred before the emergence of the genera *Bactrocera* and *Ceratitis*, thus suggesting that *Adh* duplication was based on a prespeciation rather than a postspeciation event that might have involved two independent duplication events, one in each of the two genera.

Key words: Alcohol dehydrogenase gene — Gene duplication — Tephritidae

Introduction

The principal function of the enzyme alcohol dehydrogenase (ADH; EC 1.1.1.1) in insect metabolism is the catalysis (in its homodimeric form) of the reversible conversion of various alcohols generated by microbial fermentation in larval and adult feeding sites to their corresponding aldehydes and ketones. The ADH enzyme system has been studied in several species of *Drosophila* as well as in *Ceratitis capitata* (medfly) and *Bactrocera oleae* (olive fruit fly), with the main aim of defining their functional and evolutionary relationships, detecting biochemical differences of the enzyme among different species or isozymes from the same species, and analyzing the molecular organization of the respective *Adh* gene(s).

Drosophila ADH was first purified by Sofer and Ursprung (1968) and it was shown that the best substrates are secondary rather than primary alcohols (Benach et al. 1999). *Drosophila* ADH belongs to a broad, heterogeneous family of alcohol dehydrogenases, called short-chain dehydrogenases/reductases (SDR) (Jornvall et al. 1995). The refined three-dimensional crystal structure of ADH of *D. leb-*

Correspondence to: George N. Goulielmos; *email:* goulielmos@imbb.forth.gr

anonensis was recently solved at 1.9-Å resolution (Benach et al. 1998, 1999) and it represents a unique example in a dipteran species enzyme, belonging in the SDR family, with known structural features. The *Adh* locus in most *Drosophilidae* is organized as a single gene transcribed from two spatially and temporally regulated promoters. Accordingly, in *D. melanogaster* and *D. simulans* the *Adh* gene has the same transcriptional organization in terms of the position of the two promoters and the three introns, with the distal and proximal transcripts being differentially expressed during development but producing identical proteins (Benyajati et al. 1983; Savakis et al. 1986). Since the *Adh* gene is expressed from two tandem and temporally regulated promoters in *D. melanogaster* as well as in *Scaptodrosophila lebanonensis* (Juan et al. 1994), two species that belong in different genera, this organization was taken to be the ancestral phylogenetic structure. However, some species belonging to subgenus *Drosophila*, such as *D. mulleri* and *D. hydei* (*repleta* group) or *D. virilis* (*virilis* group), encode isozymes of ADH which are the result of a gene duplication (Begun 1997). In *D. virilis* and in some species of the *virilis* group, the *Adh* locus has been duplicated and each copy consists of a functional *Adh* gene with two temporally regulated promoters (Nurminsky et al. 1996). In species of the *repleta* group there are also two *Adh* loci, each of them possessing its own promoter, expressed in both larvae and adults (Fischer and Maniatis 1985; Sullivan et al. 1989). Furthermore, species of the Hawaiian group show a single *Adh* gene with two temporally regulated promoters (Rowan and Dickinson 1986, 1988). Finally, in *D. funebris* the single *Adh* gene is transcribed by a single promoter, in both larva and adult, and exhibits qualitative and quantitative species-specific differences in tissue distribution (Amador et al. 2001). ADH activity has been identified in the fat bodies, intestines, and Malpighian tubules of both larvae and adults of most species but appreciable variation of tissue localization of ADH exists in different species (Sullivan et al. 1989).

The family Tephritidae includes more than 4000 species arranged in about 500 genera throughout the world. This family is the most important group of agricultural pests of all fly families (Foote et al. 1993) and studies of tephritids have contributed to development of our understanding of speciation and evolutionary biology (Feder et al. 1988; McPheron et al. 1988). Our knowledge of ADH in tephritids is mostly concentrated on *C. capitata* and *B. oleae*. The ADH system of these species is of major interest because of its potential use in the biological control of the insects (Zouros et al. 1982; Robinson and MacLeod 1993). *Ceratitis capitata* has two *Adh* genes, *Adh1* and *Adh2*, tightly linked (0.49 cM) on the end of the left arm of the second chromosome, suggesting that these genes are products of gene duplication, followed by subsequent divergence. *Adh1* is expressed mainly in muscle, and *Adh2* in fat body and ovary (Gasperi et al. 1992, 1994; Malacrida et al. 1992).

The cloning of the *Adh* of *B. oleae* performed recently and the subsequent sequence analysis also revealed two genes, *Adh1* and *Adh2* (Goulielmos et al. 2001). The authors attempted to address a firm scenario dealing with the evolutionary history process of the *Adh* gene duplication in tephritids. Phylogenetic trees based on amino acid sequences grouped *Adh1* of *B. oleae with Adh1* of *C. capitata* and *Adh2* of *B. oleae* with *Adh2* of *C. capitata*, thus suggesting that the duplication occurred before the splitting of the two species from their common ancestor. However, similar trees produced from nucleotide sequences grouped *Adh1* from *B. oleae* with *Adh2* of *B. oleae* and *Adh1* of *C. capitata* with *Adh2* of *C. capitata*, thus suggesting that a duplication event occurred within each species after speciation. Therefore, further results were needed for the resolution of the contradiction appearing between amino acid- and nucleotide-based trees.

A more exact picture is probably possible only by including either information on variation within species or more sequences from various species belonging to *Ceratitis* and *Bactrocera* genera. The first issue was tackled recently (Goulielmos et al. 2003). Thus, the purpose of this paper is threefold: (1) to clone and sequence the *Adh* loci of various species classified in the genera *Bactrocera* and *Ceratitis*, a prerequisite for further studying of the molecular basis of the duplication events observed in several tephritid species; (2) to analyze the putative phylogenetic relationships among the *Adh* sequences of tephritids; and (3) to examine the significance of these results in our attempt to elucidate the evolution of a common ancestral *Adh* gene in the family Tephritidae.

Materials and Methods

Ceratitis *Species*

The samples of the species *Ceratitis cosyra* (mango fruit fly) were collected from farms near Nairobi (Kenya) in spring of 2000. In particular, pupae were collected from different host fruits. *Ceratitis rosa* individuals originating from Kenya were provided by Gerald Franz of the Entomology unit of the IAEA, Seibersdorf, Austria, who also provided us with a *Ceratitis fasciventris* sample. These specimens were obtained from wild collections made in April of 1999. Fifteen *C. cosyra*, 15 *C. fasciventris,* and 15 *C. rosa* individuals were used for DNA extraction and further analysis.

Bactrocera *Species*

Five *Bactrocera* species were considered in this study. The *B. oleae* species used is the colony kept in our laboratory (Agricultural University of Athens, Greece) for about 20 years (Cosmidis 1995). The samples of *B. dorsalis* (oriental fruit fly), *B. cucurbitae* (melon

fly), *B. scutellatus*, and *B. tryoni* (Queensland fruit fly) were collected in different countries. The *B. dorsalis* flies used (15 adults) were members of a laboratory strain, originated from flies collected in Amami Oshima (Kagoshima, Japan) in 1979 and reared in the laboratory ever since. The *B. cucurbitae* flies used (15 adults) were members of a laboratory strain originated from flies captured in Amami Oshima (Kagoshima, Japan) in 1987 and, therefore, kept in the laboratory. The *B. scutellatus* sample used was represented by 15 adults, collected on May 2000 in Tsukuba (Ibaraki, Japan). A total of 12 adult individuals of *B. tryoni* was used in this study, collected from Australian tropical rain forests and kept as a laboratory colony at the University of Sydney. All specimens of the same species were used together for extraction of DNA.

Amplification of Genomic Adh1 *and* Adh2 *Genes*

Preparation of genomic DNA from each of the species was done according to the protocol described by Holmes and Bonner (1973). The cDNA *Adh* sequences of *C. capitata* (GenBank accession Nos. Z 30194 and Z 30195) and the genomic DNA *Adh* sequences of *B. oleae* (EMBL Database Accession Nos. AJ277834 and AJ277835) were used to design primers for PCR amplification of the corresponding *Adh* genomic fragments of *Ceratitis* and *Bactrocera* species. The upstream primers 5′-ACGCGTCGACG-AAATTCATGAG(C/T)TTGGCIGGIAAAAA(C/T)G-3′ and 5′-ACGCGTCGACGAATTCATGGGTTTGAGCGGCAAAAAT-3′ and the downstream 5′-ACGGAGCTC(G/A)TAIGTGGG(T/C) TCCCA(G/A)TAIAC-3′ and 5′-CGAGCTCGGATCCCTAG TTTGAATGTGGGTTGCCA-3′ were used to generate the *Adh1* and *Adh2* products, respectively. In any case the amplification was carried out using high-fidelity conditions (Kwiatowski et al. 1991). To this end, *Pwo* polymerase (a proofreading enzyme) (Boehringer–Mannheim) was used to get amplification products of a high fidelity. In the case of *Adh1*, a hot start was used, with initial heating at 94°C for 5 min, followed by the addition of the polymerase and then 35 cycles of denaturing (at 94°C for 1 min), annealing (at 63°C for 1 min), and chain extension (at 72°C for 1.5 min), followed by a final extension step at 72°C for 10 min. The *Adh2* genomic clones were generated according to the same PCR reactions but with the annealing temperature lowered at 61°C.

Cloning and Sequencing of the Adh *Genes*

The resulting PCR products were cloned into the plasmid vector pGEM (Promega). Restriction and DNA modification enzymes were provided by MINOTECH and New England Biolabs. Agarose gel electrophoresis and other recombinant DNA methods were performed essentially as described by Sambrook et al. (1989). For each genomic region both strands were completely sequenced and the consensus sequence was obtained for two different clones, generated from two independent PCR reactions. Sequencing of the double-stranded plasmids was carried out according to the dideoxy-chain termination method, using either vector-specific (T7, SP6) or custom gene-specific (internal) primers. A Li-Cor 4200L sequencer at the Laboratory of Microchemistry (IMBB-FORTH, Crete, Greece) was used as well.

The orthology of the *Adh1* and *Adh2* genes across species is based on the length of intron 1 and the number of amino acids of the predicted polypeptides, as also reported by Goulielmos et al. (2003). Thus, intron 1 of *Adh1* was longer than intron 1 of *Adh2* by about 900 bp and ADH1 was smaller than ADH2 by one amino acid.

DNA Sequence Analysis

The DNA sequences were analyzed with the GCG Sequence Analysis Software Computer Package. Alignment of the sequences was done using the Clustal X program (Thompson et al. 1997). The rates of synonymous (K_s) and nonsynonymous (K_a) substitutions were estimated using the DnaSP computer program (Rozas and Rozas 1999). The nucleotide sequences used in this study have the following accession numbers: *B. oleae Adh1*, AJ277835; *B. oleae Adh2*, AJ277834; *B. dorsalis Adh1*, AJ488554; *B. dorsalis Adh2*, AJ539542; *B. cucurbitae Adh1*, AJ539546; *B. cucurbitae Adh2*, AJ539541; *B. tryoni Adh2*, AJ539543; *B. scutellatus Adh2*, AJ539544; *C. capitata Adh1*, Z30194; *C. capitata Adh2*, Z30195; *C. cosyra Adh1*, AJ539546; *C. cosyra Adh2*, AJ539540; *C. rosa Adh2*, AJ539539; *C. fasciventris Adh2*, AJ539538; and *Sarcophaga peregrina Adh*, D63669.

Phylogenetic Tree Construction

Phylogenetic trees were constructed using the neighbor-joining (Saitou and Nei 1987), UPGMA (Sneath and Sokal 1973), minimum evolution (Rzhetsky and Nei 1992), and maximum parsimony (Fitch 1971) methods, through the MEGA-2 computer package (Kumar et al. 2001). Kimura's (1980) two-parameter distance was used. To assess the confidence of individual nodes a bootstrap analysis (Felsenstein 1985) with 1000 replications was performed using the same computer package. The transition/transversion ratio was calculated using the DnaSP 3.50 program (Rozas and Rozas 1999).

Results

The Adh1 *Locus*

The use of degenerate primers based on the cDNA sequences of the *Adh1* of *C. capitata* and *B. oleae* produced an ~2-kb fragment when 200 ng of genomic DNA from the species *C. cosyra*, *B. cucurbitae*, and *B. dorsalis* was used as template. The size of the resulting products was reminiscent of the one observed in the respective genes of *C. capitata* and *B. oleae*. Nested PCR that was performed using a series of degenerate primers based on internal sequences of the *Adh1* of these species produced the initial positive, strong evidence for the existence of an *Adh1*-like sequence. Therefore, we proceeded in the cloning and sequencing of the PCR products. To define the intron/exon splice junctions, the generally approved **(C/A)AG**GTAAGTA and YYYNYAG**G** consensus sequences were considered, thus defining the 5′ and 3′ boundaries of each intron, respectively (Breathnach et al. 1978).

The *Adh1* genomic regions of all species under examination consist of three exons interrupted by two introns, and their coding sequences, with 257 codons, exhibit a considerably high nucleotide sequence identity with the respective coding sequences of *C. capitata* and *B. oleae*. The positions of introns are identical to those found in the *Adh1* gene of *B. oleae* (Goulielmos et al. 2001), located immediately after amino acid residues 31 and 165 of the predicted protein. The organization, total length, and size of the introns of all *Adh1* genes examined are reported in Table 1 and Fig. 1.

Table 1. Size (base pairs) of the *Adh1* and *Adh2* genes determined in the tephritid species and their respective introns 1 and 2

Species	Gene	Total length	Intron 1	Intron 2
B. oleae	*Adh1*	1981	1071	136
B. oleae	*Adh2*	988	125	82
B. cucurbitae	*Adh1*	1982	1072	136
B. cucurbitae	*Adh2*	976	131	68
B. dorsalis	*Adh1*	1982	817	118
B. dorsalis	*Adh2*	967	110	80
B. tryoni	*Adh2*	974	126	68
B. scutellatus	*Adh2*	968	123	68
C. cosyra	*Adh1*	1981	1071	136
C. cosyra	*Adh2*	1011	164	70
C. rosa	*Adh2*	968	121	70
C. fasciventris	*Adh2*	1011	164	70

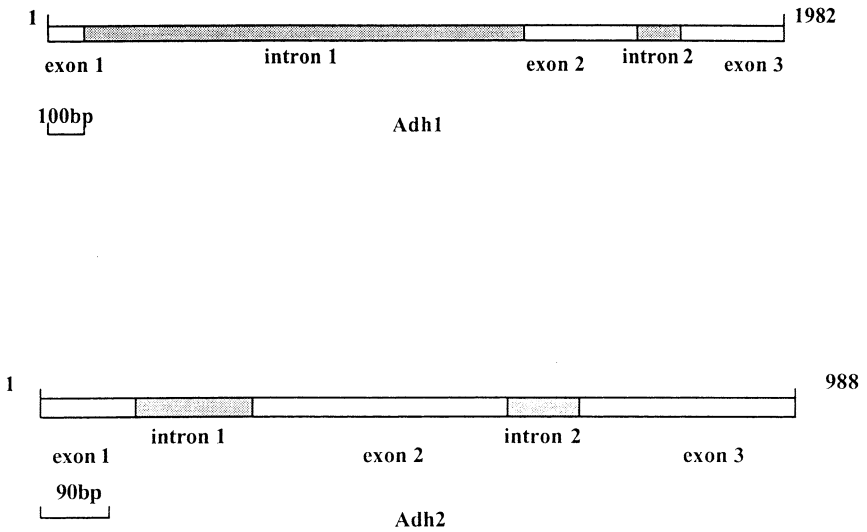

Fig. 1. Schematic representation of the *Adh1* and *Adh2* genomic regions. *White bars* represent exons; *gray bars*, introns.

The Adh2 *Locus*

We cloned and sequenced the genomic region of the *Adh2* locus of four *Bactrocera* species: *B. tryoni*, *B. scutellatus*, *B. cucurbitae*, and *B. dorsalis*. In addition, we cloned this locus from three *Ceratitis* species: *C. cosyra*, *C. fasciventris*, and *C. rosa*. In all species examined, this region consists of three exons interrupted by two introns and the coding sequence corresponds to a polypeptide of 258 amino acids. The position and size of the introns obtained by sequencing the genomic DNAs from the initial ATG codon to the stop codon, located immediately after amino acid residues 31 and 165 of the predicted protein, are reported in Table 1 and Fig. 1.

The Adh *Coding Sequence in the Family Tephritidae*

In all *Bactrocera* and *Ceratitis* species examined in this study, the amino acid coding region of *Adh* is encoded by three exons interrupted by two introns as reported above, similarly as happens in the species *B. oleae* that were analyzed previously (Goulielmos et al. 2001). The inferred amino acid sequences of the two ADH peptides of *B. cucurbitae* were found to be 77.4% identical to each other; the ADH1 and ADH2 enzymes of *B. oleae*, 79% identical; those of *B. dorsalis*, 80.5%; those of *C. cosyra*, 81.3%; and those of *C. capitata*, 81.7%. There is no evidence so far that the *Adh1* genes of *B. dorsalis*, *B. cucurbitae*, *B. oleae*, and *C. cosyra* produce active protein products. The difficulty in detecting the second ADH isozyme of the species under study is compatible with the fact that the ADH1 of *C. capitata* is expressed only in the insect's muscle, thus becoming difficult to detect (Benos et al. 2000).

The alignment of 14 sequences showed 491 (of 749) variable and 228 parsimony-informative nucleotide sites. Differences in the transition/transversion ratio and G + C composition are a potential problem when inferring phylogenies from sequence data; nevertheless, such a problem does not appear in *Adh* sequences under study (see next paragraph).

Codon Usage Bias

When all codon sites are taken into account, the G + C content of the *Adh1* sequences is virtually identical, having a value of 47%, and the same is observed for the G + C content of all *Adh2* sequences, having a value of 49%. At third codon positions, however, there was a light excess of G + C content, only in *Adh2*, ranging between 51% and 54%. The transition/transversion ratio (r = si/sv) has an average value of 1.0. In particular, according to data referred to the first, second, and third codon positions, the si/sv is 1.2, 0.6, and 1.0, respectively. Variation in *Adh* G + C content among species of the genus *Drosophila* has been noted earlier (Shields et al. 1988; Starmer and Sullivan 1989), and it reflects a bias in codon preferences that is generally encountered in well-expressed genes in other taxa, ranging from bacteria to humans (Sharp et al. 1988).

The values of codon bias for *Adh* genes were estimated by calculating the deviation from random synonymous codon usage (χ^2) (Shields et al. 1988), the codon bias index (CBI), and the effective number of codons (ENC) by summing the "effective number of alleles" used by each of the 20 amino acids (data not shown). An extremely biased gene uses only 20 codons, whereas an unbiased gene tends to use 61 codons equally (Wright 1990). The *Adh* genes of all the tephritid species examined have a very low codon bias. Two exceptions that should be mentioned deal with CCA and ACC codons (both of four possible ones), which account for 54% and 64% of all prolines and threonines, respectively.

Amino Acid Alignment of DNA Sequences and Phylogenetic Relationships in the Family Tephritidae

Amino acid sequences of the *Ceratitis* and *Bactrocera* *Adh* genes were aligned and the ADH1 sequences of all tephritid species were shorter compared to the ADH2 sequences of the same species by one amino acid at the end of the sequence. Fourteen and four-tenths percent of the ADH1 and 20.9% of the ADH2 amino acid sites are variable. Comparison of the tephritid sequences examined reveals conservation of all the protein's functionally important amino acids. Nine amino acid residues that were determined to be particularly important for enzymatic activity in *Drosophila* (Scrutton et al. 1990; Cols et al. 1993; Jornvall et al. 1995; Benach et al. 1999) were also found to occur in the same positions in all tephritid ADH products (data not shown). Further similarities in amino acid sequence among *Adh* genes of the tephritid species or among tephritids and drosophilids can be attributed either to selective constraints or to phylogenetic relationships. In a comparison of ADH1 and ADH2 sequences, 163 strictly conserved sites were detected. This conservation confers apparently to the maintenance of the protein's functionality. Among the conserved amino acids there are 18 glycines and 11 valines. Glycine is a small nonpolar amino acid that can fit in inner parts of the folded protein structure (Jornvall et al. 1984). The acid/base content (aspartic and glutamic acids/arginine and lysine) ranges between 25/25 (*B. oleae*) and 27/19 (*B. cucurbitae*) in ADH1 proteins and from 20/23 (i.e., *C. capitata*) to 20/22 (i.e., *B. tryoni*) in ADH2 proteins. These rates do not deviate highly from those calculated in ADH proteins of *D. lebanonensis* (23/22) and *D. melanogaster* (22/21).

In *Drosophila*, an ancient duplication of the alcohol dehydrogenase gene gave rise to the fat body protein-2 gene (*Fbp2*), which encodes a protein that differs substantially from ADH in its very high methionine content (20% of all amino acids in FBP2 (Rat et al. 1991; Meghlaoui and Veuille 1997). Such a high methionine content is very unusual in proteins. We attempted to answer the question of whether a similar observation might be made for the amino acid changes in the products of the *Adh* gene duplication in the species of Tephritidae studied, but the number of methionine residues in ADH1 as well as ADH2 proteins does not differ significantly, remaining at the expected very low percentage.

The *Adh* coding region in the tephritid species examined is generally conserved but, nevertheless, remains phylogenetically informative. In the present paper we are interested primarily in phylogenetic trees based on the exonic regions of *Adh* sequences without considering the respective genomic sequences. The most important reason for deciding this is because *Adh* genes were found to have introns of different sizes in each species, and in addition, the introns of *C. capitata Adh1* and *Adh2* are still unknown. Phylogenetic analyses were carried out on 14 *Adh* sequences from nine species. Given the absence of any A + T composition bias, we used Kimura two-parameter distances for constructing phylogenetic trees. However, Jukes–Cantor's distances yielded almost-identical trees (data not shown). Neighbor-joining and maximum parsimony trees were generated for (1) nucleotide sequences of the *Adh* coding region and (2) amino acid sequences. The sequence of *Sarcophaga peregrina*, a member of the family Calyptratae whose ADH protein is 36% identical to *Drosophila* protein (Horio et al. 1996), was chosen to be used as an outgroup, as reported also by Goulielmos et al. (2001) and Brogna et al. (2001).

The availability of cDNA sequences for the two *Adh* genes of *C. cosyra*, *C. capitata*, *B. oleae*, *B. cucurbitae*, and *B. dorsalis*, and of the *Adh2* gene of *C. rosa*, *C. fasciventris*, *B. scutellatus*, and *B. tryoni* can be used in an attempt to answer the question of

whether there has been one duplication event that preceded the separation of the two genera *Ceratitis* and *Bactrocera*, thus generating *Adh1* and *Adh2*, or whether there have been separate and independent duplication events in each of the species where two *Adh* genes have been found. In the first case, *Adh1* and *Adh2* genes would be considered paralogous to each other, with *Adh1* or *Adh2* of one species being orthologous to the corresponding gene of the other species. In contrast, if the second alternative is the case, *Adh1* and *Adh2* will be pairs of paralogous genes within each species.

Taking into consideration all *Adh1* and *Adh2* cDNA sequences available, we constructed the phylogenetic trees according to neighbor-joining (NJ) and maximum parsimony (MP) methods (Fig. 2) as well as UPGMA and minimum evolution (ME) (data not shown). All these trees cluster the *Adh* sequences according to "type" (i.e., *Adh1* sequences form a main cluster and *Adh2* form a second one) rather than "species." Moreover, the *Adh1* and *Adh2* sequences of *Bactrocera* and *Ceratitis* species form distinct subclusters within each of the two main clusters, with congeneric sequences thus clustering together. This type of grouping is supported by a high bootstrap value.

The trees produced from the amino acid sequences are similar to the ones obtained from cDNA sequences. Thus, all NJ and MP trees (Fig. 3) as well as the UPGMA and ME trees (data not shown) consist of two major clusters, with the ADH isozymes clustering again according to "type" (i.e., ADH1 isozymes cluster together and ADH2 isozymes form a second cluster). This grouping is supported again by high bootstrap values. Furthermore, the congeneric ADH1 and ADH2 sequences form separate subclusters within each main cluster. All these findings strongly support the idea that the *Adh* duplication in the family Tephritidae occurred before the split of this family in different genera.

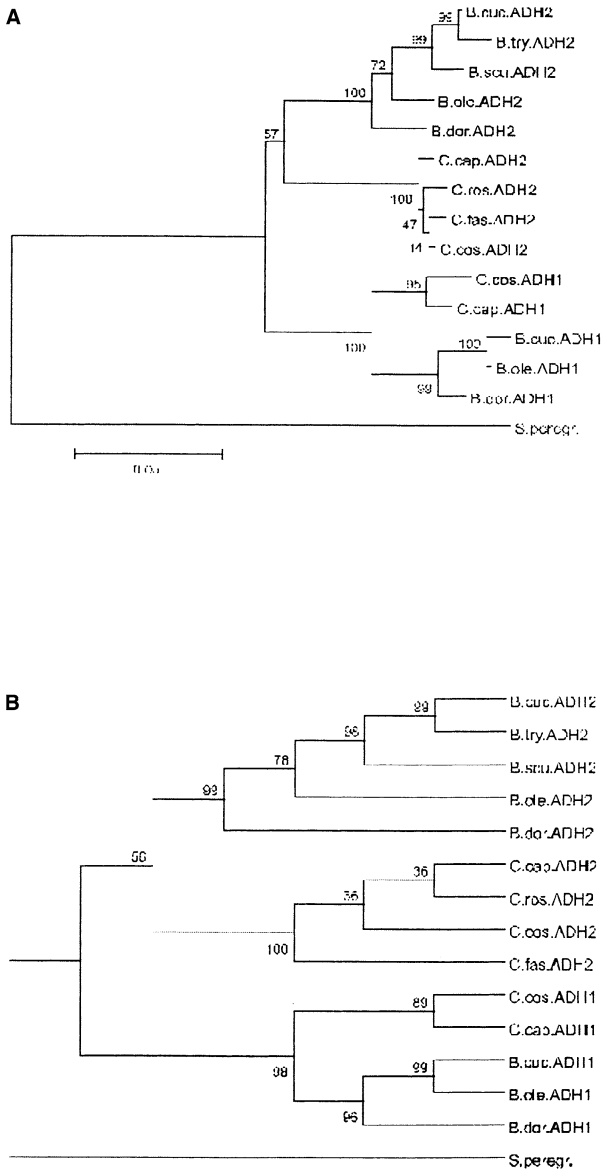

Fig. 2. Nucleotide phylogenetic trees of 14 tephritid *Adh* genes. The *Adh* gene of *Sarcophaga peregrina* was used as outgroup. Numbers are bootstrap confidence values. The bar below each tree indicates the distance measure. **A** Neighbor joining tree; **B** maximum parsimony tree.

The Age of Adh *Gene Duplication*

The *Adh* genes of tephritids have, as reported previously, very low codon usage bias. Low codon bias was also observed in the *Adh* genes of the Hawaiian *Drosophila* (Rowan and Hunt 1991). Hence, it seems justified to use the evolutionary rate of *Adh* calculated from studies carried out in Hawaiian *Drosophila* species to estimate the divergence between any conspecific *Adh1* and *Adh2* genes in tephritids. Using the time of emergence of the Hawaiian Islands, Rowan and Hunt (1991) suggested a rate of 1.5×10^{-8} synonymous substitution per year (K_s), or 0.5×10^{-8} for the total coding region (both synonymous and nonsynonymous sites), for *Adh* of Hawaiian drosophilids. Russo et al. (1995) concentrated exclusively on the total coding region substitution rate, which they revised to 1×10^{-8}.

Goulielmos et al. (2003) recently presented evidence that the *Drosophila* rate can be used as a first approximation to estimate rates of divergence in tephritids, since their estimate of divergence time was in some cases almost identical to that given by Beverley and Wilson (1984) based on immunological reaction of larval serum proteins. Therefore, the *Drosophila* evolutionary rate was used to estimate the divergence time of various *Bactrocera* species as well as the age of the ADH2 enzymatic polymorphism appearing in *B. oleae*. In the current study, the Kimura two-parameter K_a and K_s values from the

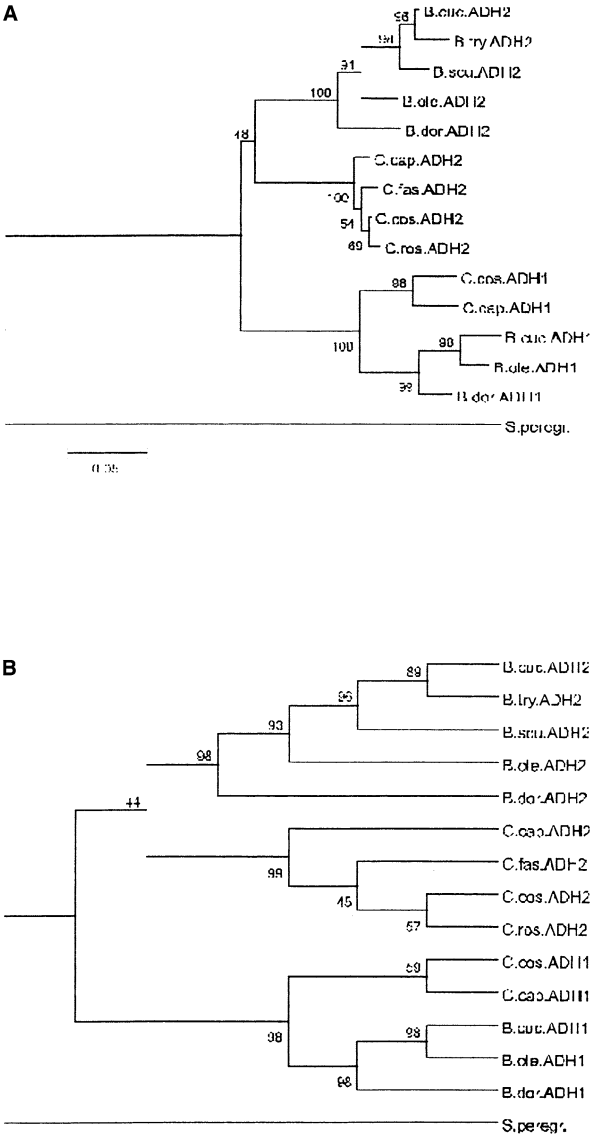

Fig. 3. Amino acid phylogenetic trees of 14 tephritid *Adh* genes. *The Adh* gene of *Sarcophaga peregrina* was used as outgroup. Numbers are bootstrap confidence values. The bar below each tree indicates the distance measure. **A** Neighbor joining tree; **B** maximum parsimony tree.

cDNA sequences of the *Adh* genes were calculated and estimates were obtained by both methods. The average divergence time for the two congeneric *Adh* genes of the tephritid species is estimated at 38.48 Myr ago. According to the existing information from the literature, the split of *Bactrocera* from *Ceratitis* is placed at between 31.5 Myr (Beverley and Wilson 1984) and 32.05 Myr (Goulielmos et al. 2003) ago. Thus, we suggest that our data support the hypothesis that the duplication of a single ancestral *Adh* gene in the Tephritidae family occurred before the emergence of the genera *Ceratitis* and *Bactrocera*. As a concequence, the *Adh* duplication was based on a prespeciation rather than a postspeciation event, a hypothesis that remains to be strengthened upon accumulation of data from more tephritid genera.

Discussion

The alcohol dehydrogenase gene is one of the central topics of interest in evolutionary and biochemical genetics, as it is a promising tool for shedding light on the biochemical and molecular factors that govern the dynamics of evolutionary processes. Thus, the study of *Drosophila Adh* genes occupies a prominent position in evolutionary studies. Furthermore, these genes have been extensively studied as tools for reconstructing the phylogeny and estimating the divergence times of drosophilid species. Nevertheless, the issue of *Adh* duplication in the genus *Drosophila* remains unsolved. In the subgenera *Sophophora* and *Scaptodrosophila* there exist a functional gene and an apparently nonfunctional gene, tightly linked to the first one. The situation appears more complicated in the subgenus *Drosophila*, where species of the *repleta* group have three genes, one of which is nonfunctional pseudogene (Sullivan et al. 1994). Russo et al. (1995) have proposed a scenario according to which as many as four independent duplications occurred in the last 180 Myr, thus yielding the present organization of the *Adh* region in drosophilids. In addition, these authors suggested that a single duplication event that happened 6–11 Myr ago could explain the origin of the *Adh* genes in the *D. mulleri* subgroup, except in *D. hydei*, where this event cannot be older than 4 Myr. Therefore, they concluded either that there have been multiple duplication events within the *repleta* group or that gene conversion may have increased the similarity of conspecific genes in some species. Thus, the species in the *repleta* group for which sequence information on *Adh1* and *Adh2* is available (i.e., *D. mulleri*, *D. mojavensis*, *D. hydei*) can be used to infer that the likely sequence of events that led to the formation of these genes in *D. mulleri* and *D. mojavensis* preceded the speciation of *D. mulleri* and *D. mojavensis*. This notion is supported by the overall analysis performed and presented in more detail by Sullivan et al. (1989). Surprisingly, Begun (1997) presented evidence that *Adh-ψ*, which was hypothesized to be a pseudogene derived from an *Adh* duplication in species of the r*epleta* group, was actually a new gene of unknown function that recruited a large number of new N-terminal amino acids, thus becoming more basic. Similarly, Brogna and Ashburner (1997) suggested that the *Adh*-related (*Adhr*) gene in the *Sophophora* subgenus is an active gene, with the function of the (paralogous) protein remaining still unknown. It should be mentioned here that all *Adhr* genes detected are highly conserved in *Drosophila*, even more conserved than *Adh* (Benos

et al. 2000). However, we do not focus on the *Adhr* gene in the present study because it has not yet been described in any of the tephritid species.

To deepen the existing knowledge dealing with the evolutionary pattern of tephritid alcohol dehydrogenase, several studies and analyses based on the overall accumulated data have been performed. It is worthwhile noting that a remarkable variation in the number of *Adh* loci was observed formerly in several tephritid species from various genera of this family. In this framework, we used the available information from the *Adh* genes of *Ceratitis* and *Bactrocera* species examined in our attempt to suggest a firm evolutionary scenario dealing with the *Adh* duplication. According to the phylogenetic trees constructed based on the ADH proteins, we conclude that *Adh1* and *Adh2* genes form two independent phylogenetic clusters. This type of clustering was produced by all methods of tree construction used and is strongly supported by bootstrap values. In addition, the fact that the ADH2 proteins are longer by one amino acid also supports the grouping observed. Furthermore, comparison of all full-length cDNAs available produced an identical picture, with the *Adh* genes clustering according to "type" and, thus, the conspecific genes being separated into two main clusters. This topography is the same irrespective of the tree construction method used, based on either the calculation of genetic distances (neighbor joining) or character-state differences (maximum parsimony). All these trees support a phylogenetic history of the *Adh* genes that is based on a unique duplication event of an ancestral gene, within the family Tephritidae, before further speciation and splitting of the genera and species. Accordingly, Benos et al. (2000) suggested that the duplication of the *Adh* gene in Tephritidae was more likely to have occurred prior to the divergence of the genera *Bactrocera* and *Ceratitis*. A similar answer has been given dealing with the *Adh* duplication in the grasses (Poaceae). In particular, phylogenetic analyses suggested that *Adh* duplicated into *Adh1* and *Adh2* before the radiation of the grasses (Gaut et al. 1999). Further studies are required to characterize fully the *Adh* gene system and, especially, to reveal whether or not *B. scutellatus*, *B. tryoni*, *C. fasciventris*, and *C. rosa* have a second copy of the *Adh* gene. The possibility that the second gene has been lost during the evolution of these species should be considered as well. Thus, the absence of a second *Adh* gene in several (closely related) species of Tephritidae should not surprise anybody. In *Drosophila*, apart from the species in the *D. virilis* group, the other two species showing *Adh* duplications, *D. montana* and *D. lacicola*, are not the ones which are the most closely related to *D. virilis*. This observation indicates that duplicate *Adh* genes, at least in *Drosophila*, have also been found in two species that apparently originated independent of *D. virilis* (Nurminsky et al. 1996).

The hypothesis suggested in this study, that the presence of two *Adh* genes in tephritid species reflects the result of an early duplication event that predates the emergence of the various genera, is also favored by the observation that many species of the family Tephritidae have two (or, more rarely, three) isoenzymes for ADH. The latter again supports the view that the *Adh* duplication is a rather early event in the radiation of the family Tephritidae. Thus, a single *Adh* locus was found initially in *Acinia fucata*, *Rachiptera limbata*, *Rhagoletis nova*, *R. conversa*, and *R. striatella*. However, two *Adh* loci were detected in *Ceratitis capitata* (Malacrida et al. 1992), in *B. oleae* (Goulielmos et al. 2001), in several *Rhagoletis* species (Berlocher and Bush 1982), and in *Anastrepha fraterculus*, *A. obliqua*, *A. bistrigata*, *A. striata*, *A. serpentina*, and *A. grandis* (Matioli et al. 1986, 1992). Surprisingly, tephritid species showing two *Adh* loci infest plant tissues during their larval stages and live inside ripening fruit (i.e., those of the genera *Anastrepha*, *Ceratitis*, and *Rhagoletis*); in contrast, other species that are members mainly of the genera *Acinia*, *Rachiptera*, and *Tomoplagia* and carry a single *Adh* gene live inside inflorescences or galls (Matioli et al. 1992). This observation may correlate ADH evolution with speciation through adaptation to various feeding niches, an issue approached by Atrian et al. (1998). Unfortunately, for all the aforementioned species, except several *Ceratitis* and *Bactrocera* species, the *Adh* genes have not been cloned, thus precluding their contributing essentially to deciphering the evolutionary process followed by the ancestral *Adh* gene.

In some cases of tephritids it was suggested previously that secondary bands resulted from the binding of protein and NAD-carbonyl derivatives. Indeed, the main effect of the addition of commercial preparations of NAD was an increase in the negative charge of the enzyme, as inferred by the changes in its mobility (Matioli et al. 1992). This is why it was supposed in the past that the presence of two *Adh* loci in several genera might not be the result of independent events of gene duplication. Apart from the high level of amino acid homology observed, ADH1 and ADH2 proteins exhibit similar enzymatic specificities and corresponding electrophoretic mobilities. The extreme hypothesis that the second *Adh* locus may be the result of an evolutionary trend toward a broad substrate specificity for the octanol dehydrogenase locus, which was not found initially in tephritids, seems unacceptable. In *Anastrepha bistrigata*, two *Adh* loci in addition to *Odh* are expressed, thus the aforementioned hypothesis that the second *Adh* locus is the result of evolutionary modifications of the *Odh* locus becomes inconceivable, at least in the case of this species (Matioli et al. 1992). In addition, the

existence of an octanol dehydrogenase locus has been reported in *B. tryoni* and *B. neohumeralis* (McKechnie 1975). The allozyme polymorphism observed exclusively in the *Adh2* locus of *B. oleae* is compatible with the reduction of selective constraints that follows a duplication event (Goodman et al. 1975). However, in the case of the α-glycerol-phosphate dehydrogenase gene duplication, an allozymic polymorphism was detected in both the standard and the duplicated gene copy (Takano et al. l989).

The deduced amino acid sequences confirmed that ADH1 and ADH2 isozymes are highly related to each other and to any known ADHs from other insects. Therefore, all ADH proteins of tephritids examined are classified as members of the short-chain dehydrogenase/reductase (SDR) family. Most members of this family, apart from their average length of 250 amino acid residues, are characterized by distant duplications and divergence, are functionally and structurally related, and lack metal ions in their active site (Persson et al. 1991; Jornvall et al. 1995). Synonymous codons are not always used with equal frequency, and in several cases various genes from the same species share similarities in codon usage preference (Grantham et al. 1981). The *Adh* genes of the tephritid species examined were found to have a very low codon bias. In accordance with our findings, Rina and Savakis (1991) noted previously that the codon usage of medfly genes is, in general, less biased than in *Drosophila*. However, some notable exceptions are the cases of two codons of arginine (of six possible codons) and two of glycine (of four possible codons), which account for almost 90% of all arginines and glycines, respectively. In contrast, highly biased codon usage in some *Drosophila* genes was reported previously (Sharp and Li 1986), and, especially, a high bias toward C- and G-ending codons was reported by Nurminsky et al. (1996) for *Adh* in *D. melanogaster* (79%). Two alternative explanations may be given, assuming the possibility either that selective pressure for specific synonymous codons is not as strong in the *Adh* genes of Tephritidae or that selection of synonymous codons may not be operating at all in some species of Tephritidae if their effective population size is small (Shields et al. 1988; Rina and Savakis 1991).

Recent progress in molecular biology has clarified not only the genetic organization of several repeated gene families but also the commonness of gene duplication for ordinary gene loci. In particular, a variety of studies on gene duplication have been performed involving the urate oxidase gene of *Drosophila virilis* (Lootens et al. 1993) and many genes of *Drosophila melanogaster*, including the α-amylase (Gemmill et al. 1985), rosy (Gelbart and Chovnick 1979), glycerol-3-phosphate dehydrogenase (Takano et al. 1989), α-esterase (Robin et al. 1996), and metallothionein (*Mtn and Mto*) (Maroni et al. 1987; Lange et al. 1990) genes. It is a widely accepted hypothesis that gene duplication is a major force in molecular evolution. Since the original function can be maintained by one of the two copies, the duplicate may easily escape the act of natural selection, and therefore, weakened selection pressure allows for the appearance of similar or new gene function (Ohno 1970). Moreover, gene duplications are often followed by an accelerated rate of evolution as observed in phylogenetic analyses of 13 lepidopteran opsin sequences that revealed two recent opsin gene duplication events within the papilionid butterfly family (Briscoe 2001). The *Papilio* Rh3 gene showed an accelerated rate of evolution and, in addition, has in fact evolved a novel function relative to its ancestral gene, *Papilio* Rh2. However, an analogous situation has not been detected in *Adh* genes of tephritids so far. For some gene families, DNA sequence analysis showed that ancestral duplications and mutations of different kinds managed to change the sequences considerably, as reported for the γ-globin genes in mammals (Slightom et al. 1980) and the human haptoglobin genes (Maeda et al. 1984), a situation that has not been observed in the case of *Adh1* and *Adh2* genes in Tephritidae species examined in the present study (see Results). In this context, it seems interesting to examine further the potential advantages that accompany the *Adh* gene duplication as well as the issue of whether the resulting duplication products are active or inactive and to include in these studies species from more tephritid genera.

Acknowledgments. We thank Dr. Masahiko Muraji (National Institute of Agrobiological Sciences, Tsukuba, Ibaraki, Japan), Dr. Merryl Robson (Fruit Fly Research Center, University of Sydney), and Dr. Susan McCobs (Department of Entomology, University of Hawaii) for providing us with *B. scutellatus*, *B. tryoni* and *B. cucurbitae*, and *B. dorsalis*, respectively. We also thank Gerald Franz of the Entomology Unit of the IAEA (Seibersdorf, Austria) for providing us with *C. rosa* and *C. fasciventris*.

References

Amador A, Papaceit M, Juan E (2001) Evolutionary change in the structure of the regulatory region that drives tissue and temporally regulated expression of alcohol dehydrogenase gene in *Drosophila funebris*. Insect Mol Biol 10(3):237–247

Atkinson PW, Mills LE, Starmer WT, Sullivan DT (1988) Structure and evolution of the *Adh* genes of *Drosophila mojavensis*. Genetics 120:713–723

Atrian S, Gonzalez-Duarte R (1982) Comparison of some biochemical features of the enzyme alcohol dehydrogenase in sixteen species of *Drosophila*. In: Lakovaara S (ed) Advances in genetics, development and evolution of *Drosophila*. Plenum Press, New York, pp 251–261

Begun DJ (1997) Origin and evolution of a new gene descended from *alcohol dehydrogenase* in *Drosophila*. Genetics 145:375–382

Benach J, Atrian S, Gonzalez-Duarte R, Ladenstein R (1998) The refined crystal structure of *Drosophila lebanonensis* alcohol dehydrogenase at 1.9 A resolution. J Mol Biol 282:383–399

Benach J, Atrian S, Gonzalez-Duarte R, Ladenstein R (1999) The catalytic reaction and inhibition mechanism of *Drosophila* alcohol dehydrogenase: Observation of an enzyme-bound NAD-ketone adduct at 1.4 A resolution by X-ray crystallography. J Mol Biol 289:335–355

Benos P, Tavernarakis N, Brogna S, Thireos G, Savakis C (2000) Acquisition of a potential marker for insect transformation: isolation of a novel alcohol dehydrogenase gene from *Bactrocera oleae* by functional complementation in yeast. Mol Gen Genet 263:90–95

Benyajati C, Spoerel N, Haymerle H, Ashburner M (1983) The messenger RNA for *Adh* in *D. melanogaster* differs in its 5′ end in different developmental stages. Cell 33:125–133

Berlocher SH, Bush GL (1984) An electrophoretic analysis of Rhagoletis (Diptera: Tephritidae) phylogeny. Syst Zool 31:136–155

Beverley SM, Wilson AC (1984) Molecular evolution in *Drosophila* and the higher Diptera. II. A time scale for fly evolution. J Mol Evol 21:1–13

Breathnach R, Benoist C, Gannon F, Chambon P (1978) Ovalbumin gene: Evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. Proc Natl Acad Sci USA 75:4853–4857

Briscoe AD (2001) Functional diversification of Lepidopteran opsins following gene duplication. Mol Biol Evol 18:2270–2279

Brogna S, Ashburner M (1997). The *Adh*-related gene of *Drosophila melanogaster* is expressed as a functional dicistronic messenger RNA: Multigenic transcription in higher organisms. EMBO J 16:2023–2031

Cols N, Marfany G, Atrian S, Gonzalez-Duarte R (1993) Effect of site-directed mutagenesis of conserved positions of *Drosophila* alcohol dehydrogenase. FEBS Lett 319:90–94

Cosmidis N (1995) Study of the selection at the alcohol dehydrogenase locus in the olive fruit fly, PhD dissertation. Agricultural University of Athens, Athens (in Greek)

Feder JL, Chilcote CA, Bush GL (1988) Genetic differentiation between sympatric host races of the apple maggot fly *Rhagoletis pomonella*. Nature 336:61–64

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

Fischer JA, Maniatis T (1985) Structure and transcription of the *Drosophila mulleri* alcohol dehydrogenase genes. Nucleic Acids Res 13:6899–6917

Fitch WM (1971) Towards defining the course of evolution: Minimum change for a specific tree topology. Syst Zool 20:406–416

Foote RH, Blanc FL, Norrbom AL (1993) Handbook of the fruit flies (Diptera: Tephritidae) of America north of Mexico. Comstock, Ithaca, NY

Gasperi G, Baruffi L, Malacrida A, Robinson AS (1992) A biochemical genetic study of alcohol dehydrogenase isozymes of the medfly *Ceratitis capitata*. Wied Biochem Genet 30:289–304

Gasperi G, Kafetzopoulos D, Christodoulidou A, Bouriotis V, Savakis C (1994) Isolation and partial characterization of two alcohol dehydrogenase isozymes from the medfly *Ceratitis capitata*. Insect Biochem Mol Biol 24(1):87–94

Gaut BS, Peek AS, Morton BR, Clegg MT (1999) Patterns of genetic diversification within the *Adh* gene family in the grasses (Poaceae). Mol Biol Evol 16(8):1086–1097

Chovnick A (1979) Spontaneous unequal exchange in the *rosy* region of *Drosophila melanogaster*. Genetics 92:849–859

Gemmill RM, Levy JN, Doane WW (1985) Molecular cloning of α-amylase genes from *Drosophila melanogaster*. I. Clone isolation by use of a mouse probe. Genetics 110:299–312

Goodman M, Moore GW, Matsuda G (1975) Darwinian evolution in the genealogy of hemoglobin. Nature 253:603–608

Goulielmos GN, Cosmidis N, Loukas M, Tsakas S, Zouros E (2001) Characterization of two alcohol dehydrogenase (*Adh*) loci from the olive fruit fly, *Bactrocera* (*Dacus*) *oleae* and implications for *Adh* duplication in dipteran insects. J Mol Evol 52(1):29–39

Goulielmos GN, Cosmidis N, Theodorakopoulou ME, Loukas M, Zouros E (2003) Tracing the history of an enzyme polymorphism: The case of alcohol dehydrogenase-2 (*Adh-2*) of the olive fruit fly *Bactrocera oleae*. Mol Biol Evol 20(3):293–306

Grantham R, Gautier C, Gouy M, Jacobzone M, Mercier R (1981) Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res 9:r43–r74

Holmes DS, Bonner J (1973) Preparation, molecular weight, base composition and secondary structure of giant nuclear ribonucleic acid. Biochemistry 12:2330–2338

Horio T, Kubo T, Natori S (1996) Purification and cDNA cloning of the alcohol dehydrogenase of the flesh fly *Sarcophaga peregrina*. A structural relationship between alcohol dehydrogenase and a 25-kDa protein. Eur J Biochem 237(3):698–703

Jornvall HH, von Bahr-Lindstrom H, Jany KD, Ulmer W, Froschle M (1984) Extended superfamily of short alcohol-polyol-sugar dehydrogenases: Structural similarities between glucose and ribitol dehydrogenases. FEBS Lett 165:190–196

Jornvall H, Persson B, Krook M, Atrian S, Gonzalez-Duarte R (1995) Short-chain dehydrogenases/reductases (SDR). Biochemistry 34:6003–6013

Juan E, Papaceit M, Quintana A (1994) Nucleotide sequence of the genomic region encompassing *Adh* and *Adh-dup* genes of *D. lebanonensis* (Scaptodrosophila): Gene expression and evolutionary relationships. J Mol Evol 38:455–467

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA 2: Molecular Evolutionary Genetics Analysis software. Bioinformatics 17:1244–1245

Kwiatowski J, Skarecky D, Hernandez S, Pham D, Quijas F, Ayala F (1991) High fidelity of polymerase chain reaction. Mol Biol Evol 8(6):884–887

Lange BW, Langley CH, Stephan W (1990) Molecular evolution of *Drosophila* metallothionein genes. Genetics 126:921–932

Lootens S, Burnett J, Friedman TB (1993) An interspecific gene duplication polymorphism of the urate oxidase gene of *Drosophila virilis:* A genetic and molecular analysis. Mol Biol Evol 10:635–646

Maeda N, Yang F, Barnett DR, Bowman BH, Smithies O (1984) Duplication within the haptoglobin Hp^2 gene. Nature 309:131–135

Malacrida AR, Gasperi G, Zacharopoulou A, Torti C, Riva-Francos E, Milani R (1992) Evidence for a genetic duplication involving alcohol dehydrogenase genes in *Ceratitis capitata*. Biochem Genet 30:35–48

Maroni G, Wise J, Young JE, Otto E (1987) Metallothionein gene duplications and metal tolerance in natural populations of *Drosophila melanogaster*. Genetics 117:739–744

Matioli SR, Morgante JS, Malavasi A (1986) Genetical and biochemical comparisons of *Anastrepha fraterculus* and *A. obliqua:* Evidence for gene duplication. Biochem Genet 18:98–100

Matioli SR, Morgante JS, Solferini VN, Frias DL (1992) Evolutionary trends of alcohol dehydrogenase isozymes in some species of tephritid flies. Rev Brasil Genet 15:33–50

McKechnie SW (1975) Enzyme polymorphism and species discrimination in fruit flies of the genus *Dacus* (Tephritidae). Austral J Biol Sci 28:405–411

McPheron BA, Smith DC, Berlocher SH (1988) Genetic differences between host races of *Rhagoletis pomonella*. Nature 336:64–66

Meghlaoui GK, Veuille M (1997) Selection and methionine accumulation in the fat body protein 2 gene (FBP2), a duplicate of the *Drosophila* alcohol dehydrogenase. J Mol Evol 44:23–32

Nurminsky DI, Moriyama EN, Lozovskaya ER, Hartl DL (1996) Molecular phylogeny and genome evolution in the *Drosophila virilis* species group: Duplications of the alcohol dehydrogenase gene. Mol Biol Evol 13:132–149

Ohno S (1970) Evolution by gene duplication. Springer, New York

Ohta T (1993) Pattern of nucleotide substitution in growth hormone-prolactine gene family: A paradigm for evolution by gene duplication. Genetics 134:1271–1276

Persson B, Krook M, Jornvall H (1991) Characteristics of short-chain alcohol dehydrogenases and related enzymes. Eur J Biochem 200:537–543

Rat L, Veuille M, Lepesant J-A (1991) *Drosophila* fat body protein P6 and alcohol dehydrogenase are derived from a common ancestral protein. J Mol Evol 33:94–103

Rina M, Savakis C (1991) A cluster of vitellogenin genes in the Mediterranean fruit fly *Ceratitis capitata:* Sequence and structural conservation in dipteran yolk proteins and their genes. Genetics 127:769–780

Robin C, Russell RJ, Medveczky KM, Oakeshott JG (1996) Duplication and divergence of the genes of the α-esterase cluster of *Drosophila melanogaster*. J Mol Evol 43:241–252

Robinson AS, MacLeod EG (1993) Genetic sexing in *Drosophila melanogaster* using the alcohol dehydrogenase locus and a Y-linked translocation. Theor Appl Genet 59:23–24

Rowan RG, Dickinson WJ (1986) Two alternate transcripts coding for alcohol dehydrogenase accumulate with different developmental specificities in different species of the *picture-winged Drosophila*. Genetics 114:435–452

Rowan RG, Dickinson WJ (1988) Nucleotide sequence of the genomic region encoding alcohol dehydrogenase in *Drosophila affinisdisjuncta*. J Mol Evol 28:49–70

Rowan RG, Hunt JA (1991) Rates of DNA change and phylogeny from the DNA sequences of the alcohol dehydrogenase gene for five closely related species of Hawaiian *Drosophila*. Mol Biol Evol 8:49–70

Rozas J, Rozas R (1999) DnaSp version 3: An integrated program for molecular population genetics and molecular evolution analysis. Bioinform Appl Note 15:174–175

Russo C, Takezaki MN, Nei M (1995). Molecular phylogeny and divergence times of *Drosophila* species. Mol Biol Evol 12:391–404

Rzhetsky A, Nei M (1993) Theoretical foundation of the minimum-evolution method of phylogenetic inference. Mol Biol Evol 10:1073–1095

Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Savakis C, Ashburner M, Willis JH (1986) The expression of the gene coding for alcohol dehydrogenase during the development of *Drosophila melanogaster*. Dev Biol 114:194–207

Scott MJ, Kriticou D, Robinson AS (1993) Isolation of cDNAs encoding 6-phosphogluconate and glucose-6-phosphate dehydrogenase from the Mediterranean fruit fly *Ceratitis capitata:* Correlating genetic and physical maps of chromosome 5. Insect Mol Biol 1:213–222

Scrutton NS, Berry A, Perham RN (1990) Redesign of coenzyme specificity of a dehydrogenase by protein engineering. Nature 343:38–43

Sharp PM, Li W-H (1986) On the rate of DNA sequence evolution in *Drosophila*. J Mol Evol 28:389–402

Sharp PM, Cowe E, Higgins DG, Shields DC, Wolfe KH, Wright F (1988) Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*; A review of the considerable within-species diversity. Nucleic Acids Res 16:8207–8211

Shields DC, Sharp PM, Higgin DG, Wright F (1988) "Silent" sites in *Drosophila* genes are not neutral: Evidence of selection among synonymous codons. Mol Biol Evol 5:704–716

Slightom JL, Blechl AE, Smithies O (1980) Human fetal G_γ- and A_γ-globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell 21:627–638

Sneath PHA, Sokal RR (1973) Numerical taxonomy. Freeman, San Francisco

Sofer W, Ursprung H (1968) *Drosophila* alcohol dehydrogenase. J Biol Chem 243:3110–3115

Spofford PA (1969) Heterosis and the evolution of duplications. Am Nat 103:407–432

Starmer WT, Sullivan DT (1989) A shift in the third-codon-position nucleotide frequency in alcohol dehydrogenase genes in the genus *Drosophila*. Mol Biol Evol 6:546–552

Sullivan DT, Atkinson PW, Starmer WT (1989) Molecular evolution of the alcohol dehydrogenase genes in the genus *Drosophila*. Evol Biol 24:107–147

Takano T, Kusakabe S, Koga A, Mukai T (1989) Polymorphism for the number of tandemly multiplicated glycerol-3-phosphate dehydrogenase genes in *Drosophila melanogaster*. Proc Natl Acad Sci USA 86:5000–5004

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

Wright F (1990) The "effective number of codons" used in a gene. Gene 87:23–29

Zouros E, Loukas M, Economopoulos A, Mazomenos B (1982) Selection at the alcohol dehydrogenase locus of the olive fruit fly *Dacus oleae* under artificial rearing. Heredity 48:169–185