J Mol Evol (1994) 38:455–467

Journal
of Molecular
Evolution

© Springer-Verlag New York Inc. 1994

Nucleotide Sequence of the Genomic Region Encompassing Adh and Adh-Dup Genes of *D. lebanonensis* (Scaptodrosophila): Gene Expression and Evolutionary Relationships

Elvira Juan, M. Papaceit, A. Quintana

Departament de Genètica, Universitat de Barcelona, Avda, Diagonal 645, 08071 Barcelona, Spain

Received: 23 September 1992 / Revised: 1 August 1993 / Accepted: 30 September 1993

Abstract. The region of the genome of *D. lebanonensis* that contains the Adh gene and the downstream Adh-dup gene was sequenced. The structure of the two genes is the same as has been described for *D. melanogaster.* Adh has two promoters and Adh-dup has only one putative promoter. The levels of expression of the two genes in this species are dramatically different. Hybridizing the same Northern blots with a specific probe for Adh-dup, we did not find transcripts for this gene in *D. lebanonensis.* The level of Adh distal transcript in adults of *D. lebanonensis* is five times greater than that of *D. melanogaster* adults. The maximum levels of proximal transcript are attained at different larval stages in the two species, being three times higher in *D. melanogaster* late-second-instar larvae than in *D. lebanonensis* first-instar larvae. The level of Adh transcripts allowed us to determine distal and proximal initiation transcription sites, the position of the first intron, the use of two polyadenylation signals, and the heterogeneity of polyadenylation sites. Temporal and spatial expression profiles of the Adh gene of *D. lebanonensis* show qualitative differences compared with *D. melanogaster.* Adh and Adh-dup evolve differently as shown by the synonymous and nonsynonymous substitution rates for the coding region of both genes when compared across two species of the melanogaster group, two of the obscura group of the subgenus *Sophophora* and *D. lebanonensis* of the victoria group of the subgenus *Scaptodrsophila.* Synonymous rates for Adh are approximately half those for Adh-dup, while nonsynonymous rates for Adh are generally higher than those for Adh-dup. Adh shows 76.8% identities at the protein level and 70.2% identities at the nucleotide level while Adh-dup shows 83.7% identities at the protein level and 67.5% identities at the nucleotide level. Codon usage for Adh-dup is shown to be less biased than for Adh, which could explain the higher synonymous rates and the generally lower nonsynonymous substitution rates in Adh-dup compared with Adh. Phylogenetic trees reconstructed by distance matrix and parsimony methods show that *Sophophora* and *Scaptodrosophila* subgenera diverged shortly after the separation from the *Drosophila* subgenus.

Key words: Alcohol dehydrogenase — *Drosophila* — *Drosophila lebanonensis* — Gene expression — Codon usage — Phylogenetic relationships

Introduction

The region of the genome that contains the alcohol dehydrogenase gene (Adh) was sequenced in species of *Drosophila* belonging to different subgenera. *Sophophora* species, in addition to Adh, show a downstream ORF, named Adh-dup, whose gene product is still unknown. The Adh genes of *Sophophora* and Hawaiian *Drosophila* species have two promoters (Benyajati et al. 1983; Schaeffer and Aquadro 1987; Rowan and Dickinson 1988) and expression from the two promoters is spatially and temporally regulated. Species of the *repleta* group of the *Drosophila* subgenus have two closely linked Adh genes and one pseudogene. Each of these

Correspondence to: E. Juan

two Adh genes is expressed at different times of development (Fischer and Maniatis 1985; Atkinson et al. 1988). The spatial and temporal expression of the Adh in species with a single copy gene has been studied in detail in species of melanogaster subgroup (Goldberg et al. 1983; Savakis et al. 1986; Thomson et al. 1991) and in Hawaiian *Drosophila* (Fang et al. 1991). The interspecific comparison of expression patterns shows some common and some differential features. Broadly, Adh is expressed primarily from the proximal promoter in larvae and from the distal promoter in adults. The common spatial pattern of Adh expression is the larval and adult fat body, gut, and Malpighian tubules. There is variable expression along the gut, crop, rectum, and gonadal tissues in different species. Previous studies using interspecific hybrids and P-element-transformed *D. melanogaster* flies have shown that differences in developmental patterns of ADH production are under the control of *cis*-acting regulatory elements linked to structural genes (Brennan et al. 1988; Wu et al. 1990). On the other hand *D. melanogaster* transformants with the Adh gene of *D. affinidisjuncta* show a number of differences that imply both *cis*-acting and *trans*-acting components in the evolutionary divergence of this expression pattern (Brennan and Dickinson 1988). The divergence patterns in Adh gene expression throughout the genus *Drosophila* constitute a model system with which to study the evolution of gene regulation.

D. lebanonensis is an interesting species since its alcohol tolerance is nearly twice as high as that of *D. melanogaster,* which is itself much more tolerant than most *Drosophila* species (David et al. 1974, 1979). In these two species the high ethanol tolerance is correlated with a high alcohol dehydrogenase activity and the higher ADH activity in *D. lebanonensis* is due to a fourfold increase in content of ADH enzyme (Villarroya and Juan 1991). *D. melanogaster* and *D. lebanonensis,* which belong to different subgenera, seem to represent two independent origins of high ADH activity and ethanol tolerance. Therefore this is a rare case revealing where to start to look for the common and distinct features concerning the way in which adaptive changes in gene regulation arise. As we are interested in the evolution of gene regulation using as a model the Adh system, we present in this paper the qualitative and quantitative characterization of the two transcripts of the Adh gene, the profile of gene expression throughout the development, and the tissue distribution of ADH activity.

The availability of the nucleotide sequence of the Adh gene and the closely linked gene Adh-dup of *D. lebanonensis* allowed us to comparatively analyze the evolution of these two genes with a common ancestor but with very diverged functions, as is revealed by the differences in nucleotide substitution rates, gene expression, and codon usage of the two genes. The first molecular datum for the species of the *Scaptodrosophila* subgenus was the amino acid sequence of *D. lebanonensis* ADH which was used to derive the phylogenetic relationships with the other species of the genus at the molecular level (Villarroya and Juan 1991). Although *D. lebanonensis* was clearly separated from the species of the other subgenus, the order of divergence was ambiguous. We investigate here the phylogenetic relationships of *D. lebanonensis* using the nucleotide sequence of the Adh and Adh-dup genes.

Materials and Methods

Drosophila Strains. The *D. melanogaster* stocks Adh^{S} (kindly supplied by M. Aguadé) and Adh^{fn6} cn; ry^{506} were used as controls. The *D. lebanonensis* stock G323 (from Gandesa, Tarragona) was used as a source of genomic DNA and RNA.

Cloning and Sequencing. A 12-kb genomic clone containing the Adh and Adh-dup genes was isolated from a partial BSKS+ plasmid library constructed with the 12-kb *Xba*I genomic fraction from strain 323G (Gandesa, Tarragona) of *D. lebanonensis* (Fig. 1). The fragment *Xba*I-*Eco*RI, which contains Adh, and the fragment *Eco*RI-*Eco*RI, which contains Adh dup, were subcloned in BSKS+ and BSKS−. The sequence of the fragment *Xba*I-*Eco*RI was obtained as described in Juan et al. (1990); 1.4 kb of the *Eco*RI-*Eco*RI fragment was single-strand sequenced with Sequenase using oligonucleotide primers along both strands (EMBL accession No. X63716).

Staging of Larvae. To determine the three larval instars precisely, the morphology of mouth hooks and anterior spiracles was studied throughout the larval development of *D. lebanonensis.* The morphology of mouth hooks and the number of teeth are different from those of *D. melanogaster,* as has been described in Quintana and Juan (1993).

RNA Isolation and Analysis. Total RNA was isolated from different developmental stages of *D. lebanonensis* and *D. melanogaster* Adh^{S} (Chirgwin et al. 1970) and spectrophotometrically quantified. Poly A RNA was purified by oligo dT (Okayama et al. 1987).

Primer extension was carried out as described in Benyajati et al. (1983) using the [γ^{32}]5′-end-labeled primer 5′CAACGAAAATAACG-TTCTTGTTGGT (from +53 to +77), which was gel purified after end labeling, and 25 μg total RNA from different developmental stages. The hybridization was performed at 45°C overnight. Extended primers were run on 7% polyacrylamide, 7 M urea gels. Dried gels were exposed to X-ray film (XAR-5; Kodak) with intensifying screens (Lighting Plus, Dupont) at −80°C.

The mapping of the 5′ ends of transcripts and the donor and acceptor 1st intron splicing sequences was carried out by RNA sequencing using the 5′end-labeled primer 5′TCCATCTTCTGCCTC-TAACTGTTTG (from +25 to +49) following Geliebter (1987). The hybridization temperature was 45°C.

S1 mapping of the 3′ end was carried out using as a probe the fragment *Xho*II-*Dra*I (2,825–3,385, 561 bp, Fig. 2) 3′end-labeled with Klenow, [α^{32}P]dATP, [α^{32}P]dCTP, [α^{32}P]TTP, and dGTP; 20 μg of total RNA and the 3′end-labeled probe (10^6 cpm) were coprecipitated with ethanol, dried, resuspended in 30 μl of hybridization buffer, and incubated at 85°C for 10 min and then for 7 h at 45°C. Annealed samples were diluted 10-fold in S1 nuclease buffer and further incubated at 37°C for 30 min in the presence of 2,200 units/ml of S1 nu-

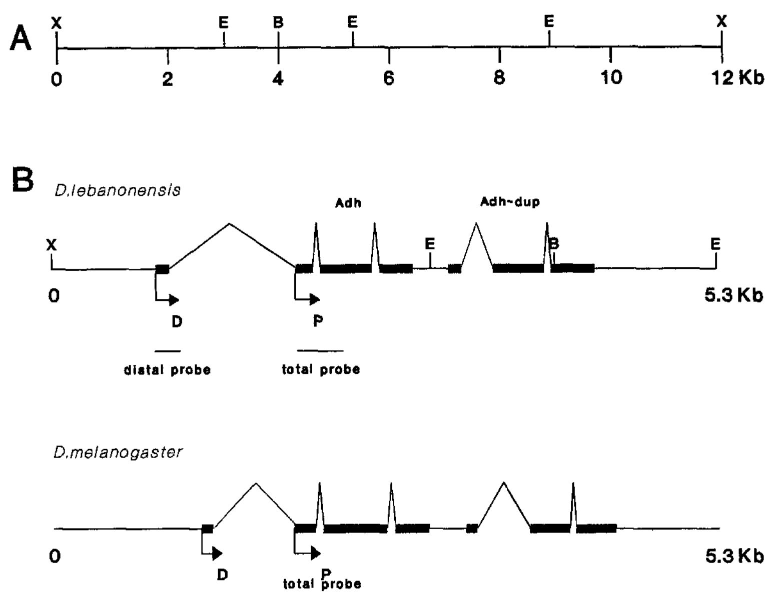

Fig. 1. **A** Restriction map of the 12-Kb *Xba*I-*Xba*I genomic clone of *D. lebanonensis* containing Adh and Adh dup genes. X, *Xba*I; E, *Eco*RI; B, *Bam*HI. **B** Gene structure of Adh and Adh dup genes of *D. lebanonensis* and *D. melanogaster.*

clease. After phenol:chloroform extractions and ethanol precipitation the DNA-protected fragments were run on 6% denaturing polyacrylamide gels. Gels were dried and exposed to X-ray film (XAR-5; Kodak) with intensifying screens (Quanta III, Dupont) at −80°C.

Northern blots were carried out with 8 μg of total RNA per lane. After electrophoresis in formaldehyde 1% agarose, the gels were blotted to a nylon filter. Total probes for Adh of *D. melanogaster* (*Hae*III-*Pvu*II, 402 bp) and *D. lebanonensis* (*Bgl*II-*Hinc*II, 366 bp) and a distal probe of *D. lebanonensis* Adh were prepared independently by random priming using 100 ng of DNA and 3,000 Ci per mmol of [α^{32}P]dCTP (Amersham) to a specific activity of 5×10^8 cpm/μg of DNA. A specific probe for *D. lebanonensis* Adh-dup was prepared by using 3,000 Ci per mmol of [α^{32}P[dATP and [α^{32}P]TTP to a specific activity of 5×10^9 cpm/μg of DNA. Filters were hybridized with equal amounts of the total probes of *D. melanogaster* and *D. lebanonensis* simultaneously at 42°C in formamide overnight. Washings were carried out at 40°C in 0.2 × SSPE, 0.5% SDS. Filters were exposed to X-ray film (XAR-5; Kodak) without intensifying screens for filters hybridized with Adh probes and with two intensifying screens for filters hybridized with the Adh-dup probe. Films were scanned by laser densitometry with an LKB 2222-020 UltroScan XL Laser Densitometer.

Histochemical Staining of Drosophila *Tissues.* ADH activity was analyzed in eggs, larvae, and adult flies as described in Ashburner (1989). The Adh nulo stock Adh^{Fm6} of *D. melanogaster* was used as a control.

Spectrophotometric Determination of ADH Activity. Cultures of *D. lebanonensis* were grown at 22°C. Individuals from different developmental stages were homogenized in 0.2 M Tris HCl pH = 8.6 at a concentration of 143 mg/ml for 30 s and centrifuged for 3 min at 14,000 rpm, and the supernatant was diluted fivefold with the same buffer. The activity was quantified as described in Juan and Gonzalez–Duarte (1980). One unit of activity reduces 1 μmol of NAD/min at 25°C.

DNA Sequence Analysis. Synonymous and nonsynonymous rates of nucleotide substitution were estimated by the method of Li et al. (1985) using a computer program which he kindly provided.

Phylogenetic trees were obtained using the PHYLIP package (version 3.4, Felsenstein 1991) and neighbor-joining method with Kimura's (1980) distances. The program for the neighbor-joining tree with bootstrap (Saitou and Nei 1987) was kindly supplied by K. Tamura.

Results

Transcriptional Analysis of Adh

The structure of the Adh gene of *D. lebanonensis* is similar to that of the previously analyzed species with a single Adh gene. It has four exons and three introns (Figs. 1, 2). RNA sequencing revealed that intron 1 has the longest sequence (1,026 bp) compared to the known Adh genes of other species of *Drosophila* (Fig. 2). Introns 2 and 3 have been deduced from the comparison of the amino acid sequence of *D. lebanonensis* ADH (Villarroya et al. 1989) and the nucleotide sequence (Juan et al. 1990) and they do not differ significantly in length from other *Drosophila* species. Primer extension analysis (Fig. 3) and RNA sequencing revealed two types of transcripts, which originate from two different promoters. The shorter one is synthesized from the proximal promoter mainly in first-instar larvae and the longer one appears in 1-day-old pupae and adults. The types of transcript appear heterogeneous over two nucleotides although there is a major initiation site from each promoter.

Two poly-A signals appear to be used both in larvae and adults; the first (AAUAUA) is less efficient than the second (AAUAAA) as has been previously shown in vitro by Sheets et al. (1990) for poly-A signals in vertebrate mRNAs. Heterogeneity is also observed in the polyadenylation addition sites of transcripts from adults and larvae (Fig. 2). Heterogeneity at the polyadenylation site has also been observed in *D. affinidisjunta* Adh mRNAs (Rowan and Dickinson 1988) and in Pig-1 gene of *D. melanogaster* (Furia et al. 1991). It seems to be a fairly common phenomenon in *Drosophila.*

Adh Gene Expression

The profile of Adh expression is shown in Figs. 4 and 5. There is no apparent transcription of the Adh gene in

```
     +.........+.........+.........+.........+.........+.........+.........+.........+.........+.........+
   1 ctagaaacacactcagtgccccctagttgagaaactgtactcgtaagtcaagtgctgacaataatctacacacatctacttgtacatatttacaaatacg
 101 ttgcatcagcgactctcagagcacatgctcgtattacatatttacatacgagtatgtactaagtacatatatagtacatagtatgcgaacacacagcact
 201 atacaagtataatgcaaactacatttatacttgtatgtaaatgagaaatctactctttcttggcattgagcactggcgtccaatctgttggcacagagc
 301 tctgctggcatcaaataccggtcattcccattcctaggcatgttacaattattggatttgcatacctccgctctattgccatataagcttagtttttgtg
 401 ctcagccttgagtggttcgctattaaaaaggcagctgctcacatcccctggatgactggggtcggctctggtacaatgtgcacctactaggggcctgttt
 501 gtgagctcggcgcacccccaatctttaaattgaaggtgatcttctggtcgctctgttgctgccctgcgcttatttctttgacattcactgaaggtgaaa
 601 acagcatgtgccgggcccggcgggcccggacccgagcgccggtcgcttgtttacatatagttcatacacatacgatacatacgagtatgtacatacgtgg
 701 gcaattgtataagcattattacaaaatatagaatgcaaaatacacacgagtcgtcgggttggcgatgctttcagctcttccgcgttcatctggctgggta

 801 tttaaagcgcgttatacgcctaggacgacgattattgtcagtgcagtggtggcgtcgtgttcagcagcgatcgagaccaacggctaattctaacggaatt
 901 acgagtatttacaaaatctattattggaaacggtgagttgagtgctagaagcgacagtgaagtgtgggagaactttagaagcagtgataaaagagaact
1001 ttaaataaccgaaaataatcaacaaattgccgataaagaacaattaaagtaatagctgtaaaaagctctacacttcgtatggaatttgtagttaatttct
1101 acgactacttttgccagcaaaatagtgatagacaccattcttgatttgttcgttcatctactatgactgcgagaatttttaatgtctaggacaatgatag
1201 agcgataaaagcgactagacactcttccaaacaattagcgactagacacagactacacactcttcctagtaattgcataatttttttttaaatttaata
1301 attaaagtcagacgataagttatcgtctgaaaagcattgaaaaagaaagcccaggcgacgctatagacatcgaactagaactagaaaaagtaagcgtaac
1401 tatgaccaagcgatttcatatcgttctctcagcaaattcaaatatacacgaatttaacgaactagtactaactataaataaagtctggcgcgcgtcttt
1501 ctgtttgctgcccgcccggctgatagcaacaaaatgcttgccggccgctcggtcggcataacgtagaagattataatgctgtggaaaaaaccaaaacaaa
1601 taatatgctggcgaagtactattttcagaatattcttttcgttgcgcgacttcggctagtgtttatcgagcacatgtagagagctacgtaatcgctatta
1701 acagagctctgttaacttaacatacatacatacatacatacatacttatgtgtgtatatgatcgtttgataagacatgaatggtgagcgtgttaatgttg
1801 taccgaaaagaaataaatacatacagacgtttacataaaacaaacgacgcgacagtgacgaggttgtttgctatttgtagcgataagaggcgacgcgatg

1901 tgtcgaaattttcacataaatagcccgactagctaaccagatctcattcgcatttcagcagccagcaaacaaacagttagaggcacaagatggatttgac
                                                                                                 M  D  L
2001 caacaagaacgttatttcgttgccgctctgggcggtattggtctcgacacgagtcgggagctcgtcaagcgtaatctgaaggtgagtccaaccgtaatc
     T  N  K  N  V  I  F  V  A  A  L  G  G  I  G  L  D  T  S  R  E  L  V  K  R  N  L  K
2101 tctggctggagatgaaacgatttaaacgattctgcgtccacagaatttgtcatcttggacagagttgagaacccgactgctcttgctgagctgaaggca
                                               N  F  V  I  L  D  R  V  E  N  P  T  A  L  A  E  L  K  A
2202 ataaatcccaaggtgaacatcaccttccatacctacgatgtgaccgtacccgttgctgagtcgaagaagctgctgaagaagatttcgatcagctgaaaa
       I  N  P  K  V  N  I  T  F  H  T  Y  D  V  T  V  P  V  A  E  S  K  K  L  L  K  K  I  F  D  Q  L  K
2301 ccgttgacattctgatcaatggtgctggtatttggatgaccatcagattgagcgcaccattgccatcaacttcactggtctggtcaacaccaccaccgc
      T  V  D  I  L  I  N  G  A  G  I  L  D  D  H  Q  I  E  R  T  I  A  I  N  F  T  G  L  V  N  T  T  T
2401 tatttggactttgggacaagcgtaagggcggtcctggtggcatcattgccaacatttgctccgttactggcttcaatgccatccaccaggtgcctgtc
     A  I  L  D  F  W  D  K  R  K  G  G  P  G  G  I  I  A  N  I  C  S  V  T  G  F  N  A  I  H  Q  V  P  V
2501 tactcggcatccaaggctgctgttgtcagcttcacaaattccctggcggtaaggagtatagcattggattcctcttatataatattaatgaacatttata
       Y  S  A  S  K  A  A  V  V  S  F  T  N  S  L  A
2601 atgtttttatatagaaactggctcccatcactggcgttactgcctattcgatcaaccctggcatcaccaggacccccttggtgcacactttcaactcctg
                       K  L  A  P  I  T  G  V  T  A  T  S  I  N  P  G  I  T  R  T  P  L  V  H  T  F  N  S
2701 gctggatgtagagccacgtgttgccgaactgctgctgtcccatccaacccagacctccgagcaatgcggtcagaactttgtcaaggccattgaggccaac
     W  L  D  V  E  P  R  V  A  E  L  L  L  S  H  P  T  Q  T  S  E  Q  C  G  Q  N  F  V  K  A  I  E  A  N
2801 aagaatggcgccatatggaaattggatctgggcacccttgaggccattgagtggaccaagcactgggactcgcacatctaagcatttccacgccattgtg
       K  N  G  A  I  W  K  L  D  L  G  T  L  E  A  I  E  W  T  K  H  W  D  S  H  I
2901 ataagatagacgtgtatacggaattgaacaattaatatgaagtaaggctgattgataaacgttccatgttgtcgaacgctgaatggtcgaatttgttatt

3001 tttatgaaaatatatcgaagagaattccaaacaaataaacaaaaactgataactctaattaaaaaatatgataaataattgactgtggatcgggagttag
             Poly A                   Poly A
3101 gtaaacagtagagcagtgctagttcaaccattgaataaacaagtggatatttttctaaataccagtatgttcgacttgacaggcaaaaatgtctgctat
                                                                            M  F  D  L  T  G  K  N  V  C  Y
3201 gtggcggactgtggaggcattgccttggagacctgcaaggtgttgatgacaaaaaatattgcagtgagtagtgaagagaatatgaatgataactggatga
       V  A  D  C  G  G  I  A  L  E  T  C  K  V  L  M  T  K  N  I  A
3301 tagtgaacatcaatattgtttaagtaattgtaatcctattgggtcagccatcatttactaaaaaattgttaaacatattaatttaaacaaatattggaa
3401 atgaaaagacttcacccacgatcgttcctttacataactctgatataatacttattaatatacttcccatatctgttgatctattattaattgcaaattt
3501 atcctctaattctgcccaaagaaattggcaatactgcatagtgttgagaacccgcaggccattgcgcaactgcaatccctcaagcctagcacgcaaatat
                                 K  L  A  I  L  H  S  V  E  N  P  Q  A  I  A  Q  L  Q  S  L  K  P  S  T  Q  I
3601 ttttctggacctatgatgtgaccatggcacgtgccgatatgcagaaatactttgacgaagtgatggttcagatggactatatcgatgtactcatctatgg
      F  F  W  T  Y  D  V  T  M  A  R  A  D  M  Q  K  Y  F  D  E  V  M  V  Q  M  D  Y  I  D  V  L  I  Y
3701 cgccaccttgtgcgatgagacggacatcgatggcaccatcaatacaaacttgacgggtatgatgaacacctgtgccactgtgttgccccatatggacaag
      G  A  T  L  C  D  E  T  D  I  D  G  T  I  N  T  N  L  T  G  M  M  N  T  C  A  T  V  L  P  H  M  D  K
3801 aaaaaggacggcagcggcggtttgatttgaatgtaacctccgtcattggattagacccatcgccagtttttgcgcctacagtgcttctaaatttggtg
       K  K  D  G  S  G  G  L  I  L  N  V  T  S  V  I  G  L  D  P  S  P  V  F  C  A  Y  S  A  S  K  F  G
3901 tgattggatttacgcgcagcttggcggtgattgaacattaaacaattatgtgtagatgaattttgaatgactctttttctctgtaggatcccctgtatta
      V  I  G  F  T  R  S  L  A                                                                D  P  L  Y
4001 tacacaaaacggagttgctgtcatggctgtgtgctgtggtcccacaaaagtatttgtagatcgcgagctaaccgccttcttgccttacggacagtccttt
     Y  T  Q  N  G  V  A  V  M  A  V  C  C  G  P  T  K  V  F  V  D  R  E  L  T  A  F  L  P  Y  G  Q  S  F
4101 gccgatcgcctgcgcactgcccccctgccagtcaacagctgtttgtggccagaacattgtgcgtgccatagaacgcggcgagaatggacagatttggatag
       A  D  R  L  R  T  A  P  C  Q  S  T  A  V  C  G  Q  N  I  V  R  A  I  E  R  G  E  N  G  Q  I  W  I
4201 cggacaagggcggcctggaactggtcaagctgcaaagctattggcatatggcggacgtattcctgcactacatgcagacaaagaggatgattagacaaat
       A  D  K  G  G  L  E  L  V  K  L  Q  S  Y  W  H  M  A  D  V  F  L  H  Y  M  Q  T  K  R  M  I  R  Q
4301 aagaagtttccaaaaaactatatacatttgacgaacaattttgtaaaaatgctgaaagaaagcagctgcttaaataacgtagacatggaatatattgaa
     I  R  S  F  Q  K  T  I  Y  I
4401 atatatgtatatattgacgtacaa
```

Fig. 2. Nucleotide sequence of the genomic region of *D. lebanonensis* which contains Adh and Adh dup genes. TATA box and polyA signals are *underlined. Arrows* indicate the most frequently used initiation sites and *asterisks* the alternative sites in the expression of the Adh gene. *Circles* show the polyadenylation sites and *black dots* the most frequently used polyadenylation sites.

embryos and it begins to be actively transcribed from the proximal promoter in 1st-instar larvae, the level of proximal transcript decreases in 2nd-instar larvae and it remains at a basal level throughout the remaining stages of development (Fig. 3). Transcription from the distal promoter is apparent in 1-day-old pupae and adults and the level of distal transcripts increases dramatically as the adults get old (Fig. 4A,B). Transcription from the distal promoter is at a basal level in larvae and pupae (Figs. 3 and 4). The shortness of the specific proximal probe (13 bp) made it impossible to obtain the profile of transcription from the proximal promoter in Northern blots. As a control of loading we used rRNA in bromide-stained gels (Fig. 4C) since none of the housekeeping genes used was sufficiently similar to give a reliable signal and the rRNA has been shown to be a reliable control (Correa-Rotter et al. 1992).

The quantification of each signal from the films by laser densitometry revealed that the amount of distal Adh mRNA is approximately five times higher in adults

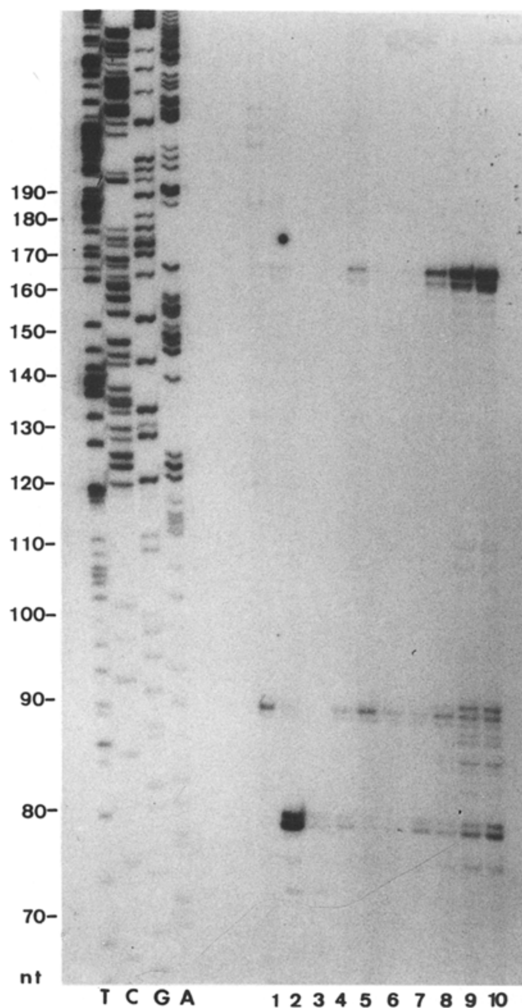

Fig. 3. Primer extension using total RNA from different developmental stages. 1, embryos; 2, 3, and 4, 1st-, 2nd-, and 3rd-instar larvae; 5, 6, and 7, white, red eye, and black pupae; 8, 9, and 10, 1-, 4-, and 8-day-old adults. A sequence reaction using the same end-labeled primer and the mRNA-like single-stranded DNA yields the molecular weight markers.

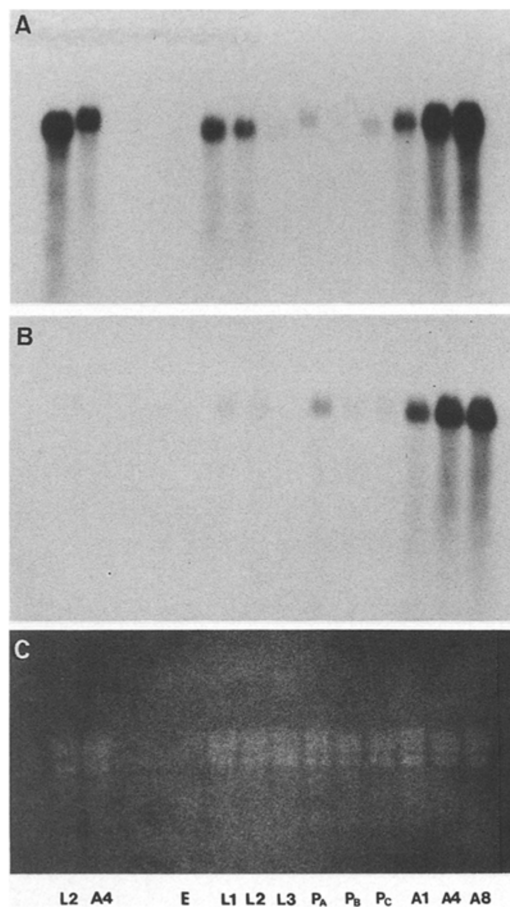

Fig. 4. Analysis of Adh RNA through the development. RNA was prepared from two developmental stages of *D. melanogaster* Adh^S (the first two lanes on the left) and from 10 different developmental stages of *D. lebanonensis* (E, embryos; L1, L2, and L3, 1st-, 2nd-, and 3rd-instar larvae; P_A, 1-day-old pupae; P_B, red eye pupae; P_C, black pupae; A1, A4, and A8, 1-, 4-, and 8-day-old adults; 8 μg of total RNA was loaded onto each track. **A** The filter was simultaneously hybridized with the total probe of *D. lebanonensis* Adh (*Bgl*II-*Hinc*II) and the total probe of *D. melanogaster* Adh (*Hae*III-*Pvu*II) (Fig. 1). **B** After washing out the hybridized total probe it was hybridized with the distal probe (*Dra*I-*Dra*I) of *D. lebanonensis* Adh. **C** Ethidium-bromide-stained gel.

of *D. lebanonensis* than in *D. melanogaster* adults, although in the latter it is much higher in larvae, as is shown in Table 1 and Fig. 4A. In *D. melanogaster* the length of the total probe is 401 bp, of which 381 bp are common to distal and proximal mRNA and the G+C% is 52.1%. In *D. lebanonensis* the length is 366 bp, of which 347 bp are common to distal and proximal mRNA, and the G+C% is 47.7%. The design of these probes and the experimental conditions indicate that the results reported are reliable.

Developmental Profile and Tissue Distribution of ADH Activity

The profile of ADH activity of *D. lebanonensis* is shown in Fig. 5. ADH activity per gram of flies rises during larval development until the third instar; at the end of this period and during the pupal stage the activity falls off, but it increases again after eclosion, reaching a maximum in 6-day-old adults. Notably, there is a delay of 4 days in the detection of the larval maximum ADH activity after the maximum of proximal mRNA is detected. ADH activity profile in adults parallels the mRNA profile.

Tissue distribution of ADH activity was analyzed in eggs in 1st-, 2nd-, and 3rd-instar larvae and in 1-day, 4-day-, and 8-day-old adults. ADH activity was detected in all developmental stages except in eggs. As in *D. melanogaster*, the major sites of ADH activity are the larval and adult guts, fat bodies, Malpigian tubules, and the crop and the rectum of the adults (Fig. 6). The ADH activity pattern of the gut changed in different stages. Second-instar larvae showed ADH activity throughout the midgut (Fig. 6B) while in third-instar larvae it was only detected in the anterior midgut (AMG) (Fig. 6C). The anterior midgut of 1st-instar larvae was also light-

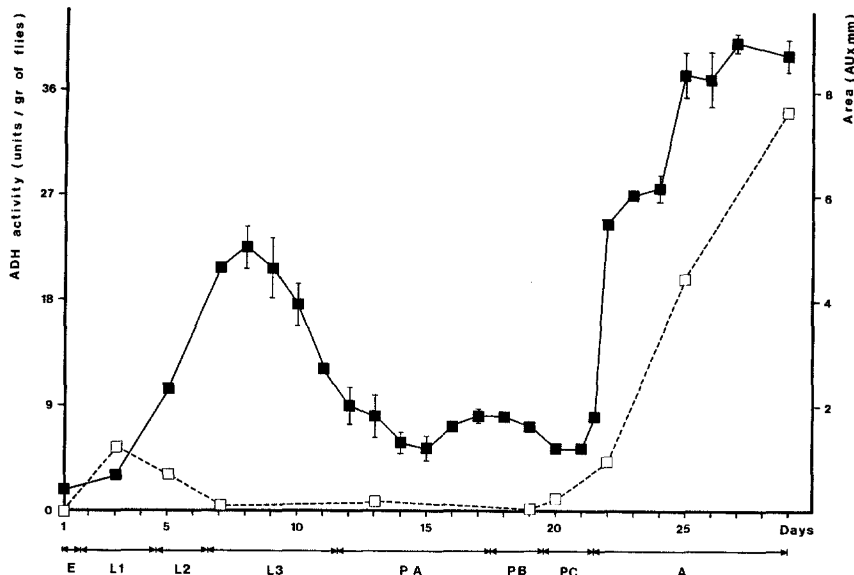

Fig. 5. Adh gene expression during the development (■) ADH activity profile. (□) mRNA profile. E, embryos; L1, L2, and L3, 1st-, 2nd-, and 3rd-instar larvae; P_A, 1–6-day-old pupae; P_B, red-eyed pupae; P_C, black pupae; A, adults.

ly stained. One-day-old adults only show staining of the anterior midgut (Fig. 6G), while 4-day and 8-day adult flies show ADH staining throughout the midgut (Fig. 6H). Ovarioles of 4-day-old adult females (Fig. 6E) and ejaculatory ducts of 4-day-old adult males (Fig. 6F) also show ADH activity.

The pattern of ADH distribution along the gut in *D. lebanonensis* is clearly different from that of *D. melanogaster*. ADH activity is detectable throughout 3rd-instar larvae midgut of *D. melanogaster* (Fig. 6D) while *D. lebanonensis* 3rd-instar larvae only show staining of the anterior midgut (Fig. 6C). Four-day-old adult midgut of this species stains over nearly all its length (Fig. 6H) while in *D. melanogaster* the adult midgut staining is restricted to the cardia (Fig. 6I).

Table 1. Adh gene expression throughout development

	ADH activity (U/gm flies)	Adh transcripts Northern (AU × mm)[b]
D. lebanonensis		
E	1.80 ± 0.00	0.03
L1	3.03 ± 0.06	1.21
L2	10.66 ± 0.21	0.68
L3	20.99 ± 0.21	0.10
PA	8.19 ± 0.85	0.18
PB	8.47 ± 0.17	0.05
PC	5.44 ± 0.19	0.24
A1	24.47 ± 0.13	0.94
A4	37.60 ± 0.98	4.41
A8	38.95 ± 0.67	7.59
D. melanogaster (Adh^s strain)		
L2–L3	ND[a]	3.98
A4	ND	1.60

[a]ND, not determined
[b]Area

Features of the D. lebanonensis *Adh-dup Gene*

Adh-dup of *D. lebanonensis* (Figs. 1, 2) has the same structure as the sequenced Adh-dup genes of a *Sophophora* species (Schaeffer and Aquadro 1987; Kreitman and Hudson 1991; Marfany and Gonzalez-Duarte 1991). The putative promoter CTAATTAAAA, which is not predicted by the method of Bucher (1990), has a transition in the first position compared with the other known Adh-dup genes. This motif overlaps with polyadenylation sites of the second polyadenylation signal of the Adh gene. The sequences GGCAATAAGGCTGATT and AATGG present in the four *Sophophora* species in which the gene has been sequenced are missing in the Adh-dup gene of *D. lebanonensis* as well as the polyadenylation signal AATAAA in the position described for those species. Exons 1 and 2 have the same length as in *Sophophora* species. Interestingly, exon 1 has a residue of phenylalanine or tyrosine in these five species. This residue at the second position in the product of the Adh-dup gene aligns with the third residue in Adh of species of the *melanogaster* group, but it is not present in the Adh of the other species whose Adh sequence is known. Exon 3 has 342 bp, which is the same length as Adh-dup of *D. ambigua* but longer than in the other species. The amino acid percent identity up to residue 264 is 87.5, 84.4, and 82.4 for exons 1, 2, and 3, respectively, but from residue 265 to the C-terminus there is a great variability in length and in amino acid composition. The length of the Adh-dup deduced amino acid sequence varies from 272 to 281 amino acids, as shown in Table 3 and Fig. 7.

The hybridization of the filter used to analyze the expression of the Adh gene with a specific probe for Adh-dup did not yield any signal although the specific activity of the probe was 10 times higher than for Adh

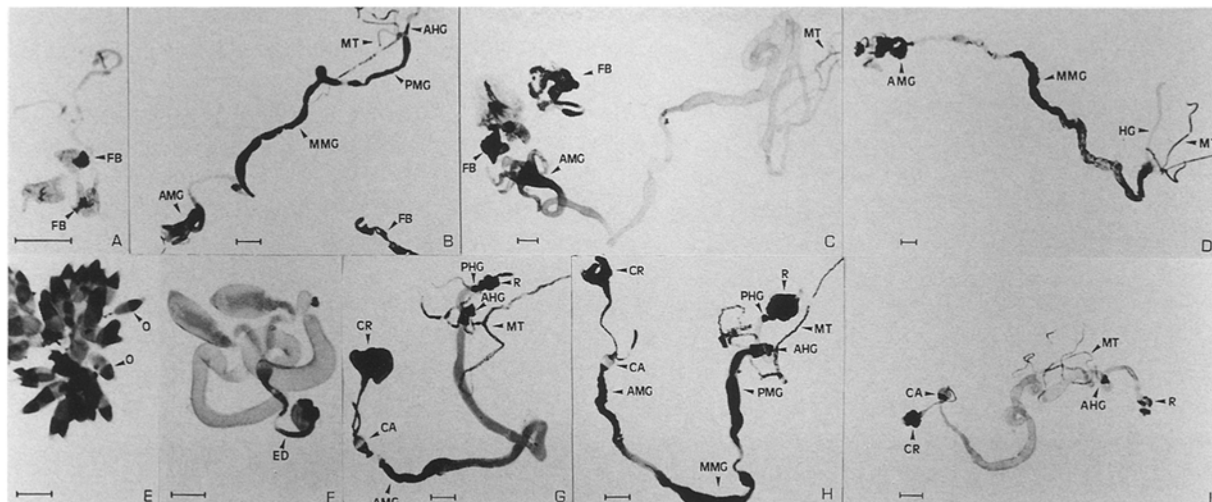

Fig. 6. Histochemical staining for ADH activity in dissected tissues of larvae and adult flies of *D. lebanonensis.* A, B, and C, digestive tract, attached Malpigian tubules, and fat body from 1st-, 2nd-, and 3rd-instar larvae, respectively. G and H, digestive tract and attached Malpigian tubules from 1-day- and 4-day-old adult flies, respectively. E, dissected ovaries from a 4-day-old female. F, reproductive system from a 4-day-old male. D and I, digestive tract and Malpigian tubules from 3rd-instar larvae and 4-day-old adults of *D. melanogaster,* respectively. FB, fat body; AMG, anterior midgut; MMG, mid midgut; PMG, posterior midgut; AHG, anterior hindgut; PHG, posterior hindgut; MT, Malpigian tubules; CR, crop; CA, cardia; R, rectum; O, ovarioles; ED, ejaculatory duct. Bar = 50 μm.

```
                          10        20        30        40        50        60        70        80        90        100       110       120
                          *         *         *         *         *         *         *         *         *         *         *         *
D.lebanonensis   MFDLTGKNVCYVADCGGIALETCKVLMTKNIAKLAILHSVENPQAIAQLQSLKPSTQIFFWTYDVTMARADMQKYFDEVMVQMDYIDVLIYGATLCDETDIDGTINTNLTGMMNTCATVL
D.melanogaster   .......H..............S..............Q.T...........I.................E..K.................N.......NN..A............V....
D.mauritiana     .Y.....H..............S..............Q.T...........I.................EE.K.................N.......NN..A............V....
D.pseudoobscura  .Y.....H..............S..............Q.....P.......I.H........F......EE.K.................N.......RN..A............V....
D.ambigua        .Y..A..H..............S..............Q....QP.......I.H........F......EE.K.................N.......RN..A............V....

 130       140       150       160       170       180       190       200       210       220       230       240       250       260
  *         *         *         *         *         *         *         *         *         *         *         *         *         *
PHMDKKKDGSGGLILNVTSVIGLDPSPVFCAYSASKFGVIGFTRSLADPLYYTQNGVAVMAVCCGPTKVFVDRELTAFLPYGQSFADRLRTAPCQSTAVCGQNIVRAIERGENGQIWIADKGGLELVKLQSYWHMA
.Y..R.IG.T....V.....................................S..............R.......K...E..........R......S.......N....S..................HW.....
.Y..R.MG......V.....................................S..............R.......K...E..........R......S.......N....S..................HW.....
.Y..R.MG......V................................................................E..........R..............N....S..............S.A.HW.....
.Y..R.MG......V.....................................N......................N...E...T......L..............N....SQ...............T.HW.....

 260       270       280
  *         *         *
DVFLHYMQTKRMIRQIRSFQKTIYI
.Q.V....---------.NDEEDQD
.Q.V....---------.NDEEDQD
.Q.VN...STDDED.---EFFLGQR
.Q.IS...STDDED.ECFLSAAYRK
```

Fig. 7 Alignment of the deduced amino acid sequences of the Adh-dup gene of five species of *Drosophila.*

probes and two intensifying screens were used for longer exposure times. Northern blots carried out with 25 μg of total RNA per lane also gave negative results. The expression of Adh-dup must be very low and perhaps specific to a small organ, so its mRNA becomes highly diluted in total RNA and is not detectable in these conditions. Thus it was not possible to detect the start site of transcription and the polyadenylation site experimentally.

Codon Usage in Adh and Adh-dup

The base composition and hence the pattern of synonymous codon usage vary considerably among different *D. melanogaster* genes. The group of genes with a high codon bias is known to have high expression levels (Shields et al. 1988). Apparently in *D. melanogaster* as in unicellular organisms the extent of codon bias is correlated with gene expressivity (Ikemura 1985).

The extent of codon bias for the Adh gene in different species of *Drosophila* varies as is shown in Table 2, the lowest value being that of *D. lebanonensis.* Additional results on the comparison of codon bias and gene expressivity for *D. melanogaster* and *D. simulans* Adh show that the former has a higher level of expression than the latter (Juan and Gonzalez–Duarte 1980; Dickinson et al. 1984); however, both species have identical codon bias. Thus it seems that the codon bias cannot be correlated with gene expressivity among species. Codon preference is less biased for Adh-dup than for Adh in the five species in which both genes have been sequenced, as is shown in Table 2. Interestingly, *D. lebanonensis* Adh-dup shows a value of codon bias similar to the species of *melanogaster* group; however, the value of Adh codon bias is approximately half that of *melanogaster* subgroup species. Thus not even the ratio between Adh and Adh-dup codon bias is the same for the species of *melanogaster* subgroup as for *D. lebanonensis.*

Starmer and Sullivan (1989) found a shift in the

Table 2. Codon bias in Adh and Adh-dup genes

	χ^2_a	
	Adh	Adh-dup
D. melanogaster	0.88	0.11
D. mauritiana	0.92	0.14
D. ambigua	0.91	0.37
D. pseudobscura	0.63	0.40
D. lebanonensis	0.47	0.18

[a] From deviation from random synonymous codon usage, scaled by gene length (Shields et al., 1988)

third-codon-position nucleotide frequency in alcohol dehydrogenase genes when the species of the two subgenera *Drosophila* and *Sophophora* were compared. When the frequency of nucleotides at the third codon position for all codons in Adh of *D. lebanonensis* is compared with the Adhs of seven *Sophophora* and nine *Drosophila* species in which Adh has been sequenced, significant differences were found between *D. lebanonensis* and the species of the *Sophophora* subgenus. However, only some *Drosophila* species genes differed from *D. lebanonensis* Adh. Significant differences were found for *D. mojavensis* and *D. mulleri* but not for *D. hydei* and *D. affinidisjuncta.* Thus the shift in the third-codon-position nucleotide frequency observed for Adh of *Sophophora* and *Drosophila* subgenera does not extend to *D. lebanonensis* of the *Scaptodrosophila* subgenus. When this comparison is performed with the five sequenced Adh-duplo genes (Table 3), no shift is detected.

Evolutionary Differences Between Adh and Adh-dup

Synonymous substitution rates among Adh-dup genes are much higher than synonymous substitution rates for Adh genes while nonsynonymous rates are generally lower among Adh-dup genes than among Adh genes, as is shown in Table 4. This result is reflected in the average percentage identity at the amino acid and nucleotide level of Adh and of Adh-dup genes of the five species compared. The average percentage identities, at the amino acid level, are 76.8 for ADH and 83.7 for the deduced amino acid sequences of Adh-dup. At the nucleotide level they are 70.2 for Adh and 67.5 for Adh-dup. The amino acid sequence for Adh-dup product is better conserved than for ADH although at the nucleotide level more divergence is observed.

Phylogenetic Relationships among Drosophila *Species*

The phylogenetic trees obtained by UPGMA with synonymous substitution rates for Adh and for Adh-dup show the same topology as the trees reconstructed by the neighbor-joining method and maximum parsimony. *D. lebanonensis* is clearly separated from the *Sophophora* species (data not shown).

The phylogenetic relationships of *D. lebanonensis* to 16 *Drosophila* species have been obtained by UPGMA and by bootstrapped neighbor-joining and maximum parsimony analyses with the sequence of the Adh gene. The tree reconstructed by UPGMA, using synonymous substitution rates (Li et al. 1985) with error estimates for the branching points, places the divergence of *D. lebanonensis* from the *Sophophora* subgenus shortly after the divergence from the *Drosophila* subgenus (Fig. 8A). The neighbor-joining method unambiguously places *D. lebanonensis* nearer to the *Sophophora* subgenus than to the *Drosophila* subgenus (Fig. 8B). The topology of the neighbor-joining tree differs from UPGMA for closely related species of the *melanogaster* and *mulleri* subgroups.

The method of maximum parsimony with bootstrap estimates (100 replicates) was used to construct a majority-rule consensus tree (Fig. 9). It has a topology which agrees with the UPGMA for *D. lebanonensis* but not with the topology given by UPGMA for species of *melanogaster* and *mulleri* subgroups, although the number of trees that show the topology of Fig. 9 at those nodes is below 50%.

Discussion

The temporal and spatial profile of expression of the Adh gene of *D. lebanonensis* shows some features distinct from *D. melanogaster* (Savakis et al. 1986):

1. The gene does not seem to be expressed in embryos. It begins to be expressed in larvae from the proximal promoter but the maximum level of transcript is reached at the 1st-instar larvae, decreasing over larval development. Interestingly, the level of transcript is not correlated with ADH enzyme activity in the three instars. ADH activity is very low in 1st larvae and reaches a maximum in 3rd-instar larvae where the level of transcript has reached a minimum. The translation seems to take place in 1st- and 2nd-instar larvae when mRNA is available but the protein is activated progressively until the 3rd instar, where it reaches the maximum activity. Similar results have been reported in *D. melanogaster* and *D. simulans* in which mRNA appears to anticipate changes in both activity and cross-reacting material (Thomson et al. 1991). However, in *D. lebanonensis,* the time interval between the maximum levels of mRNA and ADH activity is doubled. One possible explanation for this result, in larvae, is that the degradation of mRNA is associated with translation which has been described for histone mRNA (Graves et al. 1987) and that the protein would have to be modified to yield the active enzyme. We do not have further evidence to support this interpretation but if this is so it would be a case of posttranslational modification of the ADH, which has not been described up to now.

Table 3. Third-codon-position base utilization for Adh and Adh-dup genes [a]

	A	T	G	C	Total	G + C%
Adh						
Subgenus *Scaptodrosophila*						
D. lebanonensis	15	75	70	94	254	64.5
Adh-duplo						
Subgenus *Sophophora*						
D. melanogaster	50	71	82	69	272	55.5
D. mauritiana	50	68	89	74	281	58.0
D. pseudobscura	26	62	91	99	278	68.3
D. ambigua	30	62	96	93	281	67.2
Subgenus *Scaptodrosophila*						
D. lebanonensis	52	68	75	86	281	57.3

	D. mauritiana	*D. pseudobscura*	*D. ambigua*	*D. lebanonensis*
D. melanogaster	1.25	14.53*	9.76*	2.04
D. mauritiana		10.85*	7.70	2.14
D. pseudobscura			0.37	11.32*
D. ambigua				9.01*

[a] Comparisons made with 2 × 4 contingency χ^2 statistics (*$P < 0.05$, 3 df)

Table 4. Synonymous (above diagonal) and nonsynonymous (below diagonal) substitution rates for Adh-dup and Adh genes[a]

	D. melanogaster	*D. mauritiana*	*D. pseudoobscura*	*D. ambigua*	*D. lebanonensis*
D. melanogaster	—	0.16 (0.03)	1.29 (0.18)	1.14 (0.15)	2.15 (0.60)
D. mauritiana	0.006 (0.003)	—	1.18 (0.16)	1.07 (0.14)	2.23 (0.89)
D. pseudoobscura	0.043 (0.008)	0.36 (0.008)	—	0.42 (0.06)	1.60 (0.11)
D. ambigua	0.052 (0.009)	0.044 (0.008)	0.020 (0.006)	—	1.42 (0.22)
D. lebanonensis	0.101 (0.013)	0.103 (0.013)	0.108 (0.014)	0.116 (0.014)	—

	D. melanogaster	*D. mauritiana*	*D. pseudoobscura*	*D. ambigua*	*D. lebanonensis*
D. melanogaster	—	0.045 (0.016)	0.63 (0.08)	0.60 (0.08)	0.83 (0.11)
D. mauritiana	0.009 (0.004)	—	0.68 (0.09)	0.63 (0.08)	0.86 (0.11)
D. pseudoobscura	0.052 (0.009)	0.057 (0.010)	—	0.31 (0.05)	0.88 (0.12)
D. ambigua	0.043 (0.009)	0.049 (0.009)	0.028 (0.007)	—	0.85 (0.11)
D. lebanonensis	0.108 (0.014)	0.110 (0.014)	0.101 (0.014)	0.112 (0.014)	—

[a] Values in parentheses are standard errors.

2. Adh of *D. lebanonensis* is expressed at a low level from the distal promoter in 1-day-old pupae and the level of transcript decreases until the eclosion, when the gene begins to be actively transcribed from the distal promoter.
3. *In vivo* the proximal and distal transcripts accumulate to different steady-state levels and at different rates in larvae and adult flies. The amount of proximal transcript in 1st-instar larvae is one-sixth of the amount of distal transcript in 8-day-old flies and the maximum level of proximal transcript is reached in 1 day. The maximum level of distal transcript is reached on the eighth day in adult flies. So the situation seems to be very different than that of *D. melanogaster* Adh, where the maximum of proximal transcript is reached at the beginning of 3rd-instar larvae (Savakis et al. 1986) and where the proximal transcript in larvae is more abundant than the distal transcript in adults.
4. Primer extension analysis shows that *D. lebanonensis* Adh is transcribed from the proximal promoter at a basal level in adults. This situation is also different from *D. melanogaster,* in which proximal transcripts are not detected in adults (Savakis et al. 1986) and a transcriptional interference mechanism has been proposed to shut off transcription from proximal promoter at the end of 3rd-instar larvae (Corbin and Maniatis 1989). If transcriptional interference was applicable to Adh of *D. lebanonensis,* a possible explanation for the presence of proximal transcripts in adults would be that they are produced in a different tissue from the abundant distal transcript.
5. Spatial distribution of Adh in *D. lebanonensis* also shows clear differences from *D. melanogaster* both in larvae and adult guts.

Thus the differential features of the Adh expression pattern in *D. lebanonensis* predict differences in the regulation of this gene. It has been shown that regula-

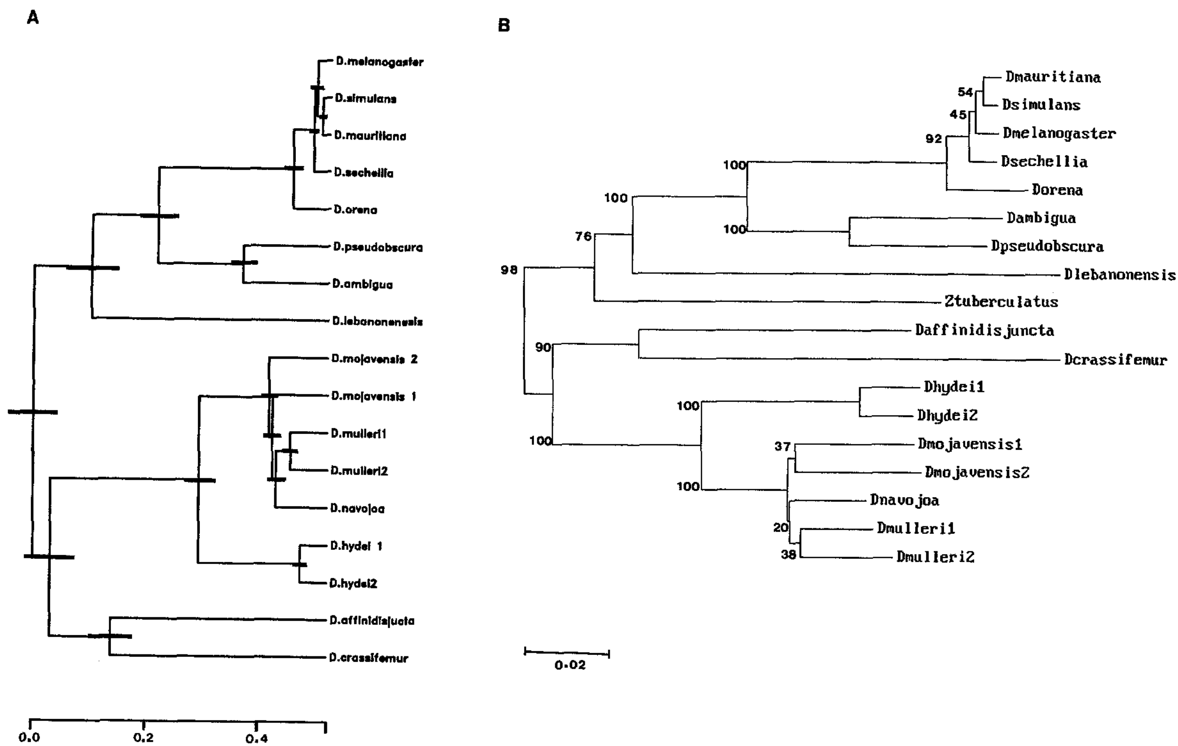

Fig. 8. A UPGMA constructed using the synonymous substitution rate (K_s) (Li et al. 1985). **B** Unrooted phylogenetic tree constructed with the coding region of Adh using the neighbor-joining method (Saitou and Nei 1987) and Kimura's two-parameter distance (Kimura 1980). The numbers on the tree are the bootstrap probability values based in 1,000 replications.

tory elements are functionally conserved in closely related species which show similar patterns of alcohol dehydrogenase expression (Moses et al. 1990). The alignment of the 5′region flanking the distal promoter of the Adh gene of *D. melanogaster* and *D. lebanonensis* is very poor, although some motifs that have been shown to bind transcription factors (Ayer and Benyajati 1990, 1992; Benyajati et al. 1992) have been conserved (Fig. 10). The octamer motif ATGCAAATTA (Singh et al. 1986) included in the negative element described by Ayer and Benyajati (1990) is 90% conserved but the second motif of this element GCAACAAC (DTF-1, Perkins et al. 1988) is not found in the sequence. The positive element DEP1-DEP2 (Ayer and Benyajati 1992) is 95% conserved. The repeats Adf2a and Adf2b which are sites for the binding of Adf2 repressor factor in *D. melanogaster* Adh-negative cells (Benyajati et al. 1992) are partially conserved in *D. lebanonensis* where Adf2a and Adf2b show 90% and 80% identity, respectively. The sequence from position 601 to 642 only is 46% identical to the Adf1 site described in *D. melanogaster* Adh (Heberlein et al. 1985), but after alignment it shows two inverted and two direct repeats at a similar position to the Adf1 motif. Although the order and the sequence of these motifs are highly conserved except for Adf1, the spacing among them is very different from that of *D. melanogaster*. Experiments in progress may demonstrate whether these elements are functionally conserved in the Adh gene of *D. lebanonensis*.

The abundance of Adh mRNA and protein in *D. lebanonensis* and a lower value of codon bias than in *Sophophora* species is surprising at first glance but the relation between codon bias and gene expressivity seems to be valid only within a species because it may depend on the G+C content of genomic DNA and on the composition of tRNA population in a particular species. What is more surprising is that the ratio of Adh and Adh-dup codon bias is very different in the different species. So in *melanogaster* subgroup species this ratio is 8; in species of the obscura group has an average of 2 and in *D. lebanonensis* 2.6. Therefore although the trend "the more biased the codon usage, the higher the expression of a gene" (Ikemura 1985) seems to be broadly true, it may not fit a perfect correlation. We would need more data on DNA composition, codon usage, composition of tRNA populations, and gene expression in the different species to draw correct conclusions.

Adh-dup of *D. lebanonensis* does not show any known motif in the spacer between Adh and Adh-dup coding region. Even the sequence that has been described as the putative promoter (Schaeffer and Aquadro 1987) is not predicted by the method of Bucher (1990), which clearly predicts the two promoters of Adh. In contrast with the Adh gene, whose product is very abundant in the fly, the product of Adh-dup is undetectable in the conditions assayed. Nevertheless, the deduced amino acid sequence of this gene product is better conserved than that of ADH in the five species compared. Adh-dup appears to be under stronger evolutionary constraints,

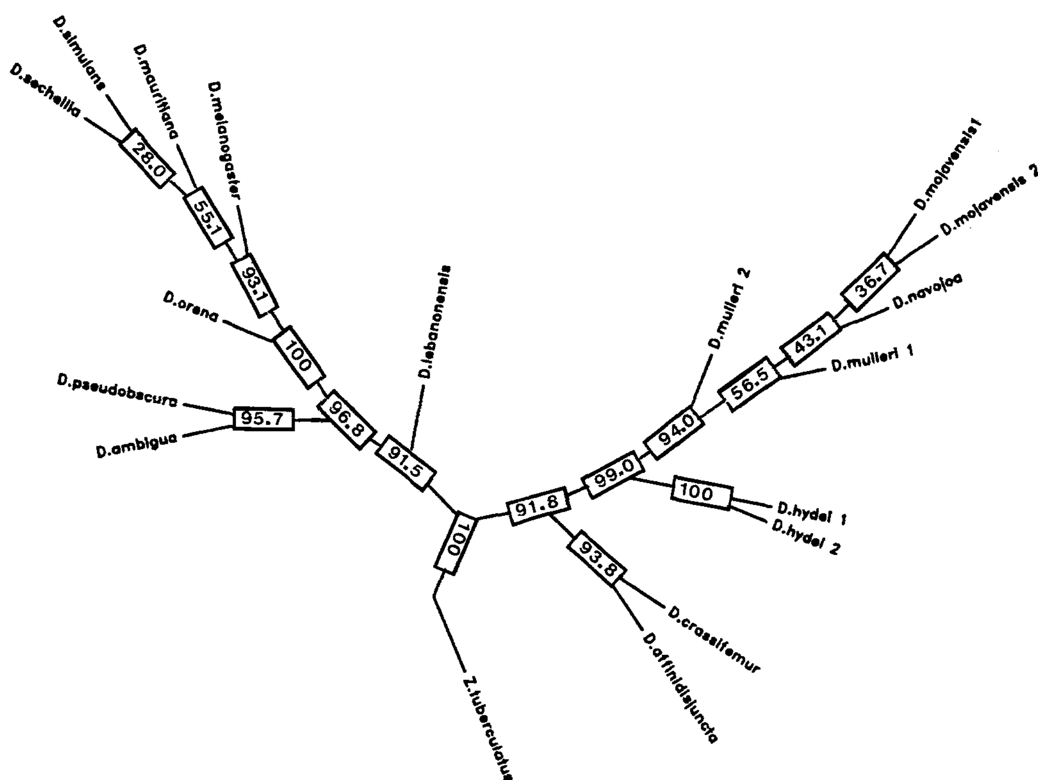

Fig. 9. Maximum parsimony three for Adh coding regions of 17 species of *Drosophila. Numbers* at the nodes show the number of times, per 100 replicates of bootstrapped parsimony, that the set of those descendent taxa occurs together.

```
                          Octamer
                 -574                            -544
D.melanogaster     tattacacgt atgcaaatta agccgaagtt
                 -618                            -588
D.lebanonensis     tacaagtata atgcaaacta catttatact


                          DEP1-DEP2
                 -500                              -470
D.melanogaster     gctatgcttg acattcac   aaggtcaaagct
                 -258                              -227
D.lebanonensis     tatttcttg acattcactg aaggtgaaaaca


                          Adf2a
                 -215                                -183
D.melanogaster     tcacttatttgtt tctcagtgca ctttctggtg
                 -730                                -698
D.lebanonensis     ttgcatcagcgac tctcagagca catgctcgta


                          Adf2b
                  -6                                  -27
D.melanogaster     tcatgcattattg tctcagtgca gtt gtcagtt
                  -6                                  -23
D.lebanonensis     acgacgattattg   tcagtgca gtggtggcgtc


                          Adf1
                  -90                                                       -41
D.melanogaster     aacgccgctg ctgctgcatc cgtcgacgtc gactgcactc gcccccacga
                 -233                                                      -184
D.lebanonensis     aaaacagcat gtgccgggcc cggcgggccc ggacccgagc gccggtcgct
```

Fig. 10. Alignment of sequences that have been proved to bind transcription factors in *D. melanogaster* Adh. *Bold letters* indicate identical nucleotides in both sequences.

which do not allow the gene product, to residue 265, to change as fast as ADH. However, from this residue to the end of the molecule these constraints seem to be more relaxed, since variation in length and composition is observed throughout the genus. Higher synonymous substitution rates than for the Adh gene could be explained by the higher bias in the codon usage (Ikemura 1985). The conservation of the amino acid sequence seems to predict that the gene product has an important function in *Drosophila,* but the low expressivity which makes it undetectable in the conditions assayed predicts a very different function of that of ADH.

The alignment of the phenylalanine or tyrosine at the second position of Adh-dup with the phenylalanine at

the third position of Adh in the species of *melanogaster* subgroup suggests that this residue was present in the ancestral gene, which gave rise to Adh and Adh-dup by duplication before the divergence of *Sophophora* and *Scaptodrosophila.* The loss of this residue in the ADH of species of other groups and subgenera would represent independent events, at least in the lineages which gave rise to the *obscura* group and *Scaptodrosophila.*

The phylogenetic tree obtained by UPGMA using synonymous substitution rates for Adh shows that *Sophophora* and *Scaptodrosophila* diverged shortly after the separation from the *Drosophila* subgenus; thus the ambiguous order of divergence obtained with the amino acid sequence data (Villarroya and Juan 1991) disappears. However, the UPGMA constructed with Kimura's two-parameter model for the whole coding region of Adh gives the same ambiguity as with amino acid data (data not shown). The neighbor-joining method (Saitou and Nei 1987) using Kimura's two-parameter model groups *D. lebanonensis* with the species of the *Sophophora* subgenus in all bootstrap replicates, while the bootstrap analysis of the parsimony tree indicates that this is supported in 91% of the replicates.

The phylogenies in Figs. 8 and 9 and the existence of Adh-dup in the same position as for the species of the *Sophophora* subgenus indicate that the phylogenetic relationships between *Sophophora* and *Scaptodrosophila* are reliable.

Acknowledgments. We thank M. Aguadé, A. Villarroya, and C. Segarra for valuable comments. We also thank N. Mendez for excellent technical assistance and R. Rycroft for English revision. Financial support was provided by a DGYCIT grant PB87-0136 to E. Juan.

References

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

Ayer S, Benyajati C (1990) Conserved enhanced and silencer elements for differential Adh transcription in *Drosophila* cell lines. Mol Cell Biol 10:3512–3523

Ayer S, Benyajati C (1992) The binding site of a steroid hormone receptor-like protein within the *Drosophila* Adh adult enhancer is required for high levels of tissue specific alcohol dehydrogenase expression. Mol Cell Biol 12:661–673

Ashburner M (1989) *Drosophila.* A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

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

Benyajati C, Ewel A, McKeon J, Chovav M, Juan E (1992) Characterization and purification of Adh distal promoter factor 2, Adf-2, a cell-specific and promoter-specific repressor in *Drosophila.* Nucleic Acids Res 20:4481–4489

Brennan MD, Wu CY, Berry AJ (1988) Tissue-specific regulatory differences for the alcohol dehydrogenase genes of Hawaiian *Drosophila* are conserved in *D. melanogaster* transformants. Proc Natl Acad Sci USA 85:6866–6869

Brennan MD, Dickinson WJ (1988) Complex developmental regulation of the *Drosophila affinidisjuncta* alcohol dehydrogenase gene in *D. melanogaster.* Dev Biol 125:64–74

Bucher P (1990) Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J Mol Biol 212:563–578

Chirgwin JM, Przybyla AE, Macdonald RJ, Rutter WJ (1970) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochem J 18:5294–5299

Corbin V, Maniatis T (1989) Role of transcriptional interference in the *D. melanogaster* Adh promoter switch. Nature 337:279–282

Correa-Rotter R, Mariash CN, Rosenberg ME (1992) Loading and transfer control for Northern hybridization. Biotechniques 12:154–158

David J, Fouillet P, Arens MF (1974) Comparison de la sensibilité à l'alcool éthilique de six espéces de *Drosophila* du sous groupe *melanogaster.* Archs Zool Exp Gén 115:401–410

David JR, Van Herrewege J, Monclus M, Prevosti A (1979) High ethanol tolerance in two distantly related *Drosophila* species: a probable case of recent convergent evolution. Comp Biochem Physiol 63C:53–56

Dickinson WJ, Rowan RG, Brennan MD (1984) Regulatory gene evolution: adaptive differences in expression of alcohol dehydrogenase in *D. melanogaster* and *D. simulans.* Heredity 52:215–225

Fang XM, Wu CY, Brennan MD (1991) Complexity in evolved regulatory variation for alcohol dehydrogenase in Hawaiian *Drosophila.* J Mol Evol 32:220–226

Fischer JA, Maniatis T (1985) Structure and transcription of the *D. mulleri* alcohol dehydrogenase genes. Nucleic Acids Res 13:6899–6916

Furia M, Digilio FA, Artiaco D, D'Avino PP, Cavaliere D, Polito LC (1991) Molecular organization of the *D. melanogaster* Pig-1. J Mol Evol 101:49–54

Felsenstein J (1991) Phylip Version 3.4 University of Washington, Seattle

Geliebter J (1987) Dideoxynucleotide sequencing of RNA and uncloned cDNA. Focus 9.1:5–8

Goldberg DA, Posakony JW, Maniatis T (1983) Correct developmental expression of a cloned alcohol dehydrogenese gene transduced into the *Drosophila* germ line. Cell 34:59–73

Graves RA, Pandey NB, Chodchoy N, Marzluff WF (1987) Translation is required for regulation of histone mRNA degradation. Cell 48:615–626

Heberlein U, England B, Tjian R (1985) Characterization of *Drosophila* transcription factors that activate the tandem promoters of the alcohol dehydrogenase gene. Cell 41:965–977

Ikemura T (1985) Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol 2:13–34

Juan E, Gonzalez-Duarte R (1980) Purification and enzyme stability of alcohol dehydrogenase from *D. simulans, D. virilis* and *D. melanogaster* Adh^S. Biochem J 189:105–110

Juan E, Papaceit M, Quintana A (1990) Nucleotide sequence of the Adh gene of *D. lebanonensis.* Nucleic Acid Res 18:6420

Kreitman M, Hudson RR (1991) Inferring the evolutionary histories of the Adh and Adh-dup loci in *D. melanogaster* from patterns of polymorphism and divergence. Genetics 127:565–582

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

Li WH, Wu CI, Luo CC (1985) A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelyhood of nucleotide and codon changes. Mol Biol Evol 2:150–174

Marfany G, Gonzalez-Duarte R (1991) The Adh genomic region of *D. ambigua:* evolutionary trends in different species. J Mol Evol 32:454–462

Moses K, Heberlein U, Ashburner M (1990) The Adh gene promoters of *D. melanogaster* and *D. orena* are functionally conserved and share features of sequence structure and nuclease-protected sites. Mol Cell Biol 10:539–548

Okayama H, Kawaichi M, Brownstein M, Lee F, Yokota T, Arai K (1987) High efficiency cloning of full length cDNA. Construction and screening of cDNA expression libraries for mammalian cells. Methods Enzymol 154:3–28

Perkins KK, Dailey GM, Tjian R (1988) In vitro analysis of the Antenapedia P2 promoter: identification of a new *Drosophila* transcription factor. Genes Dev 2:1615–1625

Quintana A, Juan E (1993) Morphology of mouth hooks and anterior spiracles through the larval development of *D. lebanonensis*. DIS 72:88

Rowan RG, Dickinson WJ (1988) Nucleotide sequence of the genomic region encoding alcohol dehydrogenase in *D. affinidisjuncta*. J Mol Evol 28:43–54

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

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

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

Schaeffer SW, Aquadro CF (1987) Nucleotide sequence of the Adh gene region of *D. pseudoobscura:* evolutionary change and evidence of an ancient gene duplication. Genetics 117:61–73

Sheets MD, Ogg SC, Wickens MP (1990) Point mutations in AAUAAA and the poly A addition site: effects on the accuracy and efficiency of cleavage and polyadenylation *in vitro*. Nucleic Acid Res 18:5799–5805

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

Singh H, Sen R, Baltimore D, Sharp PA (1986) A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature 319:154–158

Thompson MS, Jacobson JW, Laurie CC (1991) Comparison of alcohol dehydrogenase expression in *D. melanogaster* and *D. simulans*. Mol Biol Evol 8:31–48

Villarroya A, Juan E, Egestad B, Jörnvall H (1989) The primary structure of alcohol dehydrogenase from *D. lebanonensis*. Extensive variation within "short chain" alcohol dehydrogenase lacking zinc. Eur J Biochem 180:191–197

Villarroya A, Juan E (1991) ADH and phylogenetic relationships of *D. lebanonensis* (*Scaptodrosophila*). J Mol Evol 32:421–428

Wu CY, Mote J, Brennan M (1990) Tissue-specific expression phenotypes of Hawaiian *Drosophila* Adh genes in *D. melanogaster* transformants. Genetics 125:599–610