RESULTATS

CAPÍTOL I

En aquest capítol, 1) es recull la caracterització molecular de l'alcohol deshidrogenasa de classe 3 de dues espècies de cefalocordats (***Branchiostoma lanceolatum*** i ***Branchiostoma floridae***). 2) S'estudia la variabilitat genètica d'aquests gens a nivell codificant. 3) Es determina l'activitat enzimàtica enfront diferents substrats, confirmant que es tracta d'una típica ADH3. 4) S'analitza el grau de conservació estructural de la proteïna, comparant-la amb altres a nivell evolutiu. 5) L'anàlisi Southern indica que l'ADH3 és un gen de còpia única localitzat en la fracció metilada del genoma de l'amfiox. 6) Finalment, s'analitza el patró d'expressió per hibridació ***in situ*** durant el desenvolupament embrionari de ***B. floridae***, revelant un patró específic en l'intestí que contrasta amb el patró ubic descrit a mamífers.

Eur. J. Biochem. **267**, 6511–6518 (2000) © FEBS 2000

Amphioxus alcohol dehydrogenase is a class 3 form of single type and of structural conservation but with unique developmental expression

Cristian Cañestro[1], Lars Hjelmqvist[1,2], Ricard Albalat[1], Jordi Garcia-Fernàndez[1], Roser Gonzàlez-Duarte[1] and Hans Jörnvall[2]

[1]*Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Spain;* [2]*Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden*

The coding region of amphioxus alcohol dehydrogenase class 3 (ADH3) has been characterized from two species, *Branchiostoma lanceolatum* and *Branchiostoma floridae*. The species variants have residue differences at positions that result in only marginal functional distinctions. Activity measurements show a class 3 glutathione-dependent formaldehyde dehydrogenase, with k_{cat}/K_m values about threefold those of the human class 3 ADH enzyme. Only a single ADH3 form is identified in each of the two amphioxus species, and no ethanol activity ascribed to other classes is detectable, supporting the conclusion that evolution of ethanol-active ADH classes by gene duplications occurred at early vertebrate radiation after the formation of the amphioxus lineage. Similarly, Southern blot analysis indicated that amphioxus ADH3 is encoded by a single gene present in the methylated fraction of the amphioxus genome and northern blots revealed a single 1.4-kb transcript. *In situ* experiments showed that amphioxus *Adh3* expression is restricted to particular cell types in the embryos. Transcripts were first evident at the neurula stage and then located at the larval ventral region, in the intestinal epithelium. This tissue-specific pattern contrasts with the ubiquitous *Adh3* expression in mammals.

Keywords: amphioxus; ADH3; biochemical properties; *in situ* hybridization; methylation.

Alcohol dehydrogenase/aldehyde reductase (EC 1.1.1.1) was one of the first mammalian enzymes crystallized [1]. It was initially described as a liver protein, but is now known to belong to a complex enzyme system [2] with extensive multiplicity within the medium-chain dehydrogenase/reductase (MDR) family, which includes hundreds of characterized forms [3]. Different classes and isozymes have emerged from repeated gene duplications during vertebrate evolution [2]. In mammals, at least six different alcohol dehydrogenase classes are presently known [4,5], of which five have been characterized in human tissues. Of these classes, 1 and 3 are the ones best investigated and also the ones best known in species variants. The class 1 enzyme (ADH1) is the classic ethanol-active form, evolves rapidly and exhibits a considerable variability between different species [6]. The ADH3 form, identical to glutathione-dependent formaldehyde dehydrogenase [7], is present in prokaryotes and in all eukaryotes thus far investigated. The expression of ADH3 in mammals is ubiquitous, with transcripts found in all tissues analyzed [8] and increased protein levels in the large intestine, liver, and kidney [9]. It is apparent in mouse embryogenesis already by 6.5 days post-coitum and continuously nearly ubiquitously through 9.5 days post-coitum [10]. This transcript distribution pattern would be consistent with its proposed housekeeping role in cytoprotection by metabolism of formaldehyde [11]. Other vertebrate ADH classes, such as the liver enzyme of class 1, show more restricted expression at the tissue level and at different developmental stages. ADH3 exhibits little variability, both overall and in functional segments [6], compatible with a defined role in formaldehyde oxidation. Class 3 appears to be the ancient form from which class 1 has emerged by 'enzymogenesis' [12] via a gene duplication at early vertebrate times, as deduced from an apparent absence of an ethanol-active class 1 form in lines originating before bony fish [13], the presence of a class 1/3-mixed form in fish [14], and calculation of the evolutionary rate [15]. Nevertheless, no data have been gathered at the level of amphioxus or other species relevant to the origin of vertebrates. In this regard, cephalochordates are particularly interesting as they are the closest living relatives of vertebrates. This is supported by ribosomal RNA and mitochondrial DNA sequence analysis, together with much anatomical and embryological data [16–19]. Moreover, cephalochordate gene complexity and body plan organization have led to the assumption that lancelets are 'archetypal' organisms [20,21]. There are several examples where single copy genes in amphioxus are orthologous to gene families in vertebrates [21]. We now report the protein and nucleotide coding sequence of amphioxus alcohol dehydrogenase (protein, ADH3; gene, *Adh3*) and describe its biochemical features. ADH3 is encoded by a single copy amphioxus gene in the methylated genomic fraction. *Adh3* expression analyzed by whole-mount *in situ* hybridization shows a tissue-specific pattern during embryogenesis, contrary to the situation reported for vertebrates. Overall, these data show that not only gene duplications but also variations in the gene expression may have occurred at the origin of vertebrates.

Correspondence to H. Jörnvall, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden.
Fax: + 46 8337462, Tel.: + 46 87287702,
E-mail: Hans.Jornvall@mbb.ki.se

Abbreviations: ADH, alcohol dehydrogenase protein; *Adh3*, alcohol dehydrogenase/glutathione-dependent formaldehyde dehydrogenase gene; *AmphiMLC-alk*, alkali myosin light chain; MDR, medium chain dehydrogenase/reductase.

Enzyme: alcohol dehydrogenase/aldehyde reductase (EC 1.1.1.1).

(Received 17 May 2000, revised 7 August 2000, accepted 9 August 2000)

MATERIALS AND METHODS

Genomic DNA and library screenings

Branchiostoma lanceolatum animals were kindly provided by the Laboratoire Arago (Observatoire Océanologique de Banyuls-sur-mer, France). The animals were kept at −70 °C until used. Total genomic DNA was isolated using the guanidine isothiocyanate method [22] with minor modifications [23]. A two-animal *B. lanceolatum* genomic library was constructed with Lambda FIX-II/*Xho*I partial fill-in vector. A cDNA library of *B. floridae* 6–20 h embryos in Lambda Zap II was kindly provided by J. Langeland (Dept of Biology, Kalamazoo College, MI, USA) [24].

Degenerate oligonucleotides were designed to amplify a *Drosophila* genomic fragment of the *Adh3* gene [25] and this fragment was used to screen the *B. floridae* cDNA library. The probe was labeled with [α-^{32}P]dCTP by random-hexamer priming. A low stringency hybridization was carried out in phosphate/SDS solution [26] at 55 °C overnight. Washes were performed at 55 °C for 2 × 15 min in 2 × NaCl/Cit, 0.1% SDS, and 1 × 15 min in 1 × NaCl/Cit, 0.1% SDS. Positive clones were isolated and characterized by sequence analysis. The *B. floridae* ADH3 cDNA was then used as a probe to screen the *B. lanceolatum* genomic library. Hybridization and washes were performed as mentioned above but at 65 °C. DNA fragments from positive recombinant phages were isolated, subcloned into a pUC18 vector, characterized by restriction mapping and sequenced. The *B. lanceolatum* ADH3 coding sequence was deduced from the corresponding genomic structure.

Protein purification and characterization

B. lanceolatum adults (50 g) were homogenized in 200 mL buffer of 10 mM Tris/HCl, pH 8.0, and centrifuged at 20 000 ***g***. The supernatant was adjusted with Tris base to the pH of the buffer and loaded onto a DEAE Sepharose FastFlow column after dilution to reduce the conductivity. The column was washed, and then eluted with a 600 mL gradient of 0–0.5 M NaCl in the buffer. The most active fractions were pooled, Tris was added to 50 mM, and ammonium sulfate to 1.7 M. After centrifugation, the supernatant was loaded onto a Phenyl Superose High Performance column and eluted with a gradient of 1.7–0 M ammonium sulfate in 50 mM Tris/HCl, pH 8.0. The enzyme, recovered in the latter part of the gradient, was dialyzed overnight against 20 mM Tris/acetate, pH 9.0, and submitted to fast protein liquid chromatography (FPLC) on Mono Q in this buffer with a 0–0.5 M NaCl gradient. All buffers contained 0.1 mM dithiothreitol and all steps were carried out at 4 °C. Protein quantification during purification used the Bradford method [27], and enzyme activity was monitored by NAD^+ reduction with *S*-hydroxymethylglutathione as substrate [7]. Protein purity was analyzed by SDS/PAGE with Coomassie brilliant blue staining and by isoelectric focusing with Nitro blue tetrazolium/phenazine methosulfate activity staining.

Enzymatic activities were tested at 25 °C by monitoring NAD^+ reduction measured at 340 nm. K_m and k_{cat} values were determined at pH 8.0 in 0.1 M sodium pyrophosphate (with *S*-hydroxymethylglutathione and NAD^+) or at pH 10.0 in 0.1 M glycine/NaOH buffer (with 12-hydroxydodecanoate and octanol). Constants were calculated with the program ENZYME [28], and k_{cat} values are given per dimer. Protein molecular masses were analyzed by matrix-assisted laser desorption time-of-flight mass spectrometry with a Lasermat 2000 instrument.

Southern blot analysis and methylation pattern

Total genomic DNA from a single *B. floridae* animal was isolated as given above. 10 μg of *Hpa*II-digested or *Msp*I-digested DNA were resolved on 0.9% agarose gels and transferred to nylon membranes. Southern blots were hybridized with three ^{32}P-labeled probes: 750 bp of the 3′ end of the *AmphiMLC-alk* cDNA [29], a cDNA fragment corresponding to exons 2 and 3 (221 bp from nucleotide position 53–274) and a 345 bp *Eco*RI genomic segment containing exon 8 of *B. floridae Adh3*. Hybridization was performed in phosphate/SDS solution [26] at 65 °C overnight and washes were at 65 °C as for the genomic library screening.

Northern blot analysis

For northern blot analysis, total RNA from adult *B. lanceolatum* animals was isolated by the guanidinium thiocyanate method [30]. Poly(A)$^+$ RNA was purified using oligo (dT)-cellulose type 7 matrix, resolved by 1.2% agarose/formaldehyde gel electrophoresis, and transferred to nylon membranes. Northern blot hybridization was performed in phosphate/SDS solution [26] at 65 °C overnight with *B. floridae* ^{32}P-labeled cDNA probes of *Adh3* or of the 750-bp 3′ end of the *AmphiMLC-alk* [29]. Washes were at 65 °C as for the genomic library screening.

Whole-mount *in situ* hybridization

B. floridae embryos were collected, fixed, and stored [31]. Whole-mount *in situ* hybridization was performed as described [31] with minor modifications. Treatment with glycine and acetic anhydride/triethanolamine was omitted, hybridization buffer was as in [32] and at 55 °C. Embryos were washed in 4 × NaCl/Cit, 50% formamide, 0.1% Tween 20 (3 × 30 min,

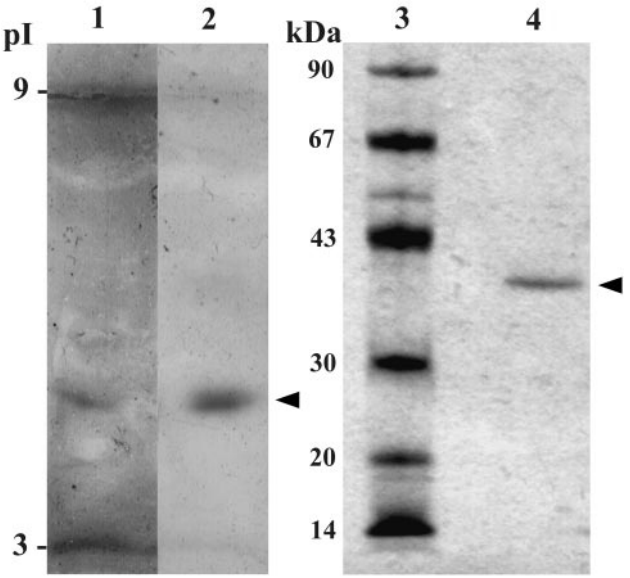


Fig. 1. Purified *B. lanceolatum* alcohol dehydrogenase class 3. Left: activity stainings with octanol (lane 1) and *S*-hydroxymethylglutathione (lane 2) after isoelectric focusing. Right: Coomassie brilliant blue staining after SDS/PAGE of protein markers (lane 3) and the purified enzyme solution (lane 4).

55 °C); 2 × NaCl/Cit, 50% formamide, 0.1% Tween 20 (3 × 30 min, 55 °C); 2 × NaCl/Cit, 0.1% Tween 20 (5 min, room temperature); followed by RNase treatment and 3 × 5 min NaCl/P_i, 0.1% Tween 20 washes. A 1354 bp clone containing the complete *Adh3* coding sequence (GenBank accession number AF154331) was used as a template to obtain the sense and antisense digoxygenin (DIG)-labeled probes, synthesized following the kit supplier's instructions (Boehringer Mannheim DIG RNA Labeling kit) and was used hydrolyzed without previous DNAse I treatment. After the labeled lancelets had been photographed as whole mounts under the microscope, they were dehydrated through an ethanol series and embedded in Spurr resin. Serial sections (3 μm each) were cut with glass knives and mounted in immersion oil.

RESULTS

Cloning and sequence analysis

A *Drosophila Adh*3 probe, constructed as described in Materials and methods, was used to screen the 6–20 h embryo *B. floridae* cDNA library. After re-screenings, three strongly hybridizing clones were selected for sequence analysis.

All three clones contained the full-length coding region but differed in the lengths of the 5′ ends. The longest clone (deposited in GenBank, accession number AF154331), with a total length of 1354 bp, 16 bp in the 5′-nontranslated region, 1131 bp in the coding region (corresponding to 377 amino-acid residues), and 187 bp to the start of the poly(A) tail, was used as a reference for further analysis. In the coding region, the three clones differed by 27 synonymous substitutions (data not shown) and six nonsynonymous substitutions corresponding to Val23 → Glu, Lys33 → Arg, Pro193 →Ser, Ala283 → Ser, Glu359 → Gln and Arg362 → Gln. Further heterogeneity at the 3′ nontranslated region was also observed. The 1354 bp *B. floridae* clone was used to screen a *B. lanceolatum* genomic library. Phages containing the *Adh3* gene were isolated and the corresponding coding region was deduced (deposited in GenBank, accession number AF156698-AF156708). The *B. lanceolatum Adh3* coding region spanned over 9.5 kb and was organized in 11 exons. Exon–intron boundaries were largely conserved when compared with those of mammalian *Adh* genes. *B. lanceolatum* and *B. floridae* ADH3 protein sequences are 91% identical. Versus the human and *Drosophila* ADH3, they are 75% and 76% identical, respectively. Interspecies (*B. lanceolatum* vs. *B. floridae*) and intraspecies (*B. floridae*) amino-acid variability does not include the substrate-binding or the coenzyme-binding positions [33,34]. Noticeably, these amino acids are highly conserved in other known ADH3 enzymes and so also in the Amphioxus variants.

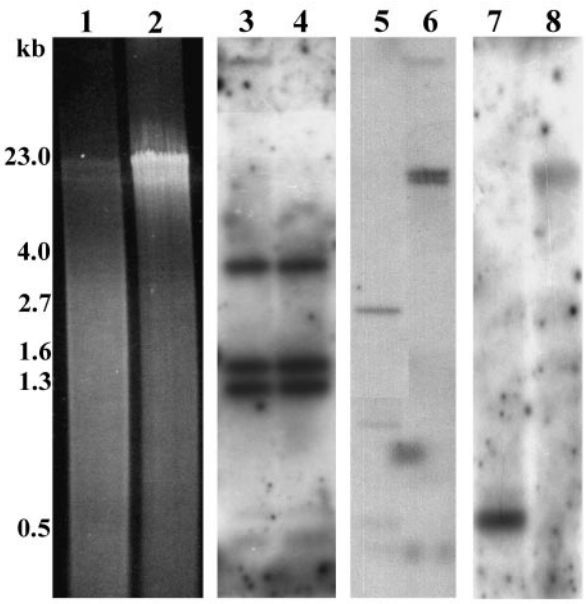


Fig. 2. Methylation status of *Adh3* gene in amphioxus. Lanes 1 and 2: fragment patterns obtained after digestion of total genomic DNA from a single *B. floridae* animal with *Msp*I and *Hpa*II, respectively. Lanes 3–8: Southern blots which were hybridized with *AmphiMLC-alk* gene (lanes 3 and 4) and *Adh3* gene: exons 2 and 3 (lanes 5 and 6) and exon 8 (lanes 7 and 8). Lanes 3, 5, 7, digestion with *Msp*I. Lanes 4, 6, 8, digestion with *Hpa*II.

Only one difference is found in these segments when the amphioxus form (Phe142) is compared to the human form (Tyr140).

Protein purification and enzymatic characterization

Class 3 alcohol dehydrogenase was purified to apparent homogeneity in batches (50 g) from *B. lanceolatum* whole animals by a three-step chromatographic procedure involving DEAE-Sepharose, Phenyl-Sepharose, and Mono Q FPLC, as given in Materials and methods. Final purification was 1400-fold in a yield of 25% and with a specific activity with S-hydroxymethylglutathione of 8.9 U·mg^{-1} (Table 1). The crude homogenate suggested the presence of only one alcohol dehydrogenase corresponding to the class 3 form, as only one band staining with octanol and formaldehyde-glutathione was resolved after iso-electric focusing (Fig. 1). No ethanol activity could be detected in the crude homogenate, directly or after attempts to induce ethanol-metabolizing enzymes by incubation of living animals in 1% ethanol for 15 h before homogenization. The purified class 3 enzyme exhibited, like the

Table 1. Purification of ADH3 from *B. lanceolatum*. Activities refer to *S*-hydroxymethylglutathione, measured at pH 8. Starting material, 50 g of whole animals.

Step	Total protein (mg)	Enzyme activity (U)	Specific activity (U·mg^{-1})	Purification (fold)	Yield (%)
Crude extract	379	2.39	0.00631	1	100
DEAE Sepharose	62	1.75	0.028	4	73
Phenyl Sepharose	1.4	1.31	0.94	150	55
Mono Q FPLC	0.067	0.59	8.9	1400	25

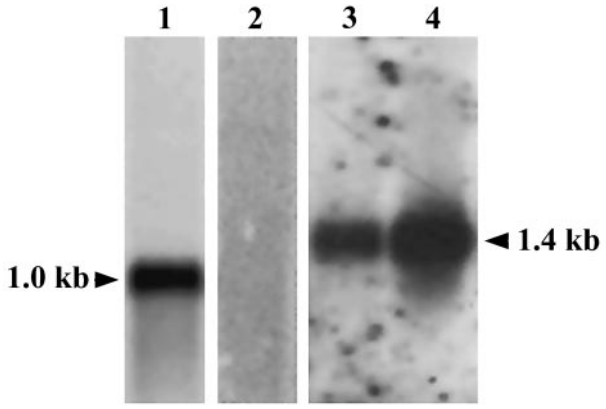


Fig. 3. Northern blot analysis of *B. lanceolatum*. 20 μg of total RNA (lane 2), 1 μg and 3 μg of poly(A)$^+$ RNA (lanes 3 and 4) were transferred and hybridized with the *Adh3* probe. Hybridization on the same membrane containing 20 μg total RNA with the *AmphiMLC-alk* probe (lane 1) was used as a control. Transcript size is indicated for *Adh3* (1.4 kb) and for *AmphiMLC-alk* (1.0 kb).

human form [35], only a weak ethanol-oxidizing activity and was unsaturable with ethanol concentrations up to at least 1 M. The combined results are compatible with a single class 3 alcohol dehydrogenase, and with an absence of class 1 forms in amphioxus. The final preparation showed a single band upon SDS/PAGE (Fig. 1). The molecular mass was estimated from SDS/PAGE to be 39.5 kDa. Mass spectrometry gave a value of 40 097 Da (calculated value for the acetylated protein chain without initiator methionine is 40 079 Da). The kinetic parameters (Table 2) were examined under conditions allowing comparisons with those obtained for other class 3 forms. The k_{cat}/K_m ratio with the basic formaldehyde dehydrogenase activity typical of class 3 is higher than for the human enzyme, and rather like that for the *Drosophila* form but overall values are related for all class 3 enzymes [11,13,34,36], corroborating conserved biochemical features of this enzyme in different animal lines.

Methylation status of the *Adh3* gene

To assess the methylation status of the *Adh3* gene the patterns of the restriction fragments generated by the methylation-sensitive enzyme, *Hpa*II, and the methylation-insensitive isoschizomer, *Msp*I, were compared. For the *Hpa*II digestion, most of the DNA was cleaved to a heterogeneous mixture of fragments resolved by electrophoresis, while an apparently undigested fraction remained at the top of the gel. This high molecular mass component was not devoid of *Hpa*II recognition sites (CCGG) as it was largely removed by digestion with *Msp*I (Fig. 2, lanes 1 and 2). Most likely this fraction was not cut by *Hpa*II because of CpG methylation. To evaluate the efficiency of the DNA digestion, Southern blots were hybridized with a probe of the *AmphiMLC-alk* gene, which belongs to the nonmethylated fraction [37]. Full digestion could be deduced from the identity of the *Hpa*II and *Msp*I restriction patterns (Fig. 2, lanes 3 and 4). The *Adh3* methylation status was mapped across the gene using as probes fragments of exons 2 and 3, and exon 8 of the gene (see Materials and methods). The probes hybridized to the fully methylated sequences of the *Hpa*II unrestricted high molecular weight fraction of amphioxus DNA (Fig. 2, lanes 6 and 8), which suggests that the 5′ and 3′ flanking regions were methylated.

Expression analysis of the amphioxus *Adh3* gene

Expression of the amphioxus *Adh3* gene was analyzed in the adult stage (Fig. 3) and in early embryonic development (Fig. 4). A single transcript of ≈ 1.4 kb was deduced from the isolated cDNAs from embryo libraries and from the northern analysis. Adult *Adh* transcripts could only be detected in RNA poly(A)$^+$ preparations, due to the relatively low abundance of the *Adh3* mRNA, in contrast to other highly transcribed genes, that is, *AmphiMLC-alk* (Fig. 3, see lanes 1 and 2 for comparison).

The earliest expression of the amphioxus *Adh3* gene was detected in the posterior endoderm of the neurula (12 h embryo) (Fig. 4B). Posterior elongation of the body continues at the end of neurulation, and embryos undergo the morphological changes that generate the larval body. At 24 h post-fertilization, the main expression domain was located beyond the branchial anlage and restricted to the posterior portion of the developing gut (Fig. 4C). At the later stages leading to the swimming larvae, the body grows at the posterior end, and the *Adh3* expressing cells are then situated far anterior of the caudal fin. When the mouth opens, at 36 h, the continuous signal of amphioxus *Adh3* seems to be interrupted by a small portion with less staining intensity, which corresponds to a small enlargement of the digestive lumen, splitting the

Table 2. Enzymatic properties of different forms of class 3 alcohol dehydrogenases. Values are from this work, *B.lan.* (*B. lanceolatum*), or from the literature [11,12,34,36]. HMGSH, S-hydroxymethylglutathione; 12-HAD, 12-hydroxydodecanoate. K_m in mM; k_{cat} in min^{-1}; k_{cat}/K_m in $min^{-1}·mM^{-1}$.

Substrate		*B.lan.*	Human	*Drosophila*
HMGSH	K_m	0.0044	0.004	0.006
(pH 8.0)	k_{cat}	680	200	960
	k_{cat}/K_m	150 000	50 000	160 000
12-HDA	K_m	0.090	0.060	0.04
(pH 10.0)	k_{cat}	680	170	840
	k_{cat}/K_m	7600	2800	21 000
Octanol	K_m	0.49	1.2	0.51
(pH 10.0)	k_{cat}	740	220	1300
	k_{cat}/K_m	1500	180	2500
NAD^+	K_m	0.011	0.009	0.13[a]
(pH 8.0)				

[a] This value was determined at pH 10 with octanol as substrate.

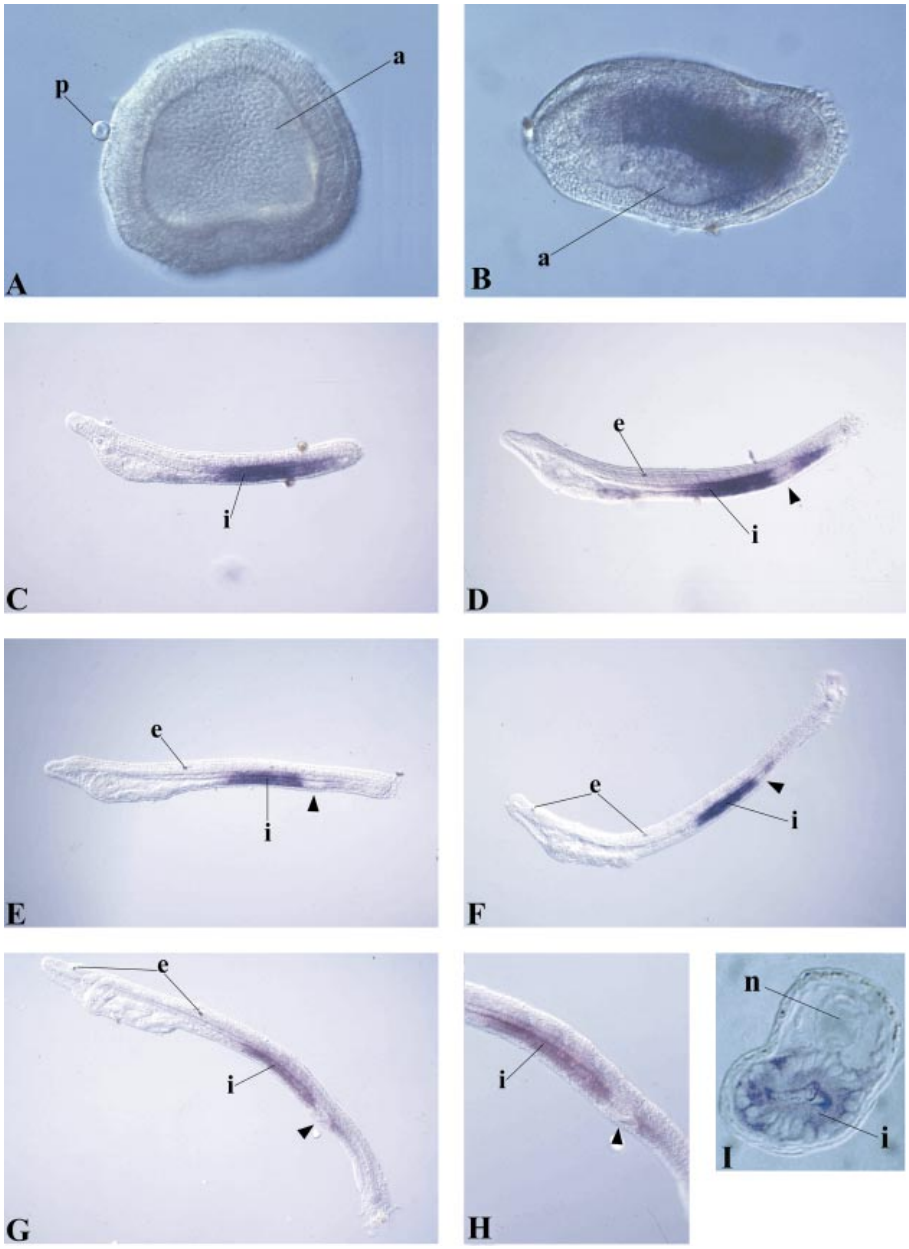


Fig. 4. Spatiotemporal expression of *Adh3* during *B. floridae* development visualized by whole-mount *in situ* hybridization. (A) 6 h amphioxus embryo at the gastrula stage. (B) Lateral view of 12 h neurula. (C–G) 24 h, 36 h, 48 h, 60 h and 6-day-old larvae. All larvae are positioned with anterior to the left and with dorsal to the top. (H) Higher magnification view of the 6-day larvae showing in detail the expression in the presumptive intestine interrupted by a nonstained region (arrowhead). (I) Transverse section through the stained region showing the signal confined to the gut. a, archenteron; e, eye spot; i, intestine; n, notochord, p, polar body.

expression pattern into two parts, where the anterior one is the largest (Fig. 4D). This expression-free zone could correspond to the region where the mid and hind guts meet, the ilio-colonic ring of adult animals. The same pattern of expression has been reported for the *AmphiNk2–2* putatively involved in regionalization of the digestive tract along the anteroposterior axis [38]. At 48 h and 60 h of development, the body has been completely elongated posteriorly and the expression pattern is maintained in the middle part of the body (Fig. 4E,F). No other sites of expression were detected in later larvae (Fig. 4G,H). Transversal sections at the level of the stained region showed the expression to be confined to the intestinal epithelium (Fig. 4I).

DISCUSSION

ADH3 relationships and properties

The knowledge of the structural and enzymatic properties of amphioxus ADH is important for the understanding of both this enzyme and the ethanol-active class 1 type of liver ADH. Previous data have suggested that many changes occurred at early vertebrate times, which involved a set of gene duplications [2] and an enzymogenesis of the ethanol activity [12]. These relationships can now be further evaluated. Observations on the functional and structural data of amphioxus ADH establish three overall relationships.

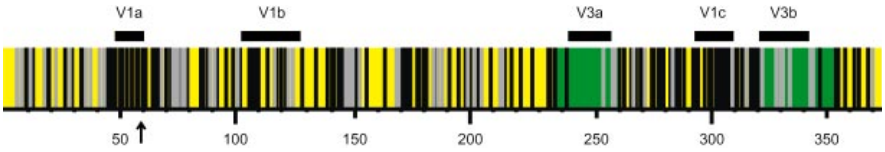


Fig. 5. Patterns of conservation and variability for ADH3 shown as a schematic representation of the residue conservations along the polypeptide chains. Strictly conserved residues are shown as black vertical lines, highly conserved residues (> 90%) as gray lines and the remaining positions as yellow. Green areas denote variable segments corresponding to those defined in [6] (black horizontal bars, V1a-c of class 1 and V3a,b of class 3). The figure is based on an alignment with CLUSTAL X [41] of ADH3 sequences from human (accession number P11766), *Uromastix* (P80467), cod *l* (P81601), hagfish (P80360), *B. floridae*, *B. lanceolatum*, *D. melanogaster* (P46415), octopus (P81431), *C. elegans* (Q17335), pea (P80572), *S. cerevisiae* (P32771) and *E. coli* (P25437). Positions are numbered according to the human enzyme. The arrow indicates a residue strictly conserved when the amphioxus sequences are excluded from the alignment.

One is the functional constancy of class 3 activities (Table 2). The amphioxus enzyme links the previously characterized mammalian forms with the still earlier yeast, insect and plant forms, and shows that, within a factor of about three, enzymatic values are constant from insects, via amphioxus to the human enzyme (Table 2). Significantly, the amphioxus values are close to those of the *Drosophila* enzyme, in which a Tyr51→Phe substitution was suggested to contribute to a weakened coenzyme binding and therefore an increased k_{cat} value [6]. In relation to the human enzyme, the amphioxus structure shows conservation not only of Tyr51 but also of the other residues ascribed coenzyme-interacting functions as deduced from crystallographic studies [33,34,39]. These apparent contradictions suggest that the differences in k_{cat} values can be derived from remote changes at 'nonfunctional' positions as concluded also for the cod isozyme differences [40].

The second is the fact that the amphioxus enzyme is a class 3 form. No evidence for more than one gene or for other ADH forms were detected. The Southern analysis (Fig. 2) is compatible with a single-copy gene and so is the finding of only a single octanol-staining band in isoelectric focusing gels of crude *B. lanceolatum* homogenates. Although three *B. floridae* cDNAs were isolated, most of their differences (27 out of 33) are synonymous substitutions, and of the non-synonymous replacements, four of six affect variable positions shared with *B. lanceolatum*. These observations are compatible with allelic variability as the explanation for the *B. floridae* differences. No ethanol dehydrogenase activity could be detected using a standard spectrophotometric assay, not even after attempts at inducing such an enzyme with ethanol in the living animal. Hence, all data are compatible with a parent nature of class 3 in relation to the later vertebrate ADH radiation. It is also well illustrated in other multigene families that duplication events in the vertebrate lineage have originated after the divergence of lancelets and that therefore only a single gene member may be present in the cephalochordate line [21]. In agreement with this, we now only find one ADH3 in amphioxus; no evidence of a class 1 form has been gathered.

The third relationship is the structural constancy of class 3 ADH. Inclusion of the amphioxus enzyme in comparisons of all class 3 forms known, from prokaryotes, fungi, plants via amphioxus to vertebrates, reveals that no less than one third of the positions are invariant, more than 45% exhibit only limited variability. This is within the range of similarity for proteins with a conserved function, compatible with the defined metabolic role for ADH3, in formaldehyde detoxification.

The conservation pattern is not uniform along the protein chain. Previous comparisons have suggested that the ADH3 variability is concentrated to two segments at the surface of the molecule, well apart from the active site and subunit interactions [6]. Addition of amphioxus to the comparisons (using the program CLUSTAL X [41]) now shows the pattern with two variable ADH3 segments (indicated as V3a and V3b and marked green in Fig. 5) to be uniform among all chordates. In contrast, the variable segments of the class 1 enzyme (V1a–c), which are near the substrate-binding pocket and participate in subunit interactions, are among the most conserved parts of the class 3 molecule, clearly showing the differences in pattern among the classes of ADH, and linking those differences to functional differences. The addition of the amphioxus sequences introduces a single substitution at an otherwise conserved position, Pro58, where a serine is found in both of the two lancelet ADHs. Although Pro58 is adjacent to Asp57, which is crucial for *S*-hydroxymethylglutathione binding [42], this proline substitution would not prevent substrate recognition, as deduced from the human class 3 tertiary structure [39] and supported by our enzymatic data. Thus, the *B. lanceolatum* enzyme is highly active with *S*-hydroxymethylglutathione, 12-hydroxydodecanoate and octanol.

Methylation status of the *Adh3* gene

A typically invertebrate methylation pattern has been reported for amphioxus [37]. The major genomic fraction comprises long domains of nonmethylated DNA interrupted by a minor portion made up of long stretches of heavily methylated DNA. This pattern differs from that of vertebrates, vastly methylated in most cell types. A classic view of genome organization links DNA methylation to gene regulation, whereas genes with a housekeeping role would not be methylated. In accordance with this view the amphioxus *Adh3* gene is methylated and shows a tissue-specific expression. However, a housekeeping role for ADH3 could not be ruled out only on the basis of the methylation status, as several housekeeping genes in amphioxus are methylated [37].

Unexpected *Adh3* expression pattern in amphioxus, suggestive of a unique ADH role in differentiation

Northern blot analysis revealed a single *B. lanceolatum* ADH3 transcript. This is similar to the situation reported for *Arabidopsis*, *Drosophila*, and mouse but dissimilar from the situation reported in the human system where two transcripts have been described [8]. Our *in situ* hybridization analysis revealed that *Adh3* expression was restricted to particular cell types, mainly in the presumptive gut, in embryos as well as in late free-swimming and feeding larvae, which already show a

basic body plan essentially identical to that of the adult animal. This result is in clear contrast with the ubiquitous class 3 expression reported in vertebrates. This difference is still more surprising if the previously proposed housekeeping role of ADH3 is considered.

Two possible scenarios would explain the differential expression patterns observed. One is that expression in the common ancestor of vertebrates and cephalochordates could have been restricted to particular cell types, as it now appears to be in amphioxus, and extension to all tissues could then have been acquired later, possibly linked to the subsequent gene duplication events and to the acquisition of new functions in the ADH family. Alternatively, the other scenario is that the original expression would have been widespread but that lancelet expression pattern of ADH3 later gained some unique traits, although lancelets retained many of the primitive chordate developmental features. In either case, the present expression analysis shows a unique pattern and suggests an interesting regulatory role for ADH3 in embryonic development. The proposed role of ADH3 as a housekeeping gene may need to be re-evaluated in view of the present data. The viability and fertility described for *Drosophila* homozygotes with the null allele Fdh^{nNCl} [43,44], and the null mutant mice generated by gene targeting [45], further emphasize this point.

ACKNOWLEDGEMENTS

We are indebted to L. Holland and N.D. Holland for helpful discussions in the *in situ* analysis. We thank J. M. Lawrence for laboratory facilities in Tampa. This work was supported by grants from DGICYT (Ministerio de Educación y Cultura, Spain, PB96-0220 and UE98-0014), the Swedish Medical Research Council (K2000-03X-03532-29 A), EU contract BIO4 CT97–2123, the Foundations of Lars Hiertas Minne, Fredrik och Ingrid Thuring, and the Swedish Medical Society, an FPI fellowship from the CIRIT (Generalitat de Catalunya, 1997FI00007) and a STINT postdoctoral fellowship. We are indebted to the Servei Cientifico-Tècnics (UB) for DNA sequence analysis and preparation of embryo sections.

REFERENCES

1. Bonnichsen, R.K. & Wassén, A.M. (1948) Crystalline alcohol dehydrogenase from horse liver. *Arch. Biochem. Biophys.* **18**, 361–363.
2. Jörnvall, H. (1994) The alcohol dehydrogenase system. In *Toward a Molecular Basis of Alcohol Use and Abuse* (Jansson, B., Jörnvall, H., Rydberg, U., Terenius, L. & Vallee, B.L., eds), pp. 221–229. Birkhäuser Verlag, Basel, Switzerland.
3. Jörnvall, H., Höög, J.-O. & Persson, B. (1999) SDR and MDR: completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS Lett.* **445**, 261–264.
4. Jörnvall, H. & Höög, J.-O. (1995) Nomenclature of alcohol dehydrogenases. *Alcohol Alcoholism* **30**, 153–161.
5. Duester, G., Farrés, J., Felder, V.C., Höög, J.O., Parés, X., Plapp, B., Yin, S.-Y. & Jörnvall, H. (1999) Recommended nomenclature for the vertebrate alcohol dehydrogenase gene family. *Biochem. Pharmacol.* **58**, 389–395.
6. Danielsson, O., Atrian, S., Luque, T., Hjelmqvist, L., Gonzàlez-Duarte, R. & Jörnvall, H. (1994) Fundamental molecular differences between alcohol dehydrogenase classes. *Proc. Natl Acad. Sci. USA* **91**, 4980–4984.
7. Koivusalo, M., Baumann, M. & Uotila, L. (1989) Evidence for the identity of glutathione-dependent formaldehyde dehydrogenase and class III alcohol dehydrogenase. *FEBS Lett.* **257**, 105–109.
8. Estonius, M., Svensson, S. & Höög, J.-O. (1996) Alcohol dehydrogenase in human tissues: localisation of transcripts coding for five classes of the enzyme. *FEBS Lett.* **397**, 338–342.
9. Haselbeck, R.J. & Duester, G. (1997) Regional restriction of alcohol/retinol dehydrogenases along the mouse gastrointestinal epithelium. *Alcohol. Clin. Exp. Res.* **21**, 1484–1490.
10. Ang, H.L., Deltour, L., Hayamizu, T.F., Zgombic-Knight, M., Duester, G. (1996) Retinoic acid synthesis in mouse embryos during gastrulation and craniofacial development linked to class IV alcohol dehydrogenase gene expression. *J. Biol. Chem.* **271**, 9526–9534.
11. Uotila, L. & Koivusalo, M. (1989) Glutathione-dependent oxidoreductases: formaldehyde dehydrogenase. In *Coenzymes Vol. 3* (Dolphin, D., Poulson, R. & Avramovic, O, eds), pp. 517–551, John Wiley and Sons, New York.
12. Danielsson, O. & Jörnvall, H. (1992) 'Enzymogenesis': classical liver alcohol dehydrogenase origin from the glutathione-dependent formaldehyde dehydrogenase line. *Proc. Natl Acad. Sci. USA* **89**, 9247–9251.
13. Danielsson, O., Shafqat, J., Estonius, M. & Jörnvall, H. (1994) Alcohol dehydrogenase class III contrasted to class I. Characterization of the cyclostome enzyme, the existence of multiple forms as for the human enzyme, and distant cross-species hybridization. *Eur. J. Biochem.* **225**, 1081–1088.
14. Danielsson, O., Eklund, H. & Jörnvall, H. (1992) The major piscine liver alcohol dehydrogenase has class-mixed properties in relation to mammalian alcohol dehydrogenases of classes I and III. *Biochemistry* **31**, 3751–3759.
15. Cederlund, E., Peralba, J.M., Parés, X. & Jörnvall, H. (1991) Amphibian alcohol dehydrogenase, the major frog liver enzyme. Relationships to other forms and assessment of an early gene duplication separating vertebrate class I and class III alcohol dehydrogenases. *Biochemistry* **30**, 2811–2816.
16. Schaeffer, B. (1987) Deuterostome monophyly and phylogeny. *Evol. Biol.* **21**, 179–235.
17. Field, K.G., Olsen, G.J., Lane, D.J., Giovanni, S.J., Ghiselin, M.T., Raft, E.C., Pace, N.R. & Raff, R.A. (1988) Molecular phylogeny of the animal kingdom. *Science* **239**, 748–753.
18. Wada, H. & Satoh, N. (1994) Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. *Proc. Natl Acad. Sci. USA* **91**, 1801–1804.
19. Spruyt, N., Delarbre, C., Gachelin, G. & Laudet, V. (1998) Complete sequence of the amphioxus (*Branchiostoma lanceolatum*) mitochondrial genome: relations to vertebrates. *Nucleic Acids Res.* **26**, 3279–3285.
20. Garcia-Fernàndez, J. & Holland, P.W.H. (1994) Archetypal organization of the amphioxus Hox gene cluster. *Nature* **370**, 563–566.
21. Holland, P.W.H. (1996) Molecular biology of lancelets: insights into development and evolution. *Israel J. Zool.* **42**, 247–272.
22. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. & Rutter, W.J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294–5299.
23. Garcia-Fernàndez, J., Baguñà, J. & Saló, E. (1993) Genomic organization and expression of the planarian homeobox genes *Dth-1* and *Dth-2*. *Development* **118**, 241–253.
24. Langeland, J.A., Tomsa, J.M., Jackman, W.R. Jr & Kimmel, C.B. (1998) An amphioxus snail gene: expression in paraxial mesoderm and neural plate suggests a conserved role in patterning the chordate embryo. *Dev. Genes Evol.* **208**, 569–577.
25. Luque, T., Atrian, S., Danielsson, O., Jörnvall, H. & Gonzàlez-Duarte, R. (1994) Structure of the *Drosophila melanogaster* glutathione-dependent formaldehyde dehydrogenase/octanol dehydrogenase gene (class III alcohol dehydrogenase). Evolutionary pathway of the alcohol dehydrogenase genes. *Eur. J. Biochem.* **225**, 985–993.
26. Church, G.M. & Gilbert, W. (1984) Genomic sequencing. *Proc. Natl Acad. Sci. USA* **81**, 1991–1995.
27. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.
28. Lutz, R.A., Bull., C. & Rodbard, D. (1986) Computer analysis of enzyme-substrate-inhibitor kinetic data with automatic model selection using IBM-PC compatible microcomputers. *Enzyme* **36**, 197–206.
29. Holland, L.Z., Pace, D.A., Blink, M.L., Kene, M. & Holland, N.D.

(1995) Sequence and expression of amphioxus alkali myosin light chain (*AmphiMLC-alk*) throughout development: implications for vertebrate myogenesis. *Dev. Biol.* **171**, 665–676.
30. Chomczynski, P. & Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156–159.
31. Holland, P.W.H., Holland, L.Z., Williams, N.A. & Holland, N.D. (1992) An amphioxus homeobox gene: sequence conservation, spatial expression during development and insights into vertebrate evolution. *Development* **116**, 653–661.
32. Agata, K., Soejima, Y., Kato, K., Kobayashi, C., Umesono, Y. & Watanabe, K. (1998) Structure of the Planarian central nervous system (CNS) revealed by neuronal cell markers. *Zool. Sci.* **15**, 433–440.
33. Eklund, H., Samama, J.P. & Jones, T.A. (1984) Crystallographic investigations of nicotinamide adenine dinucleotide binding to horse liver alcohol dehydrogenase. *Biochemistry* **23**, 5982–5996.
34. Eklund, H., Müller-Wille, P., Horjales, E., Furter, O., Holmquist, B., Vallee, B.L., Höög, J.-O., Kaiser, R. & Jörnvall, H. (1990) Comparison of three classes of human liver alcohol dehydrogenase. Emphasis on different substrate binding pockets. *Eur. J. Biochem.* **193**, 303–310.
35. Wagner, F.W., Parés, X., Holmquist, B. & Vallee, B.L. (1984) Physical and enzymatic properties of a class III isozyme of human liver alcohol dehydrogenase: χ-ADH. *Biochemistry* **23**, 2193–2199.
36. Holmquist, B. & Vallee, B.L. (1991) Human liver class III alcohol and glutathione dependent formaldehyde dehydrogenase are the same enzyme. *Biochem. Biophys. Res. Commun.* **178**, 1371–1377.
37. Tweedie, S., Charlton, J., Clark, V. & Bird, A. (1997) Methylation of genomes and genes at the invertebrate–vertebrate boundary. *Mol. Cell. Biol.* **17**, 1469–1475.
38. Holland, L.Z., Venkatesh, T.V., Gorlin, A., Bodmer, R. & Holland, N.D. (1998) Characterization and developmental expression of *Amphi Nk2–2*, an NK2 class homeobox gene from amphioxus. *Dev. Genes Evol.* **208**, 100–105.
39. Yang, Z.N., Bosron, W.F. & Hurley, T.D. (1997) Structure of human χχ alcohol dehydrogenase: a glutathione-dependent formaldehyde dehydrogenase. *J. Mol. Biol.* **265**, 330–343.
40. Danielsson, O., Shafqat, J., Estonius, M., El-Ahmad, M. & Jörnvall, H. (1996) Isozyme multiplicity with anomalous dimer patterns in a class III alcohol dehydrogenase. Effects on the activity and quaternary structure of residue exchanges at 'nonfunctional' sites in a native protein. *Biochemistry* **35**, 14561–14568.
41. Thompson, J.D., Gibson, T.J., Plewniha, F., Jeanmougin, F. & Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882.
42. Estonius, M., Höög, J.-O., Danielsson, O., Jörnvall, H. (1994) Residues specific for class III alcohol dehydrogenase. Site-directed mutagenesis of the human enzyme. *Biochemistry* **33**, 15080–15085.
43. Voelker, R.A., Langley, C.H., Leigh Brown, A.J., Ohnishi, S., Dickson, B., Montgomery, E. & Smith, S.C. (1980) Enzyme null alleles in natural populations of *Drosophila melanogaster*: frequencies in a North Carolina population. *Proc. Natl Acad. Sci. USA* **77**, 1091–1095.
44. Burkhart, B.D., Montgomery, E., Langley, C.H. & Voelker, R.A. (1984) Characterization of allozyme null and low activity alleles from two natural populations of *Drosophila melanogaster*. *Genetics* **107**, 295–306.
45. Deltour, L., Foglio, M.H. & Duester, G. (1999) Metabolic deficiencies in alcohol dehydrogenase *Adh1*, *Adh3*, and *Adh4* null mutant mice. Overlapping roles of *Adh1* and *Adh4* in ethanol clearance and metabolism of retinol to retinoic acid. *J. Biol. Chem.* **274**, 16796–16801.

CAPÍTOL II

En aquest capítol, 1) es caracteritza l'estructura gènica de l'***Adh3*** en dues espècies de cefalocordats (***Branchiostoma lanceolatum*** i ***Branchiostoma floridae***) i en una espècie d'urocordat (***Ciona intestinalis***). 2) Es determinen les taxes d'evolució de l'ADH3 i l'ADH1 en el regne animal. 3) S'estima que l'expansió de la família ADH es va produir fa 500 milions d'anys per duplicacions en tàndem probablement relacionades amb l'explosió d'isoformes proposada després de l'aparició dels àgnats, i per tant no lligada a les dues rondes de duplicacions a gran escala que tingueren lloc en l'evolució primerenca dels vertebrats. 4) La integració de totes aquestes dades permet correlacionar els destins funcionals de les diferents classes d'ADH amb l'expansió de la família gènica. 5) Finalment, basant-se en la constància evolutiva de l'ADH3 s'estima l'origen dels vertebrats fa uns 700 MA i el temps de separació entre les dues espècies d'amfiox en quasi 200 MA.

J Mol Evol (2002) 54:81–89
DOI: 10.1007/s00239-001-0020-2

JOURNAL OF MOLECULAR EVOLUTION

Ascidian and Amphioxus *Adh* Genes Correlate Functional and Molecular Features of the ADH Family Expansion During Vertebrate Evolution

Cristian Cañestro,[1] Ricard Albalat,[1] Lars Hjelmqvist,[1,2] Laura Godoy,[1] Hans Jörnvall,[2] Roser Gonzàlez-Duarte[1]

[1] Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, E-08028 Barcelona, Spain
[2] Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-17177 Stockholm, Sweden

Received: 10 April 2001 / Accepted: 23 May 2001

Abstract. The alcohol dehydrogenase (ADH) family has evolved into at least eight ADH classes during vertebrate evolution. We have characterized three prevertebrate forms of the parent enzyme of this family, including one from an urochordate (*Ciona intestinalis*) and two from cephalochordates (*Branchiostoma floridae* and *Branchiostoma lanceolatum*). An evolutionary analysis of the family was performed gathering data from protein and gene structures, exon–intron distribution, and functional features through chordate lines. Our data strongly support that the ADH family expansion occurred 500 million years ago, after the cephalochordate/vertebrate split, probably in the gnathostome subphylum line of the vertebrates. Evolutionary rates differ between the ancestral, ADH3 (glutathione-dependent formaldehyde dehydrogenase), and the emerging forms, including the classical alcohol dehydrogenase, ADH1, which has an evolutionary rate 3.6-fold that of the ADH3 form. Phylogenetic analysis and chromosomal mapping of the vertebrate *Adh* gene cluster suggest that family expansion took place by tandem duplications, probably concurrent with the extensive isoform burst observed before the fish/tetrapode split, rather than through the large-scale genome duplications also postulated in early vertebrate evolution. The absence of multifunctionality in lower chordate ADHs and the structures compared argue in favor of the acquisition of new functions in vertebrate ADH classes. Finally, comparison between *B. floridae* and *B. lanceolatum Adh*s provides the first estimate for a cephalochordate speciation, 190 million years ago, probably concomitant with the beginning of the drifting of major land masses from the Pangea.

Key words: Alcohol dehydrogenase — Evolutionary rate — Gene duplication — Vertebrate genome evolution

Introduction

Major increases in the number of genes have delineated key steps during evolution (for review see Lundin 1999). It has been suggested that extensive gene duplications occurred by genome tetraploidization after the cephalochordate/vertebrate split and that this provided the flexibility to build the more complex vertebrate body plan (Garcia-Fernàndez and Holland 1994). Moreover, gene tandem duplications (isoform burst) have been postulated to have occurred at about the same time as or immediately before the cyclostome/gnathostome split (Sharman and Holland 1996; Suga et al. 1999). The expansion of single-copy genes in vertebrates has been dated in a variety of gene families (for reviews see Lundin 1993; Skrabanek and Wolfe 1998). Irrespective of the mechanism, duplications are responsible for the extant gene families and isoforms from fish to mammals (Ohno 1970; Lundin 1993) and are believed to be related to the gains in complexity which have allowed the acquisition of new functions through enzymogenesis

Correspondence to: Hans Jörnvall; *email:* Hans.Jornvall@mbb.ki.se

(Danielsson and Jörnvall 1992) and changes in gene regulation (for a review see Shimeld 1999).

Comparisons with extensive sets of sequences from several gene families are being made to elucidate how vertebrate genomes evolved (Wang et al. 1998; Kumar and Hedges 1998; Hughes 1999; Wang and Gu 2000). However, combination of the sequence data with the exon–intron structure, chromosomal location, expression domains, and other functional features would provide a more comprehensive scenario to understand gene and genome evolution.

The alcohol dehydrogenase (ADH) family belongs to a complex protein system with extensive multiplicity within the medium-chain dehydrogenase/reductase (MDR) superfamily. Different classes and isozymes of ADHs are found in the vertebrate lineage. In mammals, at least six ADH classes are known (Jörnvall and Höög 1995; Duester et al. 1999), of which five have been characterized in human. Of these, classes 1 and 3 have been most investigated in different animal clades. The class 1 enzyme, the "classical" ethanol-active form, has evolved rapidly and exhibits a considerable variability among species (Danielsson et al. 1994a). ADH class 3 [glutathione-dependent formaldehyde dehydrogenase (Koivusalo et al. 1989)] is present in prokaryotes and in all eukaryotes and appears to be the ancestral form from which the other vertebrate ADH classes emerged (Danielsson and Jörnvall 1992; Danielsson et al. 1994a; Cañestro et al. 2000) through gene duplications during early vertebrate radiation. Supporting evidence comes from the apparent absence of an ethanol-active class 1 form in lines originating before the bony fish (Danielsson et al. 1994b) and the presence of a class 1/3-mixed form in bony fish (Danielsson et al. 1992).

To study ADH gene family expansion, we have compared the vertebrate gene classes with those found in two prevertebrates, ascidians and amphioxus. Ascidians, subphylum Urochordata, are considered the simplest chordates and a critical evolutionary link between invertebrates and vertebrates. They show compact genomes which are comparable in size to those of *Drosophila* and *Caenorhabditis elegans.* On the other hand, amphioxus, a member of the subphylum Cephalochordata, is thought to be the closest living relative to vertebrates (Garcia-Fernàndez and Holland 1994; Holland 1996) and therefore to be an ideal outgroup for evolutionary and phylogenetic studies of vertebrate gene families. We have characterized the *Adh3* genomic structure of the urochordate *Ciona intestinalis,* and the cephalochordates *Branchiostoma floridae* and *Branchiostoma lanceolatum.* Analyses of gene structure, amino acid substitutions, biochemical data, and expression domains strongly suggest that class divergence took place after the vertebrate/cephalochordate split. We estimate the first gene duplication, leading to new forms of the enzyme, to have occurred about 500 million years ago (MYA), in the early vertebrate evolution before the fish/tetrapode split. Changes in the evolutionary rate of the ancestral and the newly generated forms can be correlated with the functional evolution of vertebrate ADHs, where the ancestral ADH3 has remained biochemically constant from insects, via amphioxus to the human enzyme (Danielsson et al. 1994a; Cañestro et al. 2000), and the newly generated forms have acquired novel activities. Finally, the chromosomal location in human and mice supports extensive isoform generation by tandem duplications rather than by large-scale genome duplications for the expansion of the ADH family.

Materials and Methods

Library Screenings and Southern Blot Analysis

A *B. lanceolatum* genomic library (Cañestro et al. 2000) and a *B. floridae* genomic library (kindly provided by J. Garcia-Fernàndez) were probed with [α-^{32}P]dCTP *B. floridae Adh3* cDNA labeled by random-hexamer priming. Total RNA from *C. intestinalis* was purified by the guanidinium thiocyanate method (Chomczynski and Sacchi 1987). Synthesis of cDNA was performed with MLV/RT (Promega) and random-hexamer priming. Degenerate oligonucleotides (GA(AG)GC(ATCG)TG(TC)CA(TC)AA(AG)GG(ATCG)TGG and CC(AG)AA(ATCG)GC(ATCG)GT(ATCG)CC(CT)TTCCA) were designed to amplify a *C. intestinalis* cDNA fragment of the *Adh3* gene, and this fragment was used to screen a genomic library (kindly provided by M. Levine) and a cDNA library (kindly provided by R. Di Lauro).

Hybridizations for library screenings were carried out in phosphate–sodium dodecyl sulfate (SDS) solution (Church and Gilbert 1984) at 65°C overnight. Washes were done at 65°C for 2 × 15 min in 2× standard saline citrate (SSC), 0.1% SDS, and 1 × 15 min in 1× SSC, 0.1% SDS. DNA fragments from positive recombinant phages were isolated, subcloned into a pUC18 vector, characterized by restriction mapping, and sequenced in an ABI-Prism 377 DNA sequencer from PE Biosystems.

Sequence Data

The evolutionary distance, k_{aa}, among different groups of organisms was calculated using the sequence data from the SWISS-PROT, TrEMBL, EMBL and GenBank data banks. Accession numbers of sequences are as follow.

ADH Class 3. Primates: *Homo sapiens* (P11766). Rodents: *Mus musculus* (P28474); *Rattus norvegicus* (P12711); *Oryctolagus cuniculus,* rabbit (O19053). Mammals: Primates + rodents + *Equus caballus,* horse (P19854). Reptiles: *Uromastyx hardwickii,* lizard (P80467). Tetrapods: Mammals + reptiles. Fish: *Sparus aurata,* gilthead sea bream (P79896); *Gadus morhua callarias,* cod (P81601). Gnathostomes: Tetrapods + fish. Hagfish: *Myxine glutinosa* (P80360). Vertebrates: hagfish + gnathostomes. Cephalochordates: *Branchiostoma lanceolatum* (AF156698–AF156708); *Branchiostoma floridae* (AF154331, AF344170, AF344171). Urochordates: *Ciona intestinalis* (AF344173). Deuterostomes: Vertebrates + cephalochordates + urochordates.

ADH Class 1. Primates: *Homo sapiens* (A, P07327; B, P00325; C, P00326); *Papio hamadryas,* baboon (P14139); *Macaca mulatta,* rhesus

macaque (P28459). Rodents: *Mus musculus* (P00329); *Rattus norvegicus* (P06757); *Peromyscus maniculatus,* deer mouse (P41680); *Geomys bursarius* (Q64413); *Geomys knoxjonesi* (Q64415); *Oryctolagus cuniculus,* rabbit (Q03505). Mammals: Primates + rodents + *Equus caballus,* horse (E, P00327; S, P00328). Birds: *Gallus gallus,* chicken (P23991); *Coturnix coturnix japonica,* quail (P19631); *Apteryx australis,* kiwi (P49645); *Struthio camelus,* ostrich (P80338). Reptiles: *Uromastyx hardwickii,* Indian spiny-tailed lizard (A, P25405; B, P25406); *Alligator mississippiensis* (P80222); *Naja naja,* cobra (P80512). Amphibians: *Rana perezi,* frog (P22797). Tetrapods: Mammals + birds + reptiles + amphibians. Fish: *Gadus morhua callarias,* cod (P26325).

ADH Class 2. Homo sapiens (P08319), *Mus musculus* (Q9QYY9).

ADH Class 4. Homo sapiens (P40394); *Mus musculus* (Q64437).

Alignments, Evolutionary Distances, and Relative Rate Tests

Alignments of amino acid sequences, from Ile at position 6 to Arg at position 369 of human ADH3, were generated with the ClustalX program (Thompson et al. 1997) using default gap penalties, with adjustments made by eye.

The k_{aa} value of amino acid substitutions per site (evolutionary distance) was calculated with the MEGA (Molecular Evolutionary Genetic Analysis) package (Kumar et al. 1993) for the amino acid γ distance, with an α parameter obtained with PUZZLE version 4.0.1 (Strimmer and Haeseler 1996). This same quartet puzzling procedure was used for the maximum likelihood calculations, following the JTT model for amino acid substitutions (Jones et al. 1992), estimating the amino acid frequencies from the data set and taking into account the among-site rate heterogeneity. When the evolutionary distances between groups were based on more than one species, the average value of k_{aa} was used.

Rate differences among lineages were analyzed by the "two-cluster" and "branch-length" tests provided in the Lintre package (Takezaki et al. 1995). Both tests were applied to a phylogenetic tree according to the widest accepted phylogeny (Kuma and Miyata 1994; Naylor and Brown 1997) and the γ-corrected distribution for multiple replacements.

Results

Characterization of C. intestinalis, B. floridae, *and* B. lanceolatum Adh3 *Genes*

After library screening, restriction map analysis of *B. lanceolatum, B. floridae,* and *C. intestinalis* showed that in each case all the phages overlapped the same *Adh3* genomic region. A *B. lanceolatum* 13-kb phage containing the complete *BlAdh3* coding region was analyzed further, and intron/exon boundaries were sequenced (AF156698–AF156708). The same analysis was performed with a 15-kb phage of *B. floridae* (AF266713–AF266719) and a 6-kb phage of *C. intestinalis* (AF344173), both encompassing the entire *BfAdh3* and *CiAdh3* coding regions, respectively.

Exon–Intron Structure of Ascidian and Amphioxus ADH3

The exon–intron organization of *CiAdh3, BlAdh3,* and *BfAdh3* was established by comparing the corresponding genomic clones with the *C. intestinalis* (AF344172) and *B. floridae* (AF154331) cDNA (Fig. 1A). The *CiAdh3, BlAdh3,* and *BfAdh3* genes consisted of 11, 11, and 10 exons, respectively. Intron sizes differed greatly between the two subphyla; the average amphioxus intron length was 0.9 ± 0.6 kb, whereas it was shorter for the ascidian, 0.2 ± 0.1 kb. In comparisons of the *Adh3* exon/intron architecture of protostomes and deuterostomes, prevertebrates showed the highest number of introns (10 and 9 of the 12 introns identified; Fig. 1B). Introns 1, 2, 7, and 9 were conserved in at least one protostome and all deuterostome species. Intron 2 showed an identical 12-nucleotide (nt) slippage in the fruit fly (Luque et al. 1994) and, in ascidians, the latter with an additional 10-nt slippage in intron 4. *Drosophila subobscura* is the only protostome where intron 7 has been identified (Abad 2000). What had been considered a unique intron 6 in *C. elegans* could be regarded as a 29-nt slippage of intron 7. Introns 3, 4, 11, and 12 seemed to be restricted to the deuterostome lineage. *B. floridae* lacked intron 11, although it was present in *C. intestinalis, B. lanceolatum,* and all *Adh* vertebrate genes. Moreover, introns 5 and 8 appeared to be restricted to prevertebrates, as was intron 10 to the amphioxus lineage. Interestingly, vertebrate ADH classes 1, 2, 3, and 4 shared the same exon–intron organization (introns 1, 2, 3, 4, 7, 9, 11, and 12) (Fig. 1B). This, together with the evolutionary and functional data, suggests that class divergence took place after the vertebrate/cephalochordate split.

Evolutionary Distances and Amino Acid Substitution Rates of Chordate ADH3 and ADH1

The evolutionary distance k_{aa} was calculated for chordate ADH3 and ADH1 with a γ-corrected distribution. As this method takes into account the nonuniform among-site rate variation pattern of each enzyme, it is considered the most pertinent. The low values for the α parameter, 0.5 and 0.6 for class 3 and class 1, respectively, were in agreement with the structural pattern of ADHs, made of highly variable regions embedded in conserved segments along the polypeptide chains (Danielsson et al. 1994a). The approach to saturation of amino acid substitutions was analyzed for ADH3 by plotting the frequency of observed differences *vs* estimated substitutions (Philippe et al. 1994) (Fig. 2A). Although the curve starts to level off, it does not reach a plateau, meaning that the ADH3 sequences were informative and not mutationally fully saturated. Moreover, we compared the γ-distance *vs* ML estimates (Roger et al. 1999) and very similar values were obtained (Fig. 2B).

To evaluate the constancy of the ADH class 3 amino acid substitution rate in vertebrates and cephalochordates, the "two-cluster" and the "branch-length" statistical tests were performed, using the ascidian ADH sequence as the outgroup. The χ^2 values that combined all

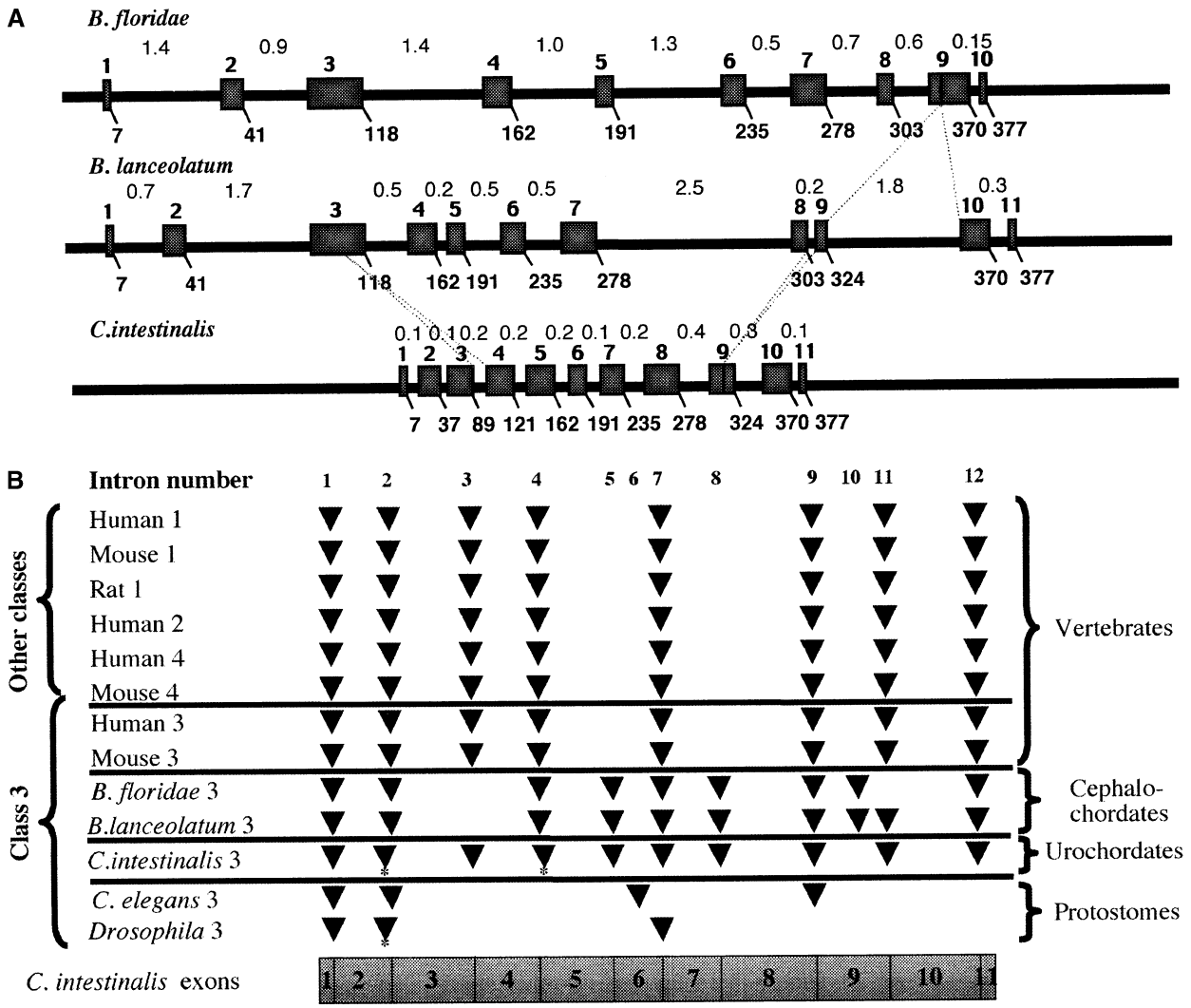


Fig. 1. Exon–intron architecture. **A** Exon (*boxed*) and intron (*non-boxed*) distribution in *B. floridae* (top), *B. lanceolatum* (middle), and *C. intestinalis* (bottom) class 3 *Adh.* The exon number is shown over the boxes and the figures below indicate the amino acid position before the intron. Intron sizes (kb) are shown above the intron line. **B** Schematic comparison of the exon–intron architecture for ADH classes in animal species. *Arrowheads* indicate the intron positions relative to the *C. intestinalis* coding region. Introns are numbered sequentially (1 to 12). Intron slippage is shown by the *asterisk.*

interior nodes and sequences showed no significant rate differences within vertebrates and cephalochordates ($p > 0.05$) (Table 1). Relative rate tests were also performed for the ADH class 1 enzyme. In this case, the amphioxus class 3 was used as the outgroup and a neighbor-joining tree topology was built. From the initial 23 sequences, 19 showed a constant rate of evolution (Table 1).

Amino acid substitution rates for ADH3 and ADH1 were estimated (Table 2). Five pairs of groups were established: (a) primates/rodents, (b) reptiles/mammals, (c) amphibians/amniotes, (d) fish/tetrapods and (e) hagfish/gnathostomes. Divergence times (T) for each pair deduced from fossil records are 65–100 MYA for a, 288–310 MYA for b, 350 MYA for c, 400 MYA for d, and 555 MYA for e (Dayhoff 1978; Archibald 1996; Benton 1997; Lee 1999; Shu et al. 1999). Multiple calibration points were used to minimize errors (Lee 1999) because too few sequences were available for some animal groups. The average evolutionary distance for each pair was plotted against the minimal divergence time. A significant linear dependence of k_{aa} on T was observed, with a correlation coefficient of 0.97 and 0.82 for ADH classes 3 and 1, respectively. From the slope of the regression line, the amino acid substitutions per site per year (ν_{aa}) for ADH3 and ADH1 were estimated to be 0.27×10^{-9} and 0.98×10^{-9}, respectively. This 3.6-fold difference could not be assigned to differential mutation rates, since both genes showed similar rates of synonymous substitutions at the DNA level (data not shown).

Class 3/1 Duplication

The amino acid differences between class 3 and class 1 and their evolutionary rates were used to estimate the

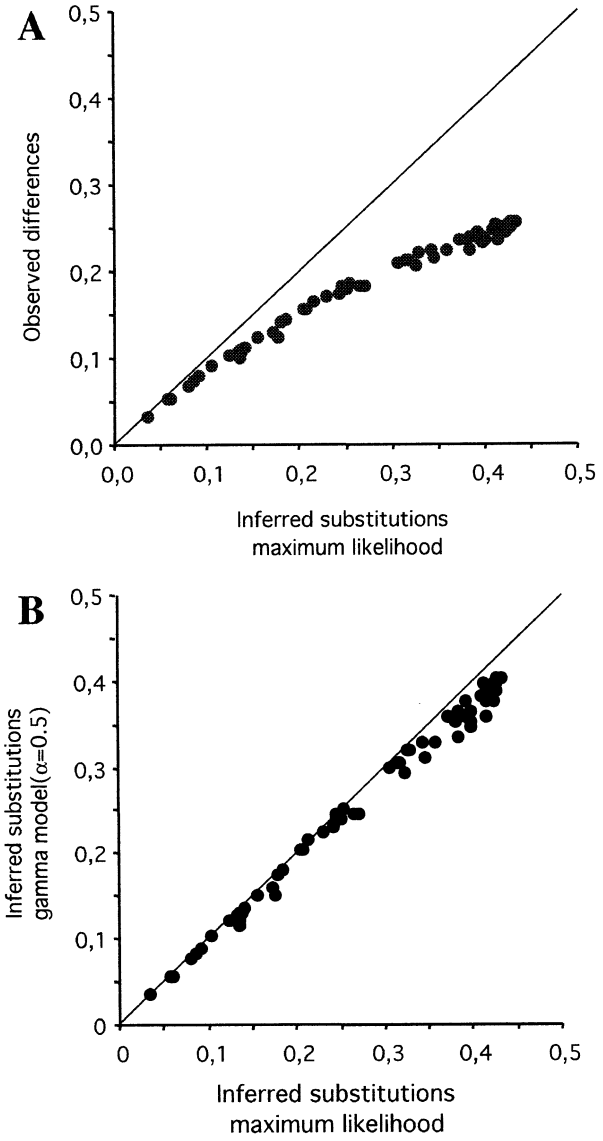


Fig. 2. Estimates of the number of amino acid substitutions per site in ADH class 3 enzymes. **A** Substitution saturation curve. The *y* axis shows the frequency of observed differences between pairs of species sequences, and the *x* axis shows the inferred distance between the same two sequences determined by maximum likelihood. Each *point* therefore represents one pair of sequences, considering the observed over the inferred number of substitutions. The resulting curve lies between the diagonal line (no saturation) and a horizontal plateau line (fully saturated), which means that the ADH3 set is informative and only moderately saturated. **B** Comparison of the numbers of ADH3 substitutions per site estimated by the γ-distance correction (*y* axis) versus those estimated by the maximum likelihood method (*x* axis).

class duplication time. The amino acid distances between both classes were calculated using a γ-corrected distribution with an α of 0.8. The α value was slightly higher than that obtained for each particular class (α = 0.5 for ADH3 and 0.6 for ADH1), due to the different arrangements of the conserved regions in the proteins compared. Since class 1 has not been found in hagfish, its ADH3 data were not used for the comparison. From all the comparisons of each class 3 *vs* class 1 from codfish to human, an average k_{aa} was obtained. The duplication time, estimated as $T = k_{aa}/(\nu_{aa}\text{-ADH3} + \nu_{aa}\text{-ADH1})$, was 493–500 MYA. Therefore, the ADH family would have expanded after the cephalochordate/vertebrate split [682–694 MYA for ADH3 (our estimate) (Table 2) and 680–700 MYA for triosephosphate isomerase and 790–860 MYA for aldolase C (Nikoh et al. 1997)].

Table 1. ADH class 3 and class 1 tests for homogeneity of divergence rates

		Test	$\chi^{2\,a}$	df	*p* value
ADH3	Vertebrates/ cephalochordates	Two-cluster	10.08	14	0.69
		Branch-length	10.85	14	0.62
ADH1	Vertebrates[b]	Two-cluster	13.77	18	0.74
		Branch-length	23.85	18	0.16

[a] Combines the rate differences for all the interior nodes (*n*) under the root, with $n - 1$ degrees of freedom (df).
[b] The rat, deer mouse, and lizard are excluded.

Discussion

Evidence in Favor of the ADH Family Expansion After the Cephalochordate/Vertebrate Split

Intron positions are relevant to elucidate gene duplication events. According to our analysis, all vertebrate ADH classes share an identical nine-exon structure (Zgombic-Knight et al. 1995; Strömberg and Höög 2000) and show marked differences from that of prevertebrates (Fig. 1B). Several hypothetical scenarios for the ancestral exon structure and the duplication events through chordate evolution have been considered (Fig. 3). Consistent with the most parsimonious process, the ancestral gene structure contained 11 exons and the duplication events leading to the various vertebrate classes took place after the cephalochordate/vertebrate split (Fig. 3A). Moreover, under this hypothesis, no recurrent events would have to be postulated. Concerning the *Adh3* architecture, ascidians and amphioxus show a higher number of introns but of a smaller size than those of vertebrates, in agreement with other reported genes (Boeddrich et al. 1999). With regard to protostomes, *Adh3* exhibits a lower intron number than deuterostomes and introns of a smaller size. More protostome *Adh* sequences should be characterized to infer the ancestral metazoan structure.

Estimates of the amino acid distances between the vertebrate ADH3 and the vertebrate ADH1 classes and of the evolutionary rates make it possible to infer the duplication event to be at about 500 MYA (Table 2), clearly after the cephalochordate/vertebrate split (682–694 MYA), probably subsequent to the hagfish divergence (555 MYA) and before the origin of tetrapods (400

Table 2. Evolutionary distances and divergence times among different animal groups estimated from the amino acid sequences of ADH3; Duplication time of the class 3 and class 1 enzymes

	Divergence time (MYA)[a]	γ distance (k_{aa})	
	Fossils (calibration points)	ADH3 (α = 0.5)	ADH1 (α = 0.6)
Animal groups			
Primates/rodents	65–100	0.06	0.21
Reptiles/mammals	288–310	0.12	0.41
Amphibians/amniotes	350	—	0.53
Fish/tetrapods	400	0.23	1.03
Hagfish/gnathostomes	555	0.30	—
	ADH3 (molecular estimations)[b]		
Cephalochordates/vertebrates	682–694	0.37	—
B. floridae/*B. lanceolatum*	186–189	0.10	—
Correlation coefficient[c]		0.97–0.96	0.82–0.81
ν_{aa} (10^{-9}/site/year)[d]		0.27–0.27	0.98–0.96
Duplication time, class 3 and class 1 (α = 0.8)	493–500	0.62	

[a] The two calibration points separated by a dash constitute the minimal confidence value from the fossil record and the widest divergence time accepted as a calibration point for each animal group (Lee 1999). Those calculations that were based on both values are separated by a dash.
[b] The molecular estimates were derived from the constant rate of amino acid substitution of ADH3.
[c] The regression line was constrained to pass through the origin.
[d] The evolutionary rate ν_{aa} was estimated from the slope of the regression line.

MYA). Biochemical data also support this duplication to be restricted to the vertebrate lineage since no ethanol-active class 1 form has been detected in amphioxus (Cañestro et al. 2000) and hagfish (Danielsson et al. 1994b).

Amphioxus ADHs: B. floridae–B. lanceolatum *Speciation*

Remarkably, *Adh3* is more similar between human and mouse than between the two amphioxus species analyzed, *B. floridae* and *B. lanceolatum.* Concerning gene structure, vertebrate species share all intron positions but this is not the case for intron 11 in amphioxus, although the other introns appear to be fully conserved (Fig. 1). At the protein level, amino acid differences amounted to 9.3% for the amphioxus species, whereas they were only 7.2% for human and mouse. The constant evolutionary rate of ADH3 makes it possible to infer that the two amphioxus species may have diverged around 190 MYA (Table 2), during the midsuperior Jurassic. Although such a high value would be unexpected on morphological grounds, since the two species appear very similar, it would be consistent with the human and mouse divergence time of 65–100 MYA. Moreover, similar data were obtained with cytoplasmic and muscle actin (data not shown). Speciation of the two cephalochordates—*B. floridae* lives along the Atlantic coast of America, whereas *B. lanceolatum* is found along the Atlantic coast of Europe and in the Mediterranean sea (Poss and Boschung 1996)—could have been a consequence of the drifting of major land masses, beginning around 200 MYA from the Pangea (Marzoli et al. 1999). Therefore, it could be argued that present similarities in the two amphioxus species would not be based on recent divergence but, rather, on the absence of major morphological changes during the last 200 million years (Myr) or thereabouts. In fact, this morphological stability supports the "archetypal" organization postulated for amphioxus (Garcia-Fernàndez and Holland 1994).

Relationship Between ADH Family Expansion and Vertebrate Genome Evolution

How does the ADH family expansion fit into the predicted genome duplications (the 2R hypothesis) in early vertebrate evolution? Paralogous chromosomal locations of some gene family members of the human and mouse have been argued to constitute remnants of the two or three rounds of tetraploidization (Lundin 1993). However, genes encoding the ADHs are not spread over different chromosomes but cluster in the human genome at 4q21-25 and in the syntenic region of mouse chromosome 3 (Giri et al. 1989). Furthermore, the 2R hypothesis would predict a (AB)(CD) phylogenetic tree topology for the four resulting genes (Hughes 1999), although this does not always hold true for all gene families that were generated during early vertebrate evolution (Wang and Gu 2000). Phylogenetic analysis of the ADH family rendered an (A(B(CD))) topology, also in agreement with the physical map of the ADH members on the human chromosome 4 (Fig. 4). In this topology, the pattern of family expansion fits better with tandem duplication

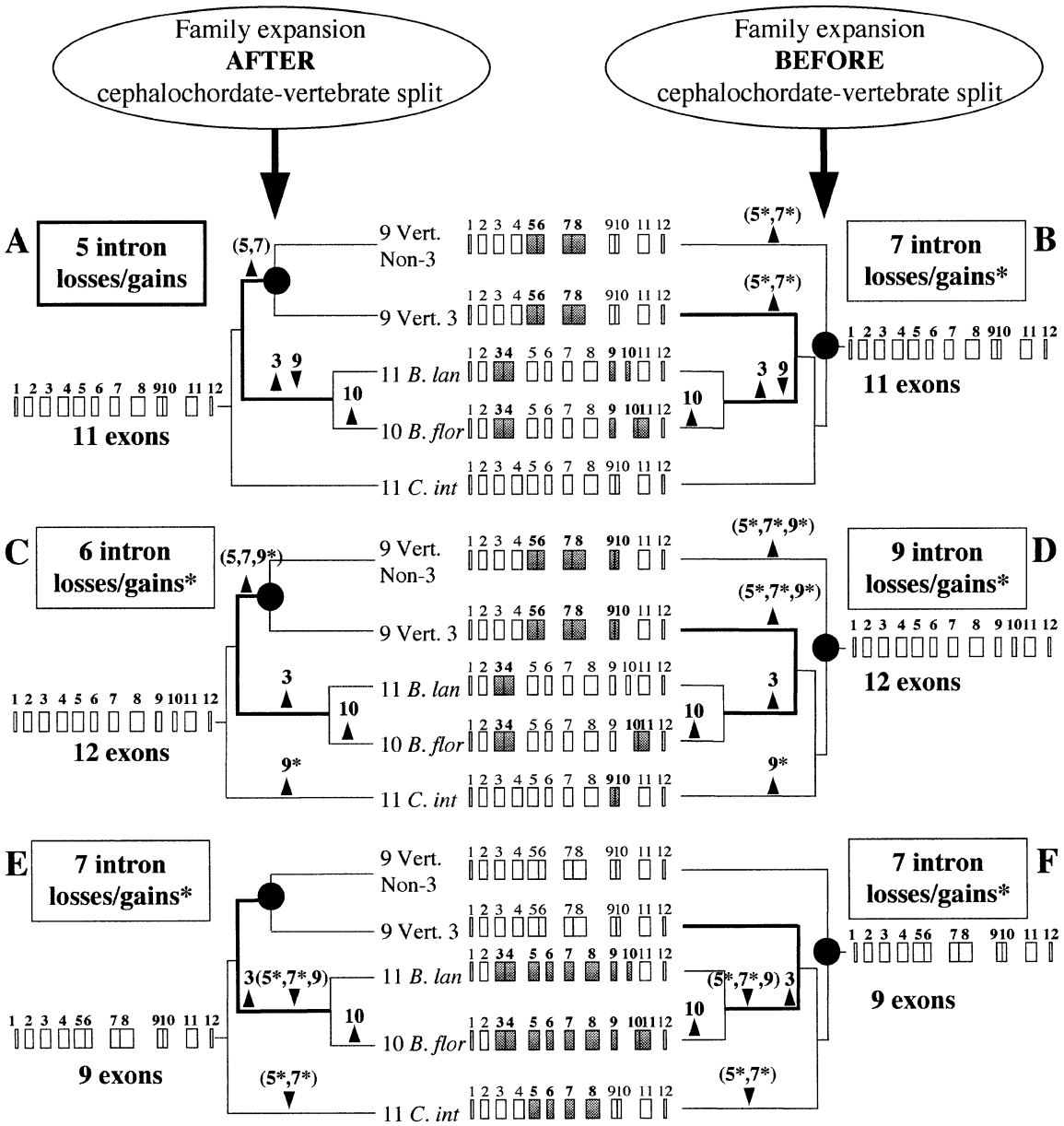


Fig. 3. *Adh* intron gains and losses assuming that the family expansion took place after **(A, C, E),** or before **(B, D, F),** the cephalochordate/vertebrate split. The putative ancestral exon structure is shown at the root of each tree [A and B, 11 exons (ascidians); C and D, 12 exons (maximum number); E and F, 9 exons (vertebrates)] and the actual-structure at the end of each branch. The duplication event and the cephalochordate/vertebrate split are shown (*filled circles* and *bold lines,* respectively). The assumed intron gains (▼) and losses (▲) are indicated (*boxed*), and the exons involved are labeled (*shaded boxes* and *bold face numbers*). Recurrent losses/gains in different lineages are marked with asterisks.

events than with whole or partial genome rearrangements assumed under the 2R hypothesis. Extensive gene duplications are thought to be responsible for the burst of isoforms during early vertebrate evolution, before the fish/tetrapode split, and probably these duplications were not directly linked to the Cambrian explosion (Suga et al. 1999).

Functional Divergence and Evolutionary Changes

Gene duplications allow the generation of new functions (Ohno 1970; Hughes 1994). In the ADH family, the data gathered on the structural and enzymological features of the different isoforms support the orthologous relationships derived from the phylogenetic analysis and provide valuable clues to discriminate between the ancestral and the new functions. The amphioxus ADH is a typical class 3 enzyme (Cañestro et al. 2000), as are the octopus (Fernandez et al. 1993), the *Drosophila* (Luque et al. 1994), and the hagfish (Danielsson et al. 1994b) forms of the enzyme. Hence, all direct analyses provide evolutionary evidence that the ancestral form of the family was endowed with ADH3 activity. Therefore, we have named

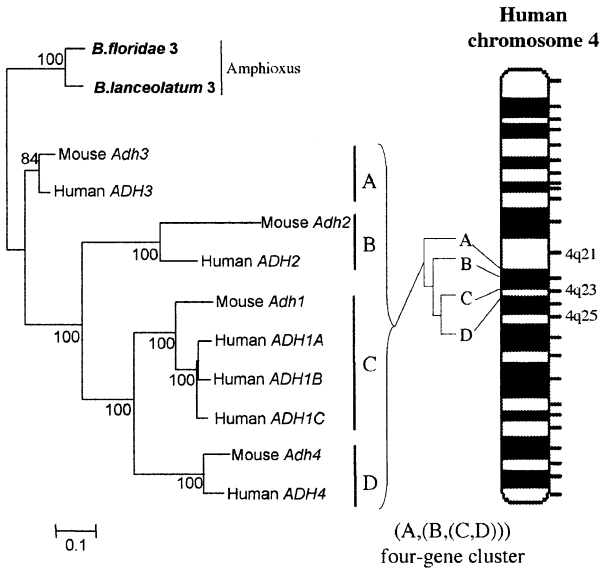


Fig. 4. Phylogenetic representation and chromosomal mapping of the *Adh* family. **Left:** γ-distance neighbor-joining tree for human/mouse *ADH* classes 1, 2, 3 and 4, rooted with the amphioxus *Adh*s. Bootstrap values are indicated for each branch. **Center:** (A(B(CD))) phylogenetic pattern of the *Adh* family. **Right:** Map position of the *ADH* gene cluster on human chromosome 4: (A) *ADH3*(χ), 420,780–404,222 bp; (B) *ADH2*(π), 476,032–453,229 bp; (C) *ADH1A*(α), 622,843–608,229 bp; *ADH1B*(β), 684,498–638,355 bp; (D) *ADH4*(μ), 761,336–744,028 bp on the reverse strand of the contig NT_022863. Information gathered at http://www.ncbi.nlm.nih.gov/AceView/ ADH* search. The ADH nomenclature is according to Duester et al. (1999).

the ascidian and amphioxus *Adh* genes *Adh3*. The enzymatic activity expressed as k_{cat}/K_m for this enzyme remains constant—within a factor of 3—in insects, amphioxus, and human. Moreover, the amino acid substitution rate for ADH3 remains unchanged after the duplication event, whereas ADH1 shows a 3.6-fold increase, also supported by their different branch lengths in the neighbor-joining tree (Fig. 4). Thus, the "functional density," defined as the fraction of a polypeptide chain involved in specific functions, which is inversely related to the evolutionary rate (Zuckerkandl 1976), could be assumed to be higher in ADH3 than in ADH1, in accordance with the classical view of a "variable" class 1 and a "constant" class 3 enzyme (Danielsson et al. 1994a). Moreover, while the segments of maximal variation of class 1 are at functionally important regions, class 3 variability is restrained to superficial, less critical positions (Persson et al. 1993; Danielsson et al. 1994a). In summary, the present data strengthen the concept of an ancestral position for the ADH3 activity.

Some of the newly generated forms deserve further consideration. ADH4 has been assumed to be crucial in vertebrates for retinoic acid synthesis via retinol oxidation (Duester 1996). Retinoic acid is thought to be involved in the anterior/posterior axis formation in both ascidian and amphioxus and, also, in vertebrates (reviewed by Shimeld 1996). But the fact that amphioxus ADH3 is not a multifunctional ADH, correlated with the absence of ADH4 in lower chordates and with data on the microsomal retinol dehydrogenase of amphioxus (Dalfó et al. 2001), points to a candidate other than ADH for carrying out the retinoic acid synthesis in animal lines in general.

Finally, in spite of the biochemical and evolutionary constancy of ADH3, changes in the promoter elements are evident when the amphioxus and vertebrate ADH3 forms are compared. Whereas a ubiquitous *Adh3* expression has been shown in mouse and human (Giri et al. 1989; Ang at al. 1996; Estonius et al. 1996), a specific expression, restricted to the midgut region, has been shown in amphioxus by *in situ* hybridization (Cañestro et al. 2000). Similarly, we have found, by *in situ* hybridization, an *Adh3* tissue specificity in *Drosophila* and ascidian (data not shown), much like that which has been demonstrated for other ADH classes in vertebrates.

Acknowledgments. We thank M. Riutort for helpful discussions on the phylogenetic analysis, W. Stone for helpful advice, M. Levine and A. Di Gregario for the *C. intestinalis* genomic library, R. Di Lauro for the cDNA *C. intestinalis* library, and J. Garcia-Fernàndez for the *B. floridae* genomic library. This work was supported by Ministerio de Ciencia y Tecnología (Spain) Grants BMC2000-0536 and UE98-0014, EU Contract BIO4 CT 97-2123, the Swedish MRC (Project 3532), CIRIT (Generalitat de Catalunya, 1997FI00007), and an FPI fellowship to C.C. We are indebted to the Serveis Científico-Tècnics (UB) for DNA sequence analysis.

References

Abad L (2000) Utilizacion del gen octanol deshidrogenasa (*Odh*) como marcador molecular para el análisis de poblaciones de *D. subobscura.* Master thesis, Department of Genetics, University of Barcelona, Barcelona, Spain.

Ang HL, Deltour L, Hayamizu TF, Zgombic-Knight M, Duester G (1996) Retinoic acid synthesis in mouse embryos during gastrulation and craniofacial development linked to class IV alcohol dehydrogenase gene expression. J Biol Chem 271:9526–9534

Archibald JD (1996) Fossil evidence for a late cretaceous origin of "hoofed" mammals. Science 272:1150–1153

Benton MJ (1997) Vertebrate paleontology. Chapman & Hall, New York, Curr Opin Genet Dev 8:694–700

Boeddrich A, Burgtorf C, Hennig S, Panopoulou C, Steffens C, Borzym K, Lehrach H (1999) Sequence analysis of an amphioxus cosmid containing a gene homologous to members of the aldo-keto reductase gene superfamily. Gene 230:207–214

Cañestro C, Hjelmqvist L, Albalat R, Garcia-Fernàndez J, Gonzàlez-Duarte R, Jörnvall H (2000) Amphioxus alcohol dehydrogenase is a class 3 form of single type and of structural conservation but with unique developmental expression. Eur J Biochem 267:6511–6518

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Anal Biochem 116:156–159

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA 81:1991–1995

Dalfó D, Cañestro C, Albalat R, Gonzàlez-Duarte R (2001) Characterization of a microsomal retinol dehydrogenase gene from amphioxus: retinoid metabolism before vertebrates. Chem-Biol Interact 130:359–370

Danielsson O, Jörnvall H (1992) "Enzymogenesis": Classical liver alcohol dehydrogenase origin from the glutathione-dependent form-

aldehyde dehydrogenase line. Proc Natl Acad Sci USA 89:9247–9251

Danielsson O, Eklund H, Jörnvall H (1992) The major piscine liver alcohol dehydrogenase has class-mixed properties in relation to mammalian alcohol dehydrogenases of classes I and III. Biochemistry 31:3751–3759

Danielsson O, Atrian S, Luque T, Hjelmqvist L, Gonzàlez-Duarte R, Jörnvall H (1994a) Fundamental molecular differences between alcohol dehydrogenase classes. Proc Natl Acad Sci USA 91:4980–4984

Danielsson O, Shafqat J, Estonius M, Jörnvall H (1994b) Alcohol dehydrogenases class III contrasted to class I. Characterization of the cyclostome enzyme, the existence of multiple forms as for the human enzyme, and distant cross-species hybridization. Eur J Biochem 225:1081–1088

Dayhoff MO (1978) Survey of new data & computer methods of analysis. eds. Dayhoff MO vol 5, suppl 3, National Biochemical Research Foundation, Washington, DC

Duester G (1996) Involvement of alcohol dehydrogenase, short-chain dehydrogenase/reductase, aldehyde dehydrogenase, and cytochrome P-450 in the control of retinoid signaling by activation of retinoic acid synthesis. Biochemistry 35:12221–12227

Duester G, Farrés J, Felder VC, Höög JO, Parés X, Plapp B, Yin SJ, Jörnvall H (1999) Recommended nomenclature for the vertebrate alcohol dehydrogenase gene family. Biochem Pharmacol 58:389–395

Estonius M, Svensson S, Höög JO (1996) Alcohol dehydrogenase in human tissues: Localisation of transcripts coding for five classes of the enzyme. FEBS Lett 397:338–342

Fernàndez MR, Jörnvall H, Moreno A, Kaiser R, Parés X (1993) Cephalopod alcohol dehydrogenase: Purification and enzymatic characterization. FEBS Lett 328:235–238

Garcia-Fernàndez J, Holland PWH (1994) Archetypal organization of the amphioxus *Hox* gene cluster. Nature 370:563–566

Giri PR, Krug JF, Kozak C, Moretti T, O'Brien SJ, Seuanez HN, Goldman D (1989) Cloning and comparative mapping of a human class III (chi) alcohol dehydrogenase cDNA. Biochem Biophys Res Commun 164:453–460

Holland PWH (1996) Molecular biology of lancelets: Insights into development and evolution. Israel J Zool 42:247–272

Hughes AL (1994) The evolution of functionally novel proteins after gene duplication. Proc R Soc Lond B Biol Sci 256:119–124

Hughes AL (1999) Phylogenies of developmentally important proteins do not support the hypothesis of two rounds of genome duplication early in vertebrate history. J Mol Evol 48:565–576

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. CABIOS 8:275–282

Jörnvall H, Höög J-0 (1995) Nomenclature of alcohol dehydrogenases. Alcohol Alcohol 30:153–161

Koivusalo M, Baumann M, Uotila L (1989) Evidence for the identity of glutathione-dependent formaldehyde dehydrogenase and class III alcohol dehydrogenase. FEBS Lett 257:105–109

Kuma K, Miyata T (1994) Mammalian phylogeny inferred from multiple protein data. Jpn J Genet 69:555–566

Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. Nature 392:917–920

Kumar S, Tamura K, Nei M (1993) *MEGA: Molecular evolutionary genetics analysis, v. 1.02.* Pennsylvania State University, University Park

Lee MSY (1999) Molecular clock calibrations and metazoan divergence dates. J Mol Evol 49:385–391

Lundin LG (1993) Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. Genomics 16:1–19

Lundin LG (1999) Gene duplication in early metazoan evolution. Semin Cell Dev Biol 10:523–530

Luque T, Atrian S, Danielsson O, Jörnvall H, Gonzàlez-Duarte R (1994) Structure of the *Drosophila melanogaster* glutathione-dependent formaldehyde dehydrogenase/octanol dehydrogenase gene. Eur J Biochem 225:985–993

Marzoli A, Renne PR, Piccirillo EM, Ernesto M, Bellieni G, De Min A (1999) Extensive 200-million-year-old continental flood basalts of the central Atlantic magmatic province. Science 284:616–618

Naylor GJP, Brown WM (1997) Structural biology and phylogenetic estimation. Nature 388:527–528

Nikoh N, Iwabe N, Kuma K, Ohno M, Sugiyama T, Watanabe Y, Yasui K, Shi-Cui Z, Hori K, Shimura Y, Miyata T (1997) An estimate of divergence time of parazoa and eumetazoa and that of cephalochordata and vertebrate by aldolase and triose phosphate isomerase clocks. J Mol Evol 45:97–106

Ohno S (1970) Evolution by gene duplication. Springer-Verlag. Heidelberg

Persson B, Bergman T, Keung WM, Waldenström U, Holmquist B, Vallee BL, Jörnvall H (1993) Basic features of class-I alcohol dehydrogenase: variable and constant segments coordinated by inter-class and intra-class variability. Eur J Biochem 216:49–56

Philippe H, Chenuil A, Adoutte A (1994) Can the cambrian explosion be inferred through molecular phylogeny? Development 120 (Suppl):15–25

Poss SG, Boschung HT (1996) Lancelets (Cephalochordata: Branchiostomatidae): How many species are valid? Israel J Zool 42:13–66

Roger AJ, Sandblom O, Doolittle WF, Philippe H (1999) An evaluation of elongation factor 1α as a phylogenetic marker for eukaryotes. Mol Biol Evol 16:218–233

Sharman AC, Holland PWH (1996) Conservation, duplication and divergence of developmental genes during chordate evolution. Neth J Zool 46:47–67

Shimeld S (1996) Retinoic acid, Hox genes and the anterior-posterior axis in chordates. BioEssays 18:613–616

Shimeld S (1999) Gene function, gene networks and the fate of duplicated genes. Semin Cell Dev Biol 10:549–553

Shu D-G, Luo H-L, Morris SC, Hu X-L, Chen L, Han J, Zhu M, Li Y, Chen LZ (1999) Lower Cambrian vertebrates from south China. Nature 402:42–46

Skrabanek L, Wolfe KH (1998) Eukaryote genome duplication—Where's the evidence? Curr Opin Genet Dev 8:694–700

Strimmer K, von Haeseler A (1996) Quartet puzzling: A quartet maximum likelihood method for reconstructing tree topologies. Mol Biol Evol 13:964–969

Strömberg P, Höög JO (2000) Human class V alcohol dehydrogenase (ADH5): A complex transcription unit generates C-terminal multiplicity. Biochem Biophys Res Commun 278:544–549

Suga H, Hoshiyama D, Kuraku S, Katoh K, Kubokawa K, Miyata T (1999) Protein tyrosine kinase cDNAs from amphioxus, hagfish and lamprey: Isoform duplications around the divergence of cyclostomes and gnathostomes. J Mol Evol 49:601–608

Takezaki N, Rzhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized trees. Mol Biol Evol 12:823–833

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Wang DY, Kumar S, Hedges SB (1999) Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. Proc R Soc Lond B Biol Sci 266:163–171

Wang Y, Gu X (2000) Evolutionary patterns of gene families generated in the early stage of vertebrates. J Mol Evol 51:88–96

Zgombic-Knight M, Foglio MH, Duester G (1995) Genomic structure and expression of the ADH7 gene encoding human class IV alcohol dehydrogenase, the form most efficient for retinol metabolism in vitro. J Biol Chem 270:4305–4311

Zuckerkandl E (1976) Evolutionary process and evolutionary noise at the molecular level. I. Functional density in proteins. J Mol Evol 7:167–183

CAPÍTOL III

En aquest capítol, 1) la comparació dels patrons d'expressió de l'***Adh3*** en diferents organismes (***Drosophila melanogaster***, ***Ciona intestinalis*** i ***Danio rerio***) amb els prèviament estudiats en amfiox i ratolí, permet inferir que el patró de l'expressió ancestral de l'***Adh3*** era específic de teixit. 2) Es caracteritza l'estat de metilació de l'ADH3 al llarg de l'evolució dels cordats, i es descarta com a possible mecanisme modulador del canvi d'expressió observat. 3) S'inicia la caracterització de les regions reguladores de l'***Adh3*** d'organismes procordats mitjançant tècniques de transgènesi, amb la transferència a embrions de ***C. intestinalis*** de construccions amb diferents promotors controlant l'expressió del gen reporter ***LacZ***.

Anàlisi evolutiva dels patrons d'expressió de la família gènica de les Alcohol deshidrogenases de cadena mitjana

Summary

The *Adh* family has evolved into a complex system of mammalian isoforms, which differ in substrate recognition and tissue specificity. We have shown previously that Amphioxus *Adh3* expression is restricted to the mid-gut, contrary to what had been reported in mammals, which is ubiquitously expressed. We now report the *in situ* hybridization analysis of *Adh3* in *Drosophila*, ascidian (*C.intestinalis*) and zebrafish (*D. rerio*). Our data show that expression is contained to the *Drosophila* fat body and to the digestive endoderm of ascidian larvae, whereas transcripts are ubiquitously detected in the vertebrate fish. These data support that the changes in the expression domains occurred in the vertebrate lineage, and possibly related to the expansion event. Moreover, methylation does not seem to contribute to the *Adh* expression pattern, as in all chordates analyzed it has been found in the methylated fraction of the genome. Finally, to characterize the regulatory elements of *Adh*3 and draw the comparisons between subphyla, constructs containing the *LacZ* reporter under the regulatory regions of ascidian and amphioxus *Adh*3 have been electroporated into ascidian embryos.

Introducció

Les duplicacions gèniques han jugat un paper clau en l'evolució proveint el material genètic necessari per l'adquisició de noves funcions. La família de les Alcohol deshidrogenases de cadena mitjana (*Adh*-MDR) es va expandir mitjançant duplicacions successives, posteriors a la separació dels llinatges dels vertebrats i cefalocordats, on els nous gens varen evolucionar per donar lloc a enzims amb activitats diverses, definint així les set classes descrites fins l'actualitat. Per contra, en els invertebrats sols s'ha caracteritzat un únic gen *Adh*-MDR, homòleg al que codifica per l'ADH3 a vertebrats amb activitat formaldehid deshidrogenasa depenent de glutatió. Això semblaria indicar que l'activitat de classe 3 correspondria a la forma ancestral i que les noves funcions haurien aparegut mitjançant canvis en les regions codificants, adquirint noves capacitats metabòliques

Nombrosos estudis bioquímics han demostrat que les propietats catalítiques de les ADH3 s'han mantingut constants al llarg de l'evolució animal, a insectes, cefalòpodes, amfioxs, rèptils i mamífers. Per contra, el patró d'expressió ha estat únicament analitzat en cefalocordats i alguns mamífers (ratolí i home). Aquests patrons presenten diferències significatives: així en cefalocordats, l'*Adh3* s'expressa de forma específica en l'intestí mitjà durant el desenvolupament, mentre que en mamífers presenta un patró ubic. Aquests resultats indiquen canvis en la regulació gènica, possiblement associats a l'expansió d'aquesta família gènica durant la transició cap als vertebrats. Per contrastar aquesta hipòtesi hem analitzat l'expressió de *l'Adh3* a altres espècies animals, estratègicament situades respecte a l'expansió: *Drosophila melanogaster*, *Ciona intestinalis* i *Danio rerio*. Les noves dades han permès comparar el patró d'expressió en un protòstom (*D. melanogaster*) i quatre deuteròstoms: dos prevertebrats (*C. intestinalis* i *B. floridae*) i dos vertebrats (*D. rerio* i *M. musculus*). També s'ha analitzat l'estat de metilació de l'*Adh3 en aquests animals*, donat que la metilació és un possible factor modulador de l'expressió gènica, especialment en la transició invertebrats-vertebrats (Tweedie *et al*, 1997). Així, els gens amb una expressió més específica estarien freqüentment metilats, mentre que els gens d'expressió basal i ubiqua estarien no metilats. Finalment, per caracteritzar els elements reguladors dels patrons d'expressió i estudiar la seva evolució, s'han realitzat experiments preliminars de transgènesi. S'han electroporat embrions d'ascidis (*C.*

intestinalis) en estadi d'una cèl·lula amb construccions que contenen el gen reporter *LacZ* sota el control de diferents dominis reguladors de l'*Adh3* d'ascidi i amfiox, iniciant així l'estudi del grau de conservació dels elements reguladors entre aquests dos subfila.

Material i mètodes

Obtenció de DNA genòmic, RNA missatger i síntesi de cDNA

A partir d'animals sencers, excepte en els ascidis on la túnica i l'estómac van ser prèviament eliminats, es va obtenir DNA genòmic i RNA de peix zebra d'aleta curta (*Danio rerio*) i d'ascidis (*Ciona intestinalis*). Les purificacions es van dur a terme seguint el mètode de l'isotiocianat de guanidina (Chirgwin *et al*, 1979), (Chomczynski & Sacchi, 1987) amb lleus modificacions (Cañestro *et al*, 2000). El cDNA es va sintetitzar utilitzant transcriptasa reversa M-MuLV (Stratagene) i l'oligonucleòtid poli(dT).

Anàlisis Southern i patrons de metilació

Entre 5-10 µg de DNA per cada individu adult de *C. intestinalis* van ser digerits amb *BamH*I, *EcoR*I i *Pst*I. La separació electroforètica es va realitzar en un gel d'agarosa al 0.9% i els fragments de DNA varen ésser transferits a membranes de niló. La sonda emprada en la hibridació corresponia a l'exó 9 de *CiAdh3*, marcada amb ^{32}P (Boehringer Mannheim Random Primed labeling kit). La hibridació i els rentats es varen dur a terme a alta rigorositat (Cañestro *et al*, 2000).

Per l'estudi de la metilació, DNA genòmic de cada espècie va ser digerit amb *Hpa*II i *Msp*I, separat electroforèticament, transferit i hibridat en les mateixes condicions prèviament descrites.

Hibridacions *in situ* "*whole mount*"

Per obtenir les sondes per les hibridacions *in situ* es van amplificar per PCR els següents fragments: *i*) 652 bp que corresponen a un fragment de cDNA des de l'exó 4 fins al 9 de l'*Adh3* de *C. intestinalis* (Cañestro *et al*, 2002) ; *ii)* 528 bp (aa 92-268) del exó 3 de l'Adh3 de *Drosophila melanogaster* (Luque *et al*, 1994); *iii*) 452 bp de l'extrem 5' de la regió codificant d'una EST de peix zebra (GeneBank AI331882) que presentava homologia de seqüència amb una ADH3, amplificats per RT-PCR utilitzant els oligonucleòtids específics *PZIII-1S*, 5' GAT CAA ATG TAA GGC AGC 3' i *PZIII1R* 5' GAT CTC AGC CAC CAC GGT 3'. Tots els fragments van ser clonats en plasmidis KS-pBluescript per *Sma*I. Les ribosondes van ser marcades amb digoxigenina (Boehringer Mannheim DIG RNA labeling kit). Com a control negatiu d'hibridació es van emprar les respectives ribosondes amb sentit. Les hibridacions *in situ* van ser efectuades seguint els protocols estàndard (Jiang *et al*, 1991, Schulte-Merker *et al*, 1992, Corbo *et al*, 1997).

Construccions de DNA per l'electroporació: Promotor *Adh3 + LacZ*

Les construccions van ser dissenyades a partir de les seqüències de fags aïllats que contenien les regions a 5' dels gens *CiAdh3* de *Ciona intestinalis*, *BlAdh3* de *Branchiostoma lanceolatum*, i *BfAdh3* de *Branchiostoma floridae* (Cañestro *et al*, 2002). Aquests fragments van ser clonats en el vector pSP1.72-27 a 5' del gen *LacZ* (Corbo *et al*, 1997). Com a control positiu d'electroporació es va emprar la construcció de 3.5 kb del promotor del gen de *Brachyury* de *Ciona intestinals* que promou una forta expressió a la notocorda (Corbo *et al*, 1997).

Per localitzar els elements reguladors de la transcripció del gen *CiAdh3*, un fragment de 3.2 kb (la distància fins el següent gen) i un fragment de 1.6 kb a 5' de la regió

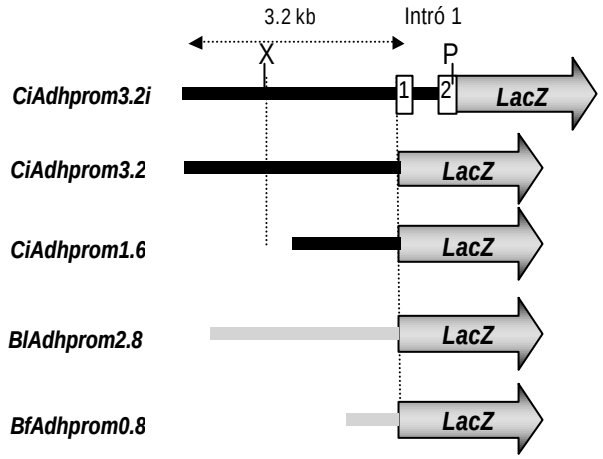


Figura 1. Representació esquemàtica de les construccions amb el gen reporter *LacZ* sota el control de diferents regions genòmiques reguladores de l'*Adh3* d'ascidis i amfioxs (*CiAdh3*: negre, *BlAdh3* i *BfAdh3*: entramat gris), emprades en les electroporacions. La distància intergènica entre l'extrem 5' de *CiAdh3* i el següent gen és de 3.2 kb. La construcció *CiAdhprom3.2i* inclou l'exó 1(blanc), l'intró 1 i part de l'exó 2 fusionat al gen *LacZ*. Les dianes *Xba*I (X) i *Pst*I (P) estan indicades.

codificant, van ser amplificats per PCR (Deep Vent DNA polimerasa, New England Biolabs) amb l'encebador *pUCrev* (5' ATA ACA ATT TCA CAC AGG 3'), localitzat sobre el fagmidi pBK-CMV (Stratagene), i l'encebador *CintAdhpromNot*, (5' A TAG TAT AGC GGC CGC AAT TTT GAT CGT TTT AAG AA 3'), sobre l'exó 1 del gen *CintAdh3* clonat en pBK-CMV. L'oligonucleòtid *CintAdhpromNot* introdueix una diana *Not*I a l'extrem 5', i substitueix l'ATG per ATT de forma que la traducció comença directament al gen *LacZ*. Aquests fragments de 3.2 kb i 1.6 kb van ser digerits i clonats en el vector pSP1.72-27 per *Sal*I-*Not*I i *Pst*I-*Not*I, generant les construccions *CiAdhprom3.2* i *CiAdhprom1.6*, respectivament.

Per generar la construcció *CiAdhprom3.2i*, que incloïa les 3.2 kb de la regió promotora més l'exó 1, l'intró 1 i part de l'exó 2 de *CiAdh3*, es va substituir un fragment *Xba*I-*Not*I del clon *CiAdhprom3.2*, per un fragment genòmic *Xba*I-*Pst*I, generant un gen de fusió de 19 codons del gen *CintAdh3* més el gen *LacZ* (fig. 1).

Les regions clonades en els fags més a 5' dels gens *BlAdh3* i *BfAdh3*, 2.5 kb i 0.8 kb respectivament, van ser amplificades amb l'encebador *Lfix_X/S-N* (5' ACT CTA GAG CTC CAG GCC GCG AGC TC 3') situat en el extrem 3' del braç curt del fag lambda FixII, el qual eliminava la diana *Not*I, i l'encebadors específic *BlAdhpromNot* (5'ATA GTA TAG CGG CCG CAG CGT CTG TCG GCC AAA GG 3') situat davant de l'inici de traducció de *BlAdh3* i l'encebador *BfAdhpromNot* (5' A TAC TAT AGC GGC CGC AAT GCT TGC AGC TTC AAC T 3') on l'ATG de *BfAdh3* quedava modificat com prèviament s'havia descrit. Els productes de PCR van ser digerits i clonats en el vector pSP1.72-27 per *Xho*I i *Not*I. Les dues construccions varen ser dissenyades de forma que l'inici de traducció fos el del gen *LacZ*.

Electroporació de DNA en embrions d'ascidis

El protocol utilitzat és una modificació de Corbo *et al.* (1997). Tots els experiments es van dur a terme a una temperatura controlada d'uns 17-18°C.

Fecundació d'ascidis

Per cada experiment, 2 o 3 animals adults van ser disseccionats i es varen obtenir els oòcits del gonoducte amb una pipeta Pasteur de vidre. Tot el material de vidre que entrà en contacte amb els embrions va ser prèviament siliconitzat aplicant Sigmacote (Sigma) i esbandit abundantment amb aigua destil·lada o submergit en aigua destil·lada durant una nit. Els ous van ser mantinguts varies hores a 17-18°C en una placa de petri de 3 cm de diàmetre sense ser diluïts en aigua de mar. L'esperma de tots els individus va ésser recollida en un únic eppendorf i guardada a 4°C. Per evitar fecundacions accidentals, tot el material que entrava en contacte amb l'esperma era rentat amb aigua destil·lada. La fecundació es va fer barrejant ous de diferents individus amb 2-5 µL d'esperma en un eppendorf siliconitzat.

Després de 10 minuts, es van afegir dues gotes de solució de decorionació (TP) per facilitar la deposició dels ous. Per preparar la solució TP es van barrejar volums iguals de 0.1% pronasa en aigua de mar artificial amb HEPES (ASWH) (420 mM NaCl, 9 mM KCl, 10 mM $CaCl_2.2H_2O$, 24.5 mM $MgCl_2.6H_2O$, 25.5 mM $MgSO_4.7H_2O$, 2.15 mM $NaHCO_3$, 10 mM HEPES pH 8) i 2% tioglicolat de sodi en ASWH, el pH es va ser ajustar a 10.0-10.4 amb 2.5 M NaOH i es va guardar a 4°C durant com a màxim una setmana. Els ous fecundats van ser dipositats en el fons del tub centrífugant durant 5 segons amb una minicentrífuga (Giralt).

Per la decorionació, es va descartar el màxim de líquid amb esperma i es va omplir l'eppendorf amb solució TP. Es va homogeneïtzar suaument amb una pipeta Pasteur. Després dels 2-3 primers minuts, el procés de decorionació va ser controlat cada 20 segons sota la lupa fins el moment en que la meitat dels embrions estigueren decorionats. Temps de decorionació superiors a 5 min varen ser letals. La decorionació s'aturà amb 4 rentats amb aigua artificial de mar (ASWH). En cada rentat, els embrions van ser dipositats molt suaument amb 2 o 3 impulsos de centrífuga a molt baixa velocitat. Aquest procediment es va realitzar ràpidament per no perdre la finestra d'electroporació (20 minuts després de la fecundació).

Electroporació

En el darrer rentat es van deixar els embrions decorionats en un volum de 200 µl d'ASWH i es van afegir 30-50 µg de DNA en 500 µl de Manitol 0.77 M. Es van barrejar suaument i van ser transferits a una cubeta d'electroporació de 0.4 cm (BioRad). Es va procedir a l'electroporació amb un únic pols elèctric (voltatge de descàrrega: 50 V, capacitància: 1400µF,

resistència: 24 Ohms) d'aproximadament 17 mseg en un electroporador BTX600.

Els embrions van ser recollits i es van deixar desenvolupar a 17-18°C en 50 µg/mL de gentamicina en ASWH en una placa de petri recoberta d'una capa fina d'agarosa al 1% en ASWH.

Detecció de l'activitat *LacZ*

Els embrions van ser fixats amb glutaraldehid 0.2% en ASWH lliure de calci i magnesi (CMF-ASW) (463 mM NaCl, 11 mM KCl, 2.15 mM $NaHCO_3$, 25.5 mM Na_2SO_4)durant 30 minuts. Fixacions més llargues inhibien l'activitat β–galactosidasa.

Es van fer 2 rentats de 10 minuts amb PBT (8.1 mM Na_2HPO_4, 1.6 mM NaH_2PO_4, 150 mM NaCl, pH 7.5) i un rentat en solució de tinció (1mM $MgCl_2.6H_2O$, 3mM $K_4Fe(CN)_6.3H_2O$, 3mM $K_3Fe(CN)_6$ en PBT) durant 5 min. La incubació es va fer a la foscor a 37 °C en 400 µg/mL X-gal en solució de tinció. Per aturar la tinció es van rentar 2 cops durant 10 minuts amb PBT i es van post-fixar amb 4% paraformaldehid en PBT.

Resultats

Anàlisi Southern i patró de metilació.

L'anàlisi Southern sobre DNA digerit amb tres enzims diferents, revelà que *CintAdh3* és un gen de còpia única. (fig. 1). Les mides de les bandes observades concordaren amb les esperades segons els clons aïllats en el cribratge genòmic

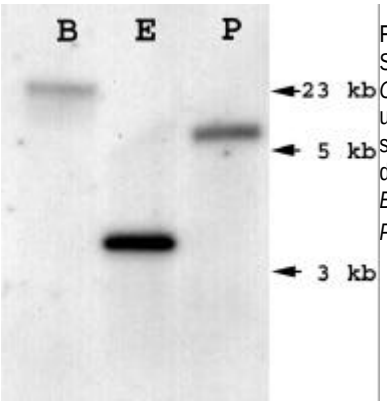


Figura 1. Anàlisi Southern del gen *CiAdh3*, hibridat amb una sonda de l'exó 9, sobre DNA genòmic d'un individu digerit amb *BamH*I (B), *EcoR*I (E) *Pst*I (P).

realitzat anteriorment (Cañestro *et al*, 2002).

Per estudiar l'estat de metilació dels membres de la família de les *Adh*s, es va fer una anàlisi Southern amb DNA genòmic de *Drosophila*, ascidi, amfiox, peix zebra i ratolí, comparant el patró d'hibridació obtingut amb *Msp*I amb el del seu esquizòmer sensible a metilació *Hpa*II. Tots els gens analitzats de la família de les *Adh*s estaven metilats amb excepció de *Drosophila*, ja que mostraren patrons de bandes diferents entre els carril digerits amb *Msp*I i *Hpa*II (fig.2).

Els gens *Adh* es troben metilats tant en els procordats com en els vertebrats, suggerint que els canvis en l'estat de metilació, un dels mecanismes propis de la regulació gènica dels vertebrats, no serien responsables dels canvis d'expressió.

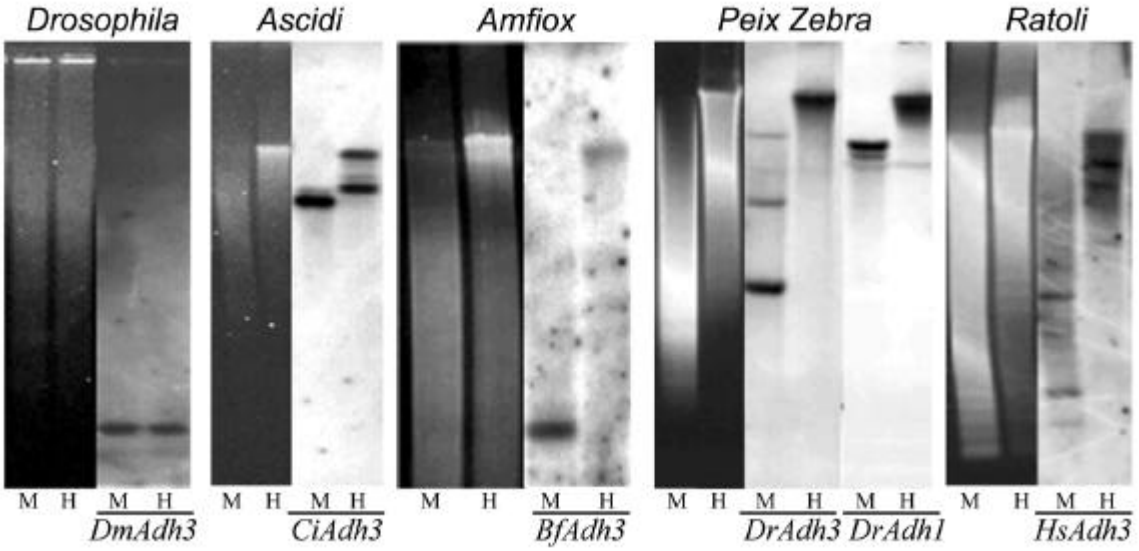


Figura 2. Patrons de metilació de l'*Adh3* i *Adh1* al llarg del regne animal: *Drosophila*, Ascidi, Amfiox, Peix Zebra i Ratolí. Per cada organisme es presenta el patró de digestió dels enzims *Msp*I (M) i *Hpa*II (H) resolt en un gel d'agarosa i la hibridació de la transferència d'aquell mateix gel amb diferents sondes (indicades sota la línia horitzontal): exó 3 de *DmAdh3* (Luque *et al*., 1994), *CiAdh3*: l'exó 9 de *CintAdh3*, l'exó 8 de *BfAdh3* (reproduït de Cañestro *et. al*., 2000), un fragment de 452 bp de l'EST AI331882 de *DrAdh3* i un fragment de 450 bp de l'EST AI476976 (450 bp) de *DrAdh1* de peix zebra, i el cDNA de *HsAdh3* humana com a sonda heteròloga amb el DNA de ratolí.

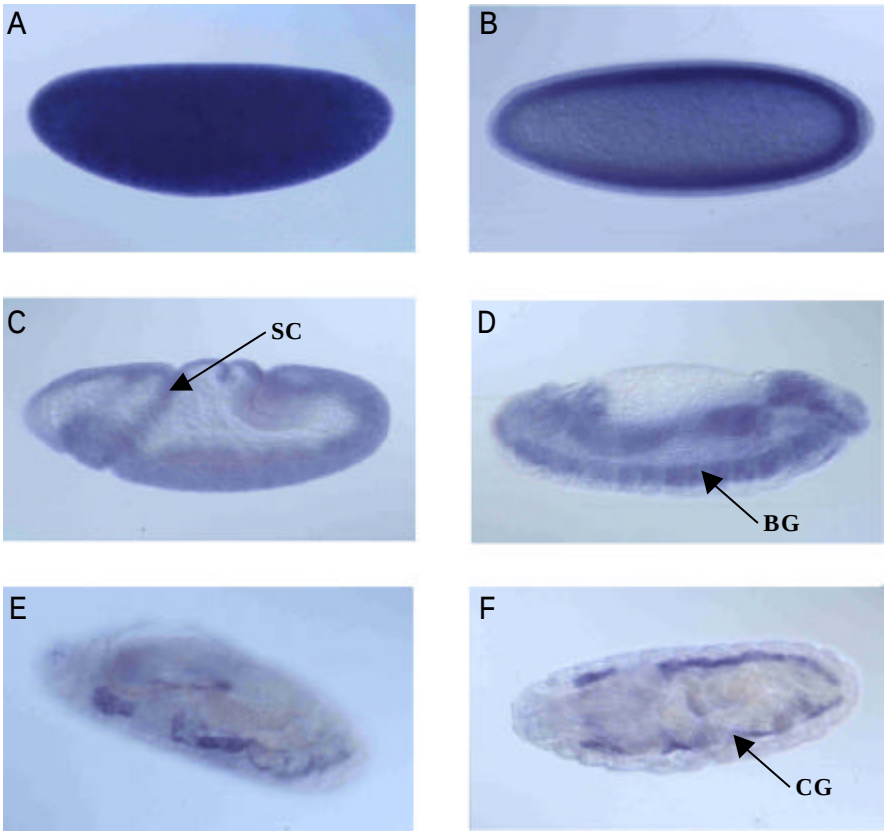


Figura 3. Anàlisi per hibridació *in situ* de l'expressió de l'*Adh3* al llarg del desenvolupament de *Drosophila melanogaster*. (A) estadi 1, ou fecundat, (B) estadi 5, blastoderm, (C) estadi 8, (D) estadi 13, (E-F) estadi 17. Tots els embrions estan orientats amb la part anterior cap a l'esquerra i la part dorsal cap a dalt, excepte la imatge F que està en visió dorsal. SC: Solc cefàlic; BG: Banda germinal; CG: Cos gras

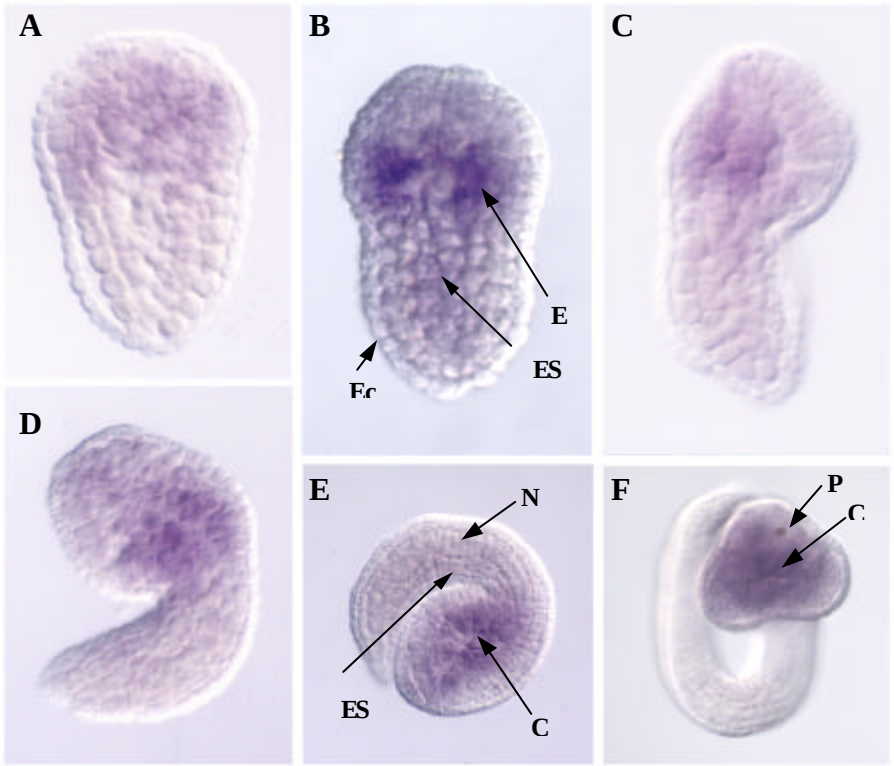


Figura 4 Anàlisi per hibridació *in situ* de l'expressió de l'*Adh3* al llarg del desenvolupament de *Ciona intestinalis*: gàstrula tardana (A), estadi de *tailbud* (B), estadi de *tailbud* tardà (C), estadi de capgrós primerenc (D), estadi de capgrós mitjà (E) i estadi de capgrós tardà (F).
CG: cavitat gàstrica, E: endoderm, ES: banda endodèrmica, Ec: ectoderm, NO: Notocordi, P: taca pigmentària.

Patrons d'expressió de l'*Adh3* a *Drosophila melanogaster*, *Ciona intestinalis* i *Danio rerio*

Per entendre les diferències en els patrons d'expressió del gen *Adh3* observades entre amfioxs i mamífers i intentar establir alguna relació entre aquestes i l'expansió de la família gènica, es van analitzar els patrons d'expressió al llarg del desenvolupament dels gens homòlegs a l'*Adh3* de diversos organismes: un protòstom (*Drosophila*), un deuteròstom prevertebrat (l'urocordat *C. intestinalis*) i un vertebrat (el peix zebra *D. rerio*).

El senyal d'hibridació observat a *D. melanogaster* fou intens i distribuït uniformement en l'ou fecundat (estadi 1) (fig. 3), fent-se restringit a mida que avançava la compartimentació cel·lular en la capa perifèrica del blastoderm (estadi 5). Durant la gastrulació, el senyal d'hibridació disminuïa, quedant localitzat majoritàriament en la banda germinal, en alguns solcs corporals transitoris (p.e., el cefàlic) i en la invaginació amnioproctodeal, una cavitat formada per la invaginació de la part posterior del precursor de l'intestí mitjà (estadi 8). El senyal es mantingué en la banda germinal durant el procés d'allargament i segmentació, mentre que l'amnioserosa es trobava lliure de senyal (estadi 13). Finalment, en els estadis més avançats en els que es diferenciaven els òrgans interns, l'expressió de l'*Adh3* va fer-se intensa i restringida al cos gras (estadi 17).

A *C. intestinalis* el senyal d'hibridació començà a ser clarament visible al final de la gastrulació, en l'endoderm més anterior (fig. 4). El baix nivell d'expressió en aquest organismes obligà a temps de tinció llargs (més de 48 hores). Quan l'embrió començà a formar l'esbós de la cua ("*tailbud*"), l'*Adh3* s'expressava en dos grups de cèl·lules de l'endoderm anterior situades simètricament a ambdós costats del pla sagital mitjà, mentre que la banda endodèrmica de la cua presentava un senyal molt feble. El senyal a l'endoderm anterior es mantingué al llarg del desenvolupament fins l'estadi de capgrós ("*tadpole*"), desapareixent a la cua. L'endoderm que delimita la cavitat gàstrica era el que presentva uns nivells d'expressió més elevats.

El patró d'expressió de l'*Adh3* a *D. rerio* era semblant al descrit a ratolí, molt generalitzat i estès per tot l'embrió (Ang *et al*, 1996) (fig. 4). No obstant, l'expressió no era completament uniforme, modificant-se al llarg del temps. A les 18 hores de desenvolupament, l'expressió de l'*Adh3* era intensa en els dos extrems de l'embrió i més feble en les regions centrals. S'expressà especialment en diferents regions del cervell anterior i mitjà, mentre que en la cua l'expressió va romandre localitzada en el teixit mesodèrmic de la cua. En estadis més tardans, a mida que les diferents regions del sistema nerviós central s'anaven diferenciant, la intensitat de l'expressió de l'*Adh3* s'estengué cap al cervell posterior, cerebel i cordó nerviós. Els primordis de les aletes pectorals també presentaren un nivell d'expressió elevat. D'altra banda, a mida que la cua s'anava allargant, l'expressió es mantingué en les regions més posteriors, continuant restringida al mesoderm muscular. També es detectà expressió en teixits que proliferaven a la part dorsal del vitel més posterior.

Patrons d'expressió en embrions de *C. intestinalis* electroporats amb el promotor de l'*Adh3* i el *LacZ*

Per caracteritzar els elements que dirigeixen l'expressió del gen *CiAdh3*, es van electroporar embrions de *C. intestinalis* amb construccions de la regió reguladora de *CiAdh3* i el reporter *LacZ*. A més, per estudiar el grau de conservació dels elements reguladors de l'expressió de l'*Adh3* al llarg de l'evolució dels cordats, construccions amb els promotors dels gens d'amfiox *BlAdh3* i *BfAdh3* també van ésser electroporades en embrions d'ascidis.

Les condicions d'electroporació descrites a Corbo *et al*. (1997) (Corbo *et al*, 1997) varen ser adaptades a l'electroporador BTX600, fent servir com a control positiu una construcció amb el promotor del gen *CiBra* (Corbo *et al*, 1997) (Fig.7). Els assajos preliminars fets amb les construccions amb la regió reguladora del gen *CiAdh3* mostraren un senyal feble, i per detectar l'activitat β-galactosidasa eren necessàries incubacions llargues de més de 72 h, mentre que per a la construcció *CiBra*, 2 hores eren suficients. En larves de 24 h la detecció del gen reporter quedà restringida principalment en l'endoderm i mesènquima del cap. La part més anterior de la cua presentava un senyal feble, mentre que la resta de la cua i l'epidermis no mostraren senyal. Cal esmentar la baixa viabilitat i l'elevat grau de malformacions observades en els

embrions electroporats amb les construccions de *CiAdh3* en comparació amb les fetes amb *CiBra*. A més, un gran percentatge d'embrions malformats o aturats en els primers estadis dels desenvolupament mostraren activitat *LacZ*.

Finalment, amb les construccions de l'*Adh3* d'amfiox mai es va aconseguir cap senyal en ésser introduïts en embrions de *C. intestinalis*.

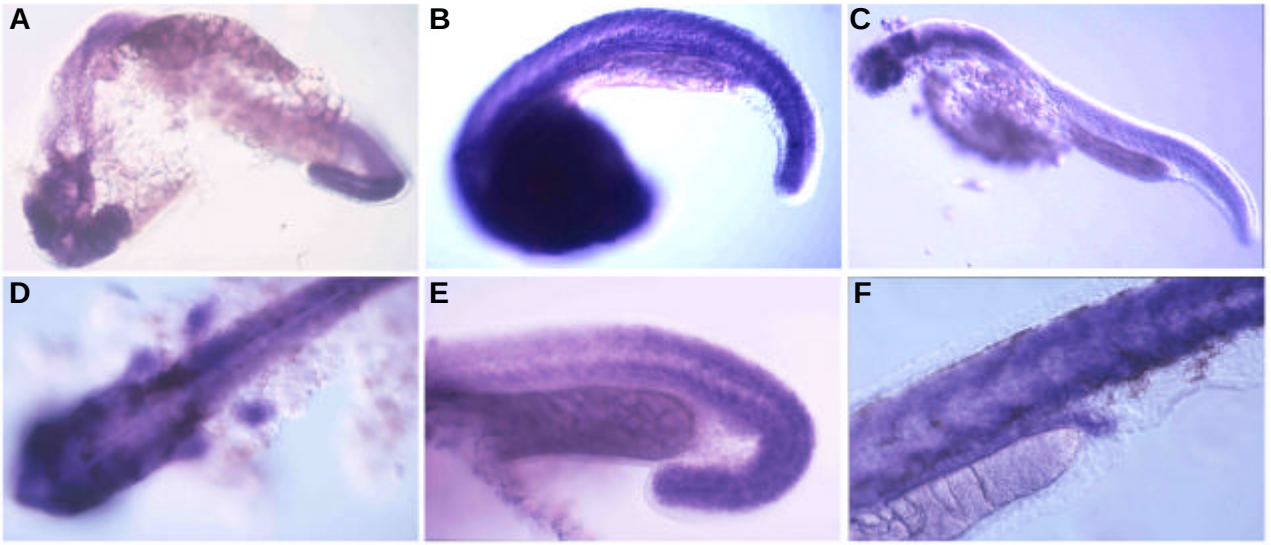


Figura 6. Anàlisi per hibridació *in situ* de l'expressió de l'Adh3 al llarg del desenvolupament de *Danio rerio*. En la part superior es mostren tres estadis diferents: 18h (A), 28 h (B) i 48 h (C). En la part inferior es mostren detalls ampliats de l'expressió de l'Adh3 en el cap i primordis de les aletes pectorals en un embrió de 72 h (D), la cua d'un embrió de 20 h (E) i la part posterior del vitel en un embrió de 60 h (F)

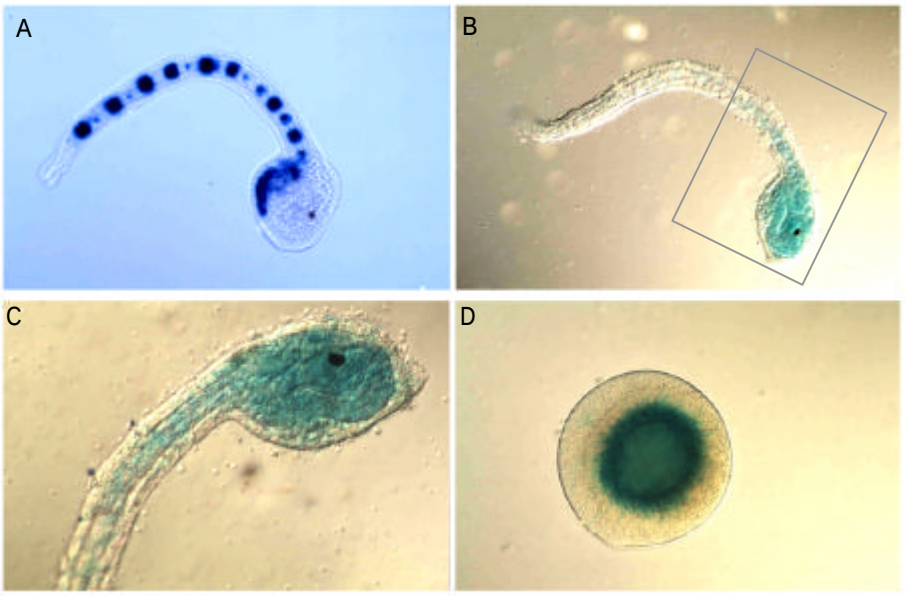


Figura 7. Detecció de l'activitat del gen reporter *LacZ* en embrions de *C. intestinalis* electroporats. (A) CiBra3.5 (control positiu): larva de 18 h amb expressió en la notocorda i part del mesènquima anterior. (B) *CiAdhprom3.2i*: larva de 24 h amb expressió a l'endoderm, mesoderm i mesènquima del cap; el rectangle indica la imatge amplificada en (C); (D): embrió no desenvolupat aturat en l'estadi d'una cèl·lula on es detecta senyal β-galactosidasa.

Discussió

L'expressió de *l'Adh3* al llarg de l'evolució animal

Els resultats de les hibridacions *in situ* indiquen que l'expressió de l'*Adh3* a *Drosophila*, ascidi i amfiox es troba restringida a determinats teixits, els quals directe o indirectament estan relacionats amb funcions digestives: cos gras, cavitat gàstrica i intestí mitjà. Per contra el patró d'expressió a peix zebra és molt més estès i generalitzat, similar al descrit a home i ratolí (Estonius *et al*, 1996, Ang *et al*, 1996). Donat que el patró d'expressió restringit és comú a protòstoms (*Drosophila*) i a deuteròstoms (ascidis i amfiox), és probable que aquest fos el patró ancestral. L'expressió ubiqua hauria de ser considerada una innovació del llinatge dels vertebrats, probablement concomitant amb l'expansió de la família gènica ara fa uns 500 MA (Cañestro *et al*, 2002). És lògic pensar que les duplicacions en tàndem de les *Adh*s haurien afectat no tan sols a les regions codificants, sinó també a les regions promotores, amb pèrdues o guanys d'elements reguladors responsables de l'especificitat tissular. En aquest sentit, és interessant destacar que l'expressió de l'*Adh3* a organismes prevertebrats és similar a la d'altres membres de la família de les *Adh*s de vertebrats, com per exemple l'*Adh*4 a l'estómac (Moreno & Parés, 1991, Moreno *et al*, 1996) i l'*Adh*1 al fetge i l'intestí (Smith *et al*, 1971). Aquest fet suggeriria que després l'evolució dels vertebrats, una còpia hauria mantingut les propietats bioquímiques ancestrals, però hauria perdut la especificitat tissular, mentre que les altres haurien adquirit noves activitats per a metabolitzar nous substrats (etanol, retinol, hormones esteroidees), però haurien mantingut el patró específic de teixit.

Regulació de l'expressió de l'*Adh3*

Investigacions recents sobre el promotor de l'*Adh3* de mamífers han descrit l'existència de nombroses caixes reguladores de la transcripció reconegudes per factors específics de teixit, C/EBP, HNF-5, E-boxes, Ap-1, AP-2, Ets-1, SRE, XRE i Heat shock proteins (revisat a Edenberg, 2000). Tot i que l'expressió de l'*Adh3* ha estat descrita com "ubiqua", els nivells entre diferents teixits varien de forma important. Alguns dels elements reguladors de l'*Adh3* són compartits amb altres membres de la família de les *Adh*s i amb altres gens que comparteixen una expressió específica al fetge de vertebrats o al cos gras de *Drosophila*. En aquest darrer organisme, caldria destacar que l'*Adh3*-MDR i l'*Adh*-SDR, responsable del metabolisme de l'etanol en aquest organisme, comparteixen una expressió comuna al cos gras. La caracterització del promotor de l'*Adh*-SDR en diferents espècies de *Drosophila* ha revelat que les C/EBP i AEF-1s són dos dels factors de transcripció determinants per la regulació de l'expressió específica al cos gras. Sorprenentment, factors homòlegs a vertebrats també regulen l'expressió de l'*Adh1*-MDR en el fetge. Anàlisis de retard en gel i protecció a la DNAsa, han mostrat com les C/EBPs de *Drosophila* i ratolí poden reconèixer els elements reguladors de l'altre organisme, suggerint un mecanisme evolutivament conservat a dos gens que codifiquen proteïnes amb convergència funcional. Anàlisis de predicció dels promotors de l'*Adh3*-MDR de *Drosophila*, ascidi i amfiox han revelat l'existència de seqüències similars a la seqüència consens dels llocs d'unió de C/EBP i AEF-1. L'estudi de la conservació d'unitats de regulació de les Adhs de diferents organismes podria aclarir com ha esdevingut l'evolució dels promotors i explicar els diversos patrons d'expressió aquí descrits.

El intens senyal de l'hibridació *in situ* en l'ou fecundat de *Drosophila*, que va disminuint en el transcurs de la compartimentació cel·lular i la gastrulació, indicaria un origen matern dels nivells d'*Adh3* en els primers estadis del desenvolupament. Cal esmentar que l'*Adh*-SDR també presenta un clar component matern (Visa *et al*, 1992). Aquest fet suposaria un altre tret comú entre la regulació de l'*Adh3*-MDR i l'*Adh*-SDR de *Drosophila*. A més, estudis per RT-PCR quantitativa sobre l'expressió de l'*Adh3* de peix zebra al llarg del desenvolupament indiquen que els nivells de mRNA des dels primers estadis té un origen matern, suggerint que el component matern també tindria un paper important en la regulació de l'*Adh3* en vertebrats (Dasmahapatra *et al*, 2001). El nivell d'expressió en estadis d'una o poques cèl·lules no ha estat estudiat per hibridació *in situ* en urocordats, cefalocordats ni vertebrats. Donat que la importància de l'efecte matern depèn de l'estratègia de desenvolupament adoptada per cada tipus d'organisme (mida de l'ou, velocitat i tipus de divisió cel·lular, ...), més estudis serien necessaris per

determinar la importància de l'efecte matern en la regulació dels nivells d'ADH3 i la seva necessitat durant el desenvolupament embrionari.

Caracterització del promotor de l'*Adh3* a procordats

Els experiments per caracteritzar el promotor de l'*Adh3* de procordats mitjançant electroporacions de construccions de DNA amb el gen *LacZ* sota el control de la regió reguladora de l'*Adh3* són encara molt preliminars, i el fet que no reprodueixin el mateix patró obtingut per anàlisi d'hibridació *in situ* dificulta avançar en aquesta direcció. No obstant, el senyal observat en embrions electroporats en estadis d'una o dues cèl·lules podria suggerir que els nivells de transcripció son elevats després de la fecundació. A més, la detecció de l'activitat β-galactosidasa té efecte de "memòria", degut a que la vida mitjana de l'enzim és de 48 hores (Smith *et al*, 1995), i la seva detecció pot allargar-se molt després de la seva síntesis interferint amb la detecció del gen reporter en estadis posteriors. Estudis d'hibridació *in situ* en estadis molt primerencs del desenvolupament i millores tècniques, com la detecció *in situ* de l'mRNA del *LacZ* enlloc de la detecció de l'activitat enzimàtica, o la utilització de sistemes de proteïnes fluorescents que canvien de color segons el temps transcorregut des de la seva síntesi (Living Colors™ Fluorescent Timer, Clontech), serien necessaris per validar els nostres resultats i arribar a poder utilitzar la tècnica dels transgènics per estudiar l'evolució funcional de les *Adh*s.

Agraïments

A R. Valero i a T. Luque per haver-nos cedit DNA genòmic de ratolí (*Mus musculus*) i de *D. melanogaster*, respectivament. To U. Oppermann for providing the human *Adh3* cDNA. A S. Gonzàlez-Crespo per la seva contribució en les hibridacions *in situ* de *Drosophila*. To S. Shimeld, P. Lemaire and C. Hudson for sharing with us their invaluable knowledge in the field of ascidian *in situ* hybridization and electroporation.

Referències

- Ang, H. L., Deltour, L., Hayamizu, T. F., Zgombic-Knight, M., Duester, G. (1996) Retinoic acid synthesis in mouse embryos during gastrulation and craniofacial development linked to class IV alcohol dehydrogenase gene expression. *J Biol Chem* 271:9526-34.
- Cañestro, C., Albalat, R., Hjelmqvist, L., Godoy, L., Jörnvall, H., Gonzàlez-Duarte, R. (2002) Ascidian and Amphioxus Adh genes correlate functional and molecular features of the ADH Family expansion during vertebrate evolution. *J Mol Evol* 54:81-89
- Cañestro, C., Hjelmqvist, L., Albalat, R., Garcia-Fernàndez, J., Gonzàlez-Duarte, R., Jörnvall, H. (2000) Amphioxus alcohol dehydrogenase is a class 3 form of single type and of structural conservation but with unique developmental expression. *Eur J Biochem* 267:6511-8.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., Rutter, W. J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18:5294-9.
- Chomczynski, P., Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162:156-9.
- Corbo, J. C., Levine, M., Zeller, R. W. (1997) Characterization of a notochord-specific enhancer from the Brachyury promoter region of the ascidian, Ciona intestinalis. *Development* 124:589-602.
- Dasmahapatra, A. K., Doucet, H. L., Bhattacharyya, C., Carvan, M. J., 3rd (2001) Developmental Expression of Alcohol Dehydrogenase (ADH3) in Zebrafish (Danio rerio). *Biochem Biophys Res Commun* 286:1082-6.
- Edenberg, H. J. (2000) Regulation of the mammalian alcohol dehydrogenase genes. *Prog Nucleic Acid Res Mol Biol* 64:295-341
- Estonius, M., Svensson, S., Höög Jan, O. (1996) Alcohol dehydrogenase in human tissues: Localisation of transcripts coding for five classes of the enzyme. *FEBS Letters* 397:338-342
- Jiang, J., Kosman, D., Ip, Y. T., Levine, M. (1991) The dorsal morphogen gradient regulates the mesoderm determinant twist in early Drosophila embryos. *Genes Dev* 5:1881-91.
- Luque, T., Atrian, S., Danielsson, O., Jörnvall, H., Gonzalez-Duarte, R. (1994) Structure of the Drosophila melanogaster glutathione-dependent formaldehyde dehydrogenase/octanol dehydrogenase

gene (class III alcohol dehydrogenase). Evolutionary pathway of the alcohol dehydrogenase genes. *Eur J Biochem* 225:985-93.

- Moreno, A., Farrés, J., Parés, X., Jörnvall, H., Persson, B. (1996) Molecular modelling of human gastric alcohol dehydrogenase (class IV) and substrate docking: differences towards the classical liver enzyme (class I). *FEBS Lett* 395:99-102.
- Moreno, A., Parés, X. (1991) Purification and characterization of a new alcohol dehydrogenase from human stomach. *Journal of Biological Chemistry* 266:1128-1133
- Schulte-Merker, S., Ho, R. K., Herrmann, B. G., Nusslein-Volhard, C. (1992) The protein product of the zebrafish homologue of the mouse T gene is expressed in nuclei of the germ ring and the notochord of the early embryo. *Development* 116:1021-32.
- Smith, M., Hopkinson, D. A., Harris, H. (1971) Developmental changes and polymorphism in human alcohol dehydrogenase. *Ann Hum Genet* 34:251-71.
- Smith, R. L., Geller, A. I., Escudero, K. W., Wilcox, C. L. (1995) Long-term expression in sensory neurons in tissue culture from herpes simplex virus type 1 (HSV-1) promoters in an HSV-1-derived vector. *J Virol* 69:4593-9.
- Tweedie, S., Charlton, J., Clark, V., Bird, A. (1997) Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Mol Cell Biol* 17:1469-75.
- Visa, N., Fibla, J., Gonzalez-Duarte, R., Santa-Cruz, M. C. (1992) Progressive redistribution of alcohol dehydrogenase during vitellogenesis in Drosophila melanogaster: characterization of ADH-positive bodies in mature oocytes. *Cell Tissue Res* 268:217-24.

CAPÍTOL IV

En aquest capítol, 1) es revela l'existència d'activitat **b**–galactosidasa endògena en els cefalocordats. 2). Es caracteritza aquesta activitat al llarg del desenvolupament de l'amfiox. 3) Es descriu la seva utilitat com a marcador histoquímic del sistema digestiu i per a exemplificar-ho, s'analitzen els efectes posteriorit zadors del tractament amb àcid retinoic en el desenvolupament de l'amfiox, on es posa de manifest la inhibició de la diferenciació funcional de tot el sistema digestiu anterior. 4) Es revela l'existència d'estructures funcionalment diferenciades al llarg del tracte intestinal. 5) Finalment, es faciliten solucions per evitar la detecció de l'activitat **b**–galactosidasa endògena en assajos en que el gen ***LacZ*** s'utilitzi com a reporter.

Dev Genes Evol (2001) 211:154–156
DOI 10.1007/s004270100137

EXPRESSION NOTE

Cristian Cañestro · Ricard Albalat · Hector Escrivà
Roser Gonzàlez-Duarte

Endogenous β-galactosidase activity in amphioxus: a useful histochemical marker for the digestive system

Received: 8 December 2000 / Accepted: 3 January 2001 / Published online: 20 February 2001

Abstract Endogenous β-galactosidase activity has been shown in the digestive tract of amphioxus from the larval to the adult stage and it can be easily followed as a histochemical marker. Enzymatic activity first appeared in 30-h larvae, became evident in 36-h larvae and remained in adults. In situ detection of β-galactosidase activity was used to monitor morphological and functional differentiation of the digestive system and the posteriorization of the endodermal structures in retinoic-acid treated embryos. The endogenous β-galactosidase activity was distinguished from the bacterial *lacZ* reporter by incubation at low pH.

Keywords Amphioxus · β-Galactosidase activity · Endodermal marker

Introduction

β-Galactosidases (β-gal) are lysosomal enzymes that cleave non-reducing β-D-galactose residues in β-D-galactosides. They contribute to glycolipid metabolism and their deficiency is associated with G_{M1} gangliosidosis, an inherited metabolic disorder (Gossrau et al. 1991). Endogenous β-gal activity has been reported in *Drosophila* and mammals (Lodja 1970; Schnetzer and Tyler 1996). Although there is some intra-specific variation, tissues that are rich in β-gal include intestine, kidney and epididymis (Conchie et al. 1958; Pearson et al. 1963). Furthermore, β-gal activity at pH 6.0 has been reported as a marker for cell senescence (Dimri et al. 1995).

Edited by J. Campos-Ortega

C. Cañestro · R. Albalat · R. Gonzàlez-Duarte (✉)
Departament de Genètica, Facultat de Biologia,
Universitat de Barcelona, 645 Av. Diagonal, 08028 Barcelona, Spain
e-mail: roser@bio.ub.es
Fax: +34-93-4110969

H. Escrivà
UMR 49 du CNRS, Ecole Normale Supérieure de Lyon,
46 Allée d'Italie, 69364 Lyon Cedex 07, France

Results and discussion

In the research reported here, endogenous β-gal activity has been detected in the chordate amphioxus, the closest living relative to vertebrates. This enzymatic activity appeared after staining incubation for over 48 h, at the ventral posterior endoderm of *Branchiostoma floridae* early larvae (30 h; Fig. 1A,B).

This was before larval feeding since the mouth had not yet opened. At 36 h, all the larvae showed β-gal activity in the midgut (Fig. 1C), and a strong signal was observed from the midgut to the anus at 50 h, whereas the pharyngeal portion remained unstained (Fig. 1D). In addition, a conspicuous signal was detected at the mid-hindgut junction, where the iliocolon ring develops. Although the digestive tract morphologically appeared as an uncompartmentalized tube the uneven staining pattern of 6-day and 15-day larvae revealed functional differentiation (Fig. 1E–H). The high β-gal activity at the mid-hindgut junction was located in two circular segments boundering the iliocolon ring (Fig. 1H). An intensely stained region was observed in the esophagus-midgut junction, from which the hepatic diverticulum will differentiate after metamorphosis. Another intensely stained small group of cells in the posterior left wall of the midgut could correspond to a new landmark of asymmetric amphioxus development (Fig. 1G,H), which could be used to draw new body part homologies between amphioxus and vertebrates (Holland and Holland 1999). The appearance of the 11th gill slit marks the beginning of amphioxus metamorphosis. Observations from this time onwards showed that the β-gal pattern remained in the gut and expanded into the developing hepatic diverticulum (Fig. 1F).

The involvement of retinoic acid (RA) in chordate anterior-posterior axis formation is well established. Exogenous RA affects endodermal development in ascidians

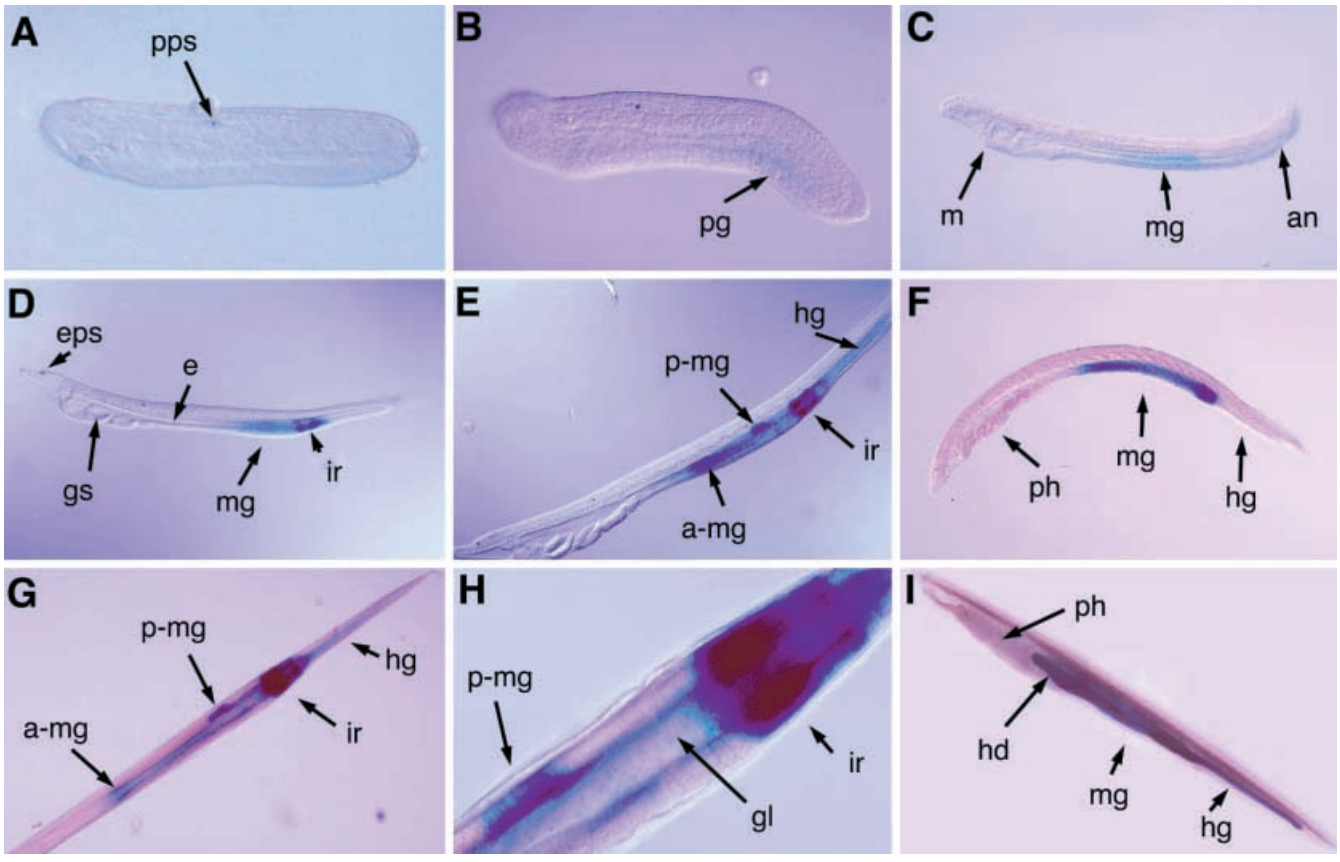


Fig. 1A–I β-Galactosidase (β-gal) activity detection during *Branchiostoma floridae* development. Animals were fixed in 0.2% glutaraldehyde in calcium/magnesium-free diluted artificial sea water (308 mM NaCl, 7.3 mM KCl, 1.4 mM $NaHCO_3$ and 17 mM Na_2SO_4) for 30 min at room temperature. After two washes in PBT (phosphate-buffered saline pH 7.2, Tween-20 0.1%) for 10 min each and once in staining buffer [1 mM $MgCl_26H_2O$, 3 mM $K_4Fe(CN)_63H_2O$ and 3 mM $K_3Fe(CN)_6$ in PBT] the samples were stained in fresh staining buffer with 0.4 mg/ml X-gal at 37°C. At alkaline pHs, 100 mM Tris-HCl buffer for the staining solution, instead of PBT, was used. The samples were fixed in 4% paraformaldehyde in PBT overnight at 4°C, washed three times in PBT and mounted in 80% glycerol in PBT. All specimens were oriented with the anterior end of the animal toward the left; in lateral views dorsal is to the *top*. **A** At the 24-h larval stage, β-gal activity was not detected. **B** β-Gal activity was weakly detected in the presumptive gut of 30-h larvae. **C** β-Gal stained the midgut of 36-h larvae uniformly. **D–F** Uneven staining showed gut compartmentalization in 4-day (**D**), 6-day (**E**) and 15-day (**F**) larvae. **G–H** Ventral view of 8-gill-slit larvae. Asymmetrical differentiation was detected on the left posterior midgut. **I** Strong β-gal activity was detected in the hepatic diverticulum of adult animals. *a-mg* Anterior midgut, *an* anus, *gs* gill slits, *gl* gut light, *hd* hepatic diverticulum, *hg* hindgut, *e* esophagus, *eps* eye pigment spot, *ir* iliocolon ring, *mg* mid-gut, *m* mouth, *ph* pharynx, *p-mg* posterior midgut, *pps* primary pigment spot

(urochordates; Hinman and Degnan 1998) and induces the loss of pharyngeal arches in vertebrates, due to *Hox* code alterations in neural crest migrating cells (Kraft et al. 1994; Lee et al. 1995). In amphioxus, RA treatment affects the pharyngeal development not by changing the *Hox* code but by inducing the overexpression of *Pax-1* and possibly other genes (Holland and Holland 1996). This posteriorization effect could be followed by the β-gal pattern of treated animals (Fig. 2A,B): their digestive tract resembled a blue tube that extended along the anterior-posterior axis (Fig. 2C,D). β-gal activity was detected in the digestive tract but not in the pharynx of wild-type embryos. Detection of β-gal activity in the presumptive pharyngeal region of treated animals indicated that RA not only prevented the formation of gill slit and mouth but probably induced an alteration of gene expression in the anterior endoderm and hence functional posteriorization.

The β-gal of amphioxus and mammals have acidic pH optima. They both behave poorly at weakly alkaline conditions, which are very favorable to the prokaryotic enzymes. On the other hand, an extensive antibiotic treatment (0.5 µg/ml penicillin, 0.5 µg/ml gentamycin and 1 µg/ml streptomycin) did not alter β-gal staining. Besides, β-gal activity detection preceded the mouth opening. Taken together, these findings indicated that β-gal activity is endogenous in amphioxus.

The *Lac Z* gene has been extensively used as a marker in gene transfer assays. In this case, histochemical discrimination of the prokaryotic activity from endogenous β-gal activity would be attainable at pH >8.5 under the conditions already described for mammals (Weiss et al. 1999). At these pHs the amphioxus β-gal activity was always negligible whereas *Lac-Z* expression of transformed *Escherichia coli DH5α* was clearly identified (data not shown). The discrimination between these two activities, β-gal endogenous and bacterial *Lac-Z*, is of relevance before using this prokaryotic reporter in gene transfer assays.

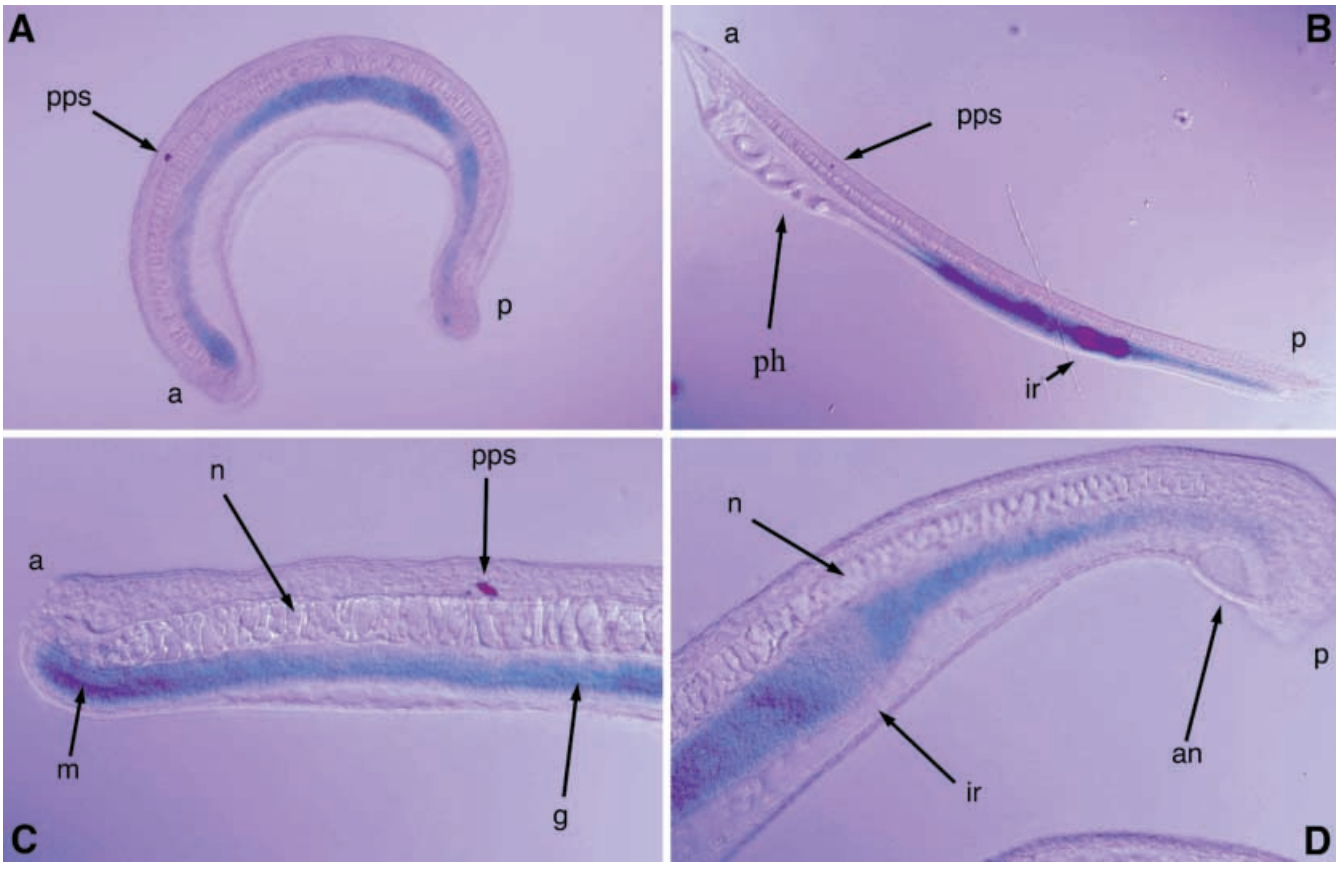


Fig. 2A–D β-Gal activity in retinoic acid (RA)-treated 6-day larvae: 2.5-h embryos were continuously treated in starvation with 10^{-6} M RA (1:1,000 dilution of 10^{-3} M RA in 100% DMSO) until processed, as described in Holland and Holland (1996). **A,B** RA-treated and untreated 6-day larvae, respectively. **C,D** Magnified view of the anterior and posterior ends of an RA-treated animal, respectively. *a* Anterior end, *an* anus, *g* gut, *ir* iliocolon ring, *n* notochord, *ph* pharynx, *p* posterior end, *pps* primary pigment spot

Acknowledgements We are indebted to L. Holland and N.D. Holland for helpful discussions. We thank J.M. Lawrence for laboratory facilities and Ray Martinez for technical support in Tampa, Florida (US) and Robin Rycroft for the English version. This work was supported by grants from DGICYT to R.G. (Ministerio de Educación y Cultura, Spain, PB96–0220, BMC2000–0536 and UE98–0014), EU contract BIO4 CT97–2123, and a FPI fellowship to C.C. from the CIRIT (Generalitat de Catalunya, 1997FI00007).

References

Conchie J, Findlay J, Levvy GA (1958) Mammalian glycosidases: distribution in the body. J Biol Chem 195:239–249

Dimri PG, Lee X, Basile G, Acosta M, Scott G, Roskelley C, Medrano EE, Linskens M, Rubel I, Pereira-Smith O, Peacock M, Campisi J (1995) A biomarker that identifies senescent human cells in culture and aging skin in vivo. Proc Natl Acad Sci USA 92:9363–9367

Gossrau R, Lodja Z, Stoward PJ (1991) Glycosidases. In: Stoward PJ, Everson Pearse AG (eds) Histochemestry: theoretical and applied. Churchill Livingstone, Edinburgh, pp 241–279

Hinman VF, Degnan BM (1998) Retinoic acid disrupts anterior ectodermal and endodermal development in ascidian larvae and postlarvae. Dev Genes Evol 208:336–345

Holland ND, Holland LZ (1999) Amphioxus and the utility of molecular genetic data for hypothesizing body part homologies between distantly related animals. Am Zool 39:630–640

Holland LZ, Holland ND (1996) Expression of *AmphiHox-1* and *AmphiPax-1* in amphioxus embryos treated with retinoic acid: insights into evolution and patterning of the chordate nerve cord and pharynx. Development 122:1829–1838

Kraft JC, Willhite CC, Juchau MR (1994) Embryogenesis in cultured whole rat embryos after combined exposures to 3,3′,5-triiodo-L-thyonine (T3) plus all-*trans*-retinoic acid and to T3 plus 9-*cis*-retinoic acid. J Craniofac Genet Dev Biol 14:75–86

Lee YM, Osumi-Yamashita N, Ninomiya Y, Moon CK, Eriksson U, Eto K (1995) Retinoic acid stage-dependently alters the migration pattern and identity of hindbrain neural crest cells. Development 121:825–837

Lodja Z (1970) Indiogenic methods for glycolisidases. II. An improved method for β-D-galactosidase and its applications to localization studies of the enzymes in the intestine and in other tissues. Histochemie 23:266–288

Pearson B, Wolf PL, Vazquez J (1963) A comparative study of a series of new indolyl compounds to localize β-galactosidase in tissues. Lab Invest 12:1249–1259

Schnetzer JW, Tyler MS (1996) Endogenous beta-galactosidase activity in the larval, pupal, and adult stages of the fruit fly, *Drosophila melanosgaster*, indicates need for caution in *lac-Z* fusion-gene studies. Biol Bull 190:173–187

Weiss DJ, Liggitt D, Clark JG (1999) Histochemical discrimination of endogenous mammalian beta-galactosidase activity from that resulting from *lac-Z* gene expression. Histochem J 31:231–236

CAPÍTOL V

En aquest capítol, 1) es caracteritzen les regions intragèniques de les ***Adh3*** de ***Branchiostoma lanceolatum*** i ***Branchiostoma floridae***. 2) Es revela per anàlisi Southern l'elevat grau de polimorfisme d'aquests locus. 3) La caracterització de les seqüències intròniques demostra l'existència d'abundat DNA repetit. 4) Es descriu un nou tipus de minisatèl·lits, els ***miratges***, que es caracteritzen per incloure en les seves subunitats de repetició part dels extrems dels exons i les senyals donadores i acceptores de ***splicing***. 5) La detecció de la creació de noves variants al·lèliques d'origen somàtic i l'aï llament d'al·lels recombinants, demostren que l'amfiox és un organisme mosaic. 6) La caracterització dels al·lels recombinants suggereix que l'entrecreuament desigual inter-al·lèlic podria ser el mecanisme generador del polimorfisme genètic en aquest organisme. 7) Finalment, es descriu el primer element mòbil no autònom, l'***AmphiEXE***, en els amfioxs.

Minisatellite instability at the *Adh* locus reveals somatic polymorphism in amphioxus

Cristian Cañestro, Roser Gonzàlez-Duarte and Ricard Albalat*

Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal, 645, 08028-Barcelona. Spain

ABSTRACT

Minisatellite instability causes genetic polymorphism in vertebrates. Aberrant meiotic crossovers are involved in the germline polymorphism, whereas DNA polymerase slippage and unequal sister chromatid exchange is thought to account for the less frequent somatic variability. A high level of polymorphism has been reported in amphioxus (subphylum Cephalochordata) but, neither the molecular basis nor the nature of the polymorphic sequences is known. Here we describe a high level of polymorphism related to repetitive DNA sequences (minisatellites) in the Alcohol dehydrogenase (*Adh*) gene in two amphioxus species (*Branchiostoma lanceolatum* and *Branchiostoma floridae*). In contrast to vertebrates, amphioxus repeat instability was not preferentially linked to germline processes but appeared during the animal lifetime causing somatic polymorphism, probably because of inter-allelic crossovers. Furthermore, most *Adh* minisatellites belonged to a novel class of minisatellites whose distinct feature was that the repeat subunit spanned the exon-intron boundaries and generated splice site duplications. However, function was not compromised as no aberrant mRNA variants were detected.

INTRODUCTION

Vertebrate genomes contain a large variety of repeated DNA elements, which constitute one of the main driving forces of genome evolution. Among these repeats, minisatellites are intermediate size tandem arrays made up of repeat units of 6-100 bp. The informativeness of these repeats in human is of special interest for DNA typing, genetic mapping and for the study of genetic diversity and population structuring. However, there is still scarce information about the instability processes that generate and maintain minisatellite variability in natural populations. Repeat dynamics have been studied through pedigree analysis (1) and Small Pool-PCR (SP-PCR) assays (2) to characterize new length variants. These approaches, mainly done in human and mouse, have been hampered because of the limited frequency of minisatellite mutations. These studies have shown that germline instability, which is strongly biased towards the male germline (1), frequently involves complex inter-allelic conversion-like events at meiosis (3). In contrast, somatic instability occurs by other mechanisms that involve polymerase slippage or, more probably, intra-allelic unequal crossover (3).

The cephalochordate amphioxus, the closest living relative to vertebrates, has frequently been considered the ideal outgroup for phylogenetic reconstructions and developmental genetic analyses of vertebrate gene families (4, 5). In recent years, several amphioxus genes have been reported to display inter-individual variability (6-18), indicative of a high level of allelic polymorphism. Here we analyzed one of these genes, the Alcohol dehydrogenase class 3, in two amphioxus species, *Branchiostoma floridae* (*BfAdh3*) and *Branchiostoma lanceolatum* (*BlAdh3*), and showed that the origin of this variability is the instability of repeated elements in the intragenic regions. The detection of somatic repeat instability and recombinant alleles in mosaic animals suggested that intra- or inter-allelic unequal exchange could be responsible for these phenomena. Moreover, we described a novel class of minisatellites, termed *miratges*, whose distinct feature was that the repeat

* Correspondence address: Ricard Albalat, Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, E 08028 Barcelona, Spain Phone number: +34.934021502 Fax number: +34.934110969 E-mail: albalat@bio.ub.es

subunit spanned the exon/intron boundaries. Therefore, because of the extensive polymorphism found in natural populations, amphioxus could be a valuable model to monitor repeat dynamics, and provide an evolutionary perspective of the mechanisms involved in repeat instability.

MATERIAL AND METHODS

Genomic DNA and Southern analysis

Adult *Branchiostoma lanceolatum* were kindly provided by the Laboratoire Arago (Observatoire Océanologique de Banyuls-sur-mer, France) and adult *Branchiostoma floridae* were harvested in Tampa Bay, Florida (USA). Animals were kept at -70°C until used. Genomic DNA was isolated using the guanidine thiocyanate method (19) with minor modifications (20). For Southern analysis, 7-10 µg *Pst*I digested genomic DNA from single animals were resolved in agarose gels, transferred to nylon membranes and hybridized with six distinct probes under high stringency conditions (16).

Detection of intron length variation by SP-PCR

Somatic and gonadal-enriched tissue was dissected from four *B. lanceolatum* males. Purified genomic DNA from each sample was kept separately. To detect minisatellite variability in each sample, the conventional SP-PCR (2), with minor modifications, was used. DNA (700 ng) was fully solubilized by digestion with 10 U *Bam*HI (1 h at 37°C in 4 mM spermidine trichloride), centrifuged at 10,000 g for 2 min to remove particulate matter, and adjusted to a final concentration of 4 ng/µl in 10 mM Tris-HCl pH 8.0/ 1 mM EDTA. Around 2,280 amphioxus haploid genomes, assuming a molecular mass of 1 pg per haploid genome, were diluted into 200 µl of PCR-mix: 1 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 µM of each forward (F) and reverse (R) primer (F_1 + R_2 and F_6 + R_7 for intron 1 and 6, respectively, Table 1), 1xPCR-buffer and 0.1 U/µl of Biotherm™ (Genecraft, Germany) Taq polymerase. Multiple 7 µl aliquots (80 haploid genomes) of this solution were SP-PCR amplified on a GeneAmp PCR system 9600 thermal cycler (Perkin Elmer) at 94°C (20 s), 55°C (1 min), 72°C (1 min) for 26 cycles, followed by a chase at 72°C (5 min). The PCR products were resolved in a 1.7% agarose gel and detected by Southern blot hybridization with a ^{32}P-labeled probe encompassing the analyzed region. To compare the frequency of the new forms between different introns, somatic and gonadal-enriched DNA, homozygous and heterozygous animals and individuals, a statistical analysis of the homogeneity of the data was assessed with a chi-squared test. Additional PCR experiments were performed to discard artifactual amplifications. As mutant-free control DNA, two intron 6 cloned variants (L, large, and S, short, 679 and 646 bp from primer F_6 to R_7, respectively) were amplified using the SP-PCR conditions described. DNA from single clones or from an L-S mixture were used as templates. To compare the frequency of the new forms that appeared with genomic DNA *vs.* mutant-free DNA in SP-PCR assays, a contrast of proportions was performed with the Fisher's exact test.

Detection of recombinant alleles

A PCR assay was devised to detect recombinant variants from the L and S alleles of intron 6, taking advantage of a polymorphic *Hae*III restriction site. After exhaustive *Hae*III digestion of the genomic DNA of a heterozygous individual, PCR was performed with the F_6 and R_{7L} primers (Table 1 and Fig. 1). Since R_{7L} is primer-specific for the L allele, and the *Hae*III site is only present in the L variant, only S-L recombinant alleles are subjected to PCR amplification. Thirty ng of genomic DNA from a single L/S heterozygous animal were exhaustively digested with 24 U of *Hae*III at 37°C for 36 h (8 U dispensed every 12 h). The SP-PCR and hybridization conditions were as described above except for 200 haploid genomes per reaction and annealing at 58°C for 35 cycles without final extension.

To assess the extent of *Hae*III digestion, additional PCR reactions (100 haploid genomes each) were carried out on the same digested DNA with primers corresponding to a close genomic region (F_3 and R_5 from exon 3 and 5, respectively, Table 1 and Fig. 1). Since exon 4 contained a non-polymorphic *Hae*III site, only undigested molecules were amplified. To

minimize experimental variability, sample and controls were resolved on the same agarose gel and hybridized with a single ^{32}P-labeled probe, which overlapped 427 nt with F_3-R_5 and 400 nt with F_6-R_{7L} amplified fragments.

To assess the number of input amplifiable recombinant alleles per SP-PCR reaction, serial dilutions of a cloned L allele were mixed with 100 haploid genomes from a single animal which did not have the L-allele. Under the same experimental conditions as before, as few as 1 or 2 L allele molecules were amplified.

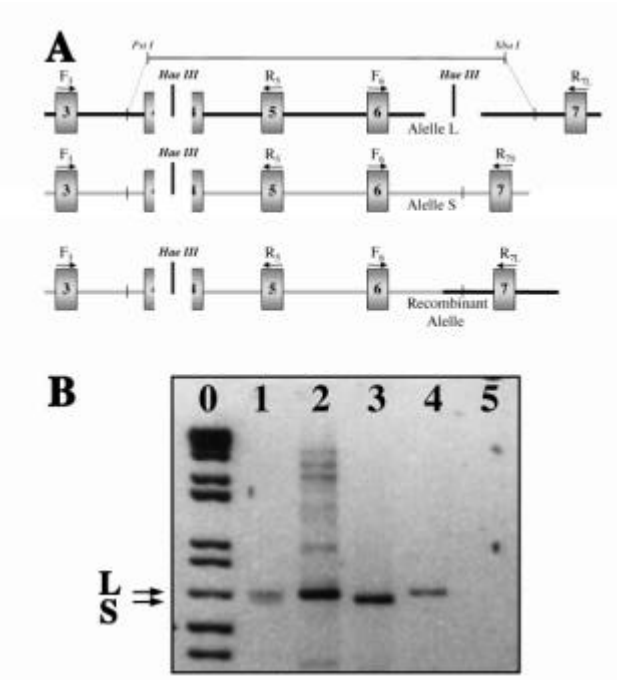


Figure 1. (A) Experimental procedure to characterize recombinant alleles. The structure of *BlAdh3* from exon 3 to 7 (boxes) of three allelic variants (L, S and a putative recombinant form) is illustrated after *Hae*III restriction (gaps). The *Pst*I-*Xba*I probe (a line over the genomic structure) to verify amplification with primers (arrows) F_3-R_5 and F_6-R_{7L} is shown. (B) Controls for PCR amplification were performed on genomic DNA of the L- and S-heterozygous animal (Fig. 3) and the cloned alleles. PCR products were resolved on a 1.7% agarose gel and stained with ethidium bromide. Lane 0: 1 kb ladder DNA marker (Promega); lane 1: co-amplification of the L- and S-cloned alleles with primers F_6-R_{7L}-R_{7S}; lane 2: amplification of the L-allele over undigested genomic DNA with F_6-R_{7L}; lane 3: amplification of the S-allele with F_6-R_{7S} over *Hae*III-restricted genomic DNA; lanes 4 and 5: amplification with F_6-R_{7L} over 10^3 and 10^2 L-allele cloned molecules, respectively.

We estimated the frequency of amplification of artifactual chimeric molecules, which were sporadically generated because incomplete chain elongation and subsequent priming on the heterologous allele along the co-amplification reaction (21). The L- and S-allele clones, oppositely oriented in the pUC18 vector, were amplified with the single M13-reverse primer. In this case, only artifactual chimeras were amplifiable.

Finally, primers F_6 and R_7 were used to amplify the *BlAdh3* intron 6 over genomic DNA from a single animal. PCR reactions were performed on a GeneAmp PCR system 9600 thermal cycler (Perkin Elmer) at 94 °C (20 s), 55 °C (1 min), 72 °C (1 min) for 26 or 40 cycles, followed by a chase at 72 °C (5 min). PCR products were cloned into *Sma*I-digested pUC18 vector and several clones were sequenced to identify recombinant alleles.

Sequence analysis

Accession numbers for the genomic DNA sequences used in this study were *BlAdh3*: AF156698, AF156700, AF156705, and *BfAdh3* AF266713-AF266719. For sequence comparisons with public databases, the WU-BLAST-X 2.0+BEAUTY and TBLASTN 2.1.2 search programs were used. The FGENES-M 1.5 program was used for gene prediction.

F_1	5'ACGGCACGCCAAAACCTC 3'
F_3	5'CCCACAGTGCAAGGAGTGC 3'
F_6	5'CATGGGGGCAAAGGTCGC 3'
R_2	5'GATGCGGACTTCGTGCGC 3'
R_5	5'CCTTAGCTGTGTTGAGAGCA 3'
R_7	5'CTGGATGGGCTTGTCATG 3'
R_{7L}	5'ACTCCGTCATGCCAAAC 3'
R_{7S}	5'ACTCCGTCATGCCGAAT 3'

Table 1. Primer sequences.

RESULTS

Amphioxus genomic variability

We had previously characterized the *Adh* genomic region of two amphioxus species (*B. lanceolatum* and *B. floridae*) and shown that *Adh* is a single-copy gene (16). Library screenings yielded 50 (*B.*

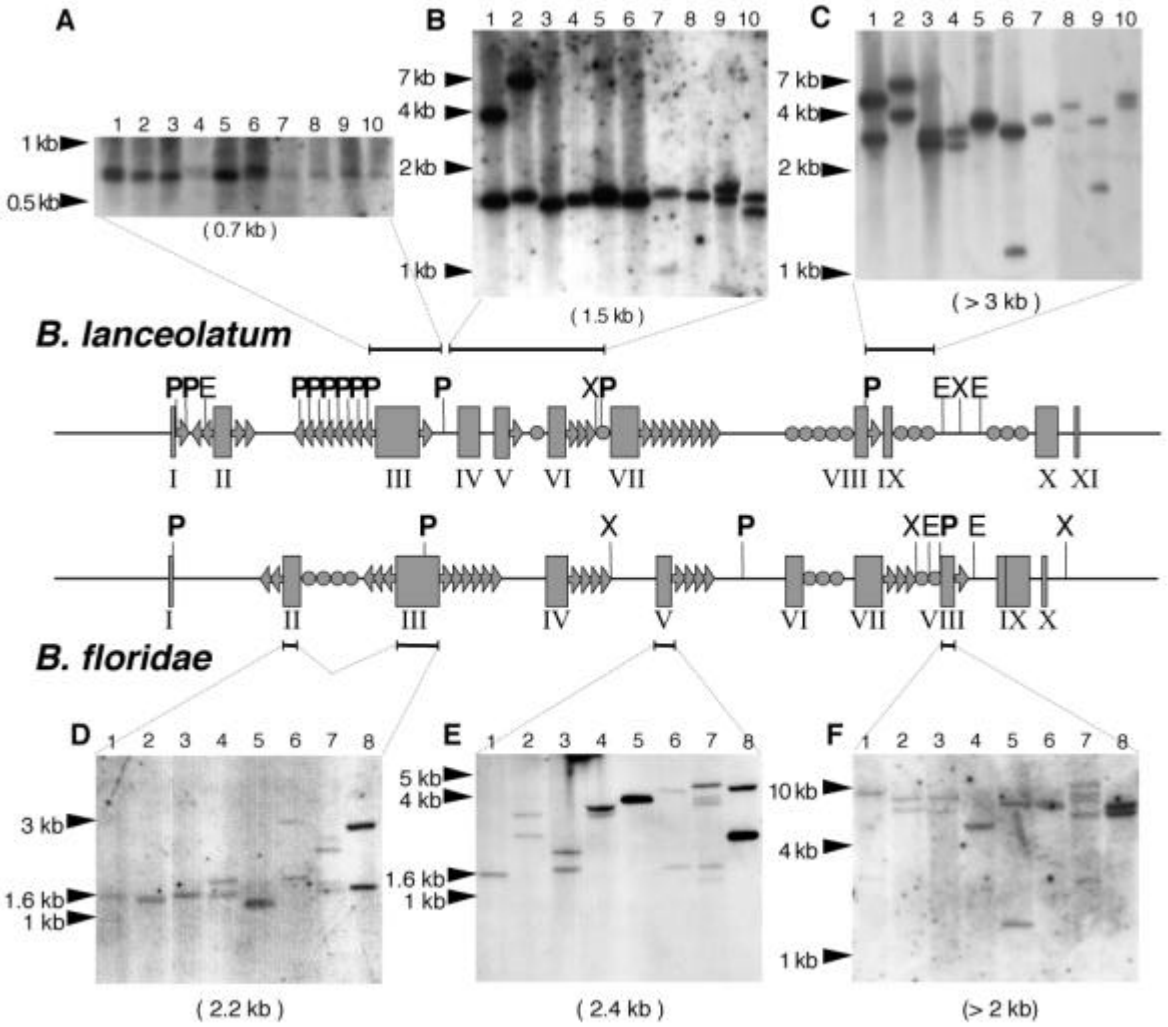


Figure 2. Southern analysis of *BlAdh3* and *BfAdh3*. Genomic DNA from single individuals was *Pst*I digested and probed with DNA fragments containing exon III (A), exons IV-VI (B), and exons VIII-IX (C) of *BlAdh3*, and exons II-III (D), exon V (E), and exon VIII (F) of *BfAdh3*. The expected band size from the characterized positive phage (under each panel) and the DNA size markers (at the left of each panel) are depicted. Genomic structure of *Adh3* (central part): exons (boxes); *Eco*RI (E) *Pst*I (P, in bold) and *Xba*I (X) restriction sites; repeated minisatellites, named *miratges*, (arrows); other minisatellites (circles) and probes (horizontal lines) are shown.

lanceolatum) and 12 (*B. floridae*) recombinant phages encompassing the same genomic region, whose restriction patterns revealed an unexpected high variability (data not shown). When *Adh* probes comprising various segments of the gene were used for Southern analysis on genomic DNA, none of the 18 animals shared an overall hybridization pattern (Fig. 2). Moreover, closer examination of the banding pattern revealed that this variability was not uniformly distributed across the *Adh* locus, ranging from a single band of similar size (region containing exon 3, Fig. 2A) to a multiplicity of bands of distinct sizes (probe containing exon 9, Fig. 2C). Remarkably, one *B. floridae* animal displayed more than two bands per lane (Fig. 2D-F, lane 7), which could suggest somatic polymorphism.

Structural characterization of the *Adh* repeats: "*Miratges*"

Sequence analysis of isolated phages revealed that the restriction length polymorphism shown by Southern was due to repetitive DNA in the *Adh* introns. Closer inspection of the intron sequences showed mainly direct tandem repeats (minisatellites, VNTRs) with,

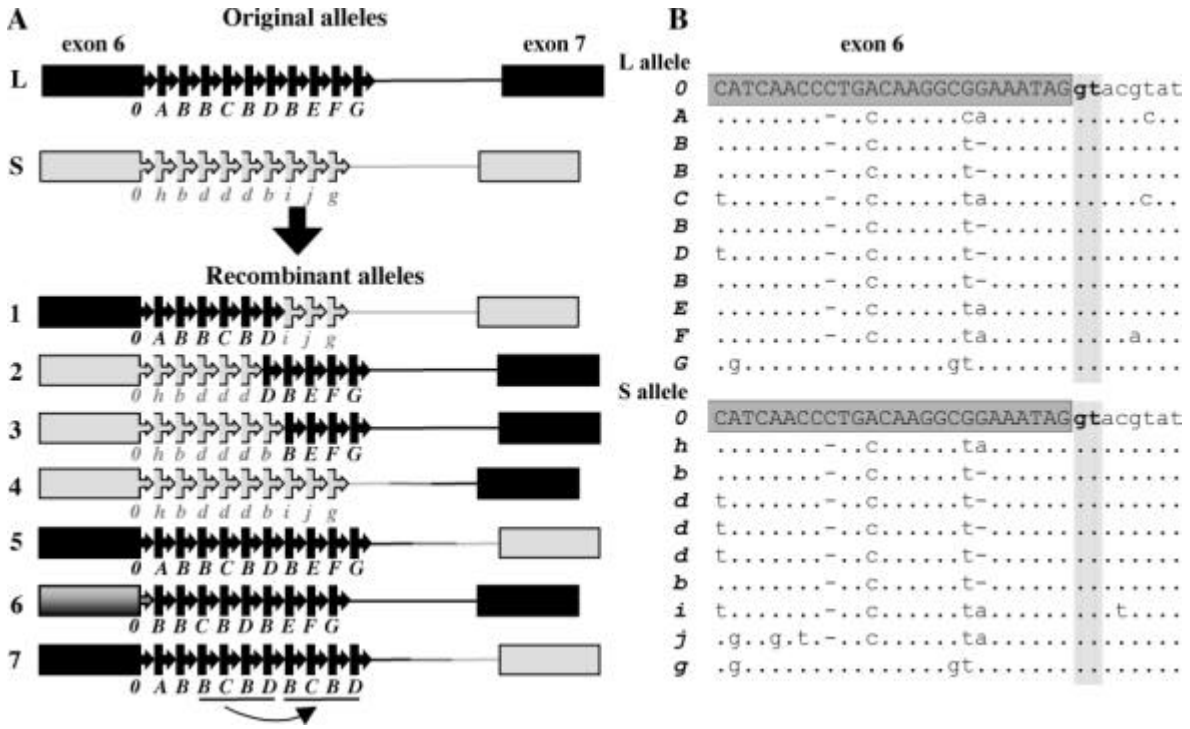


Figure 3. Structures of intron 6 *BlAdh3* original and recombinant alleles of an animal heterozygous for the L (black) and S (gray) alleles. L and S alleles contain 10 and 9 repeats, respectively and show multiple nucleotide substitutions. (A) Schematic representation of a GT-*miratge* overlapping the donor splice site at the boundary of exon 6-intron 6. Arrow-boxes represent repeat subunits where the boxed part corresponds to the sequence present in the exon. The first repeat subunit (numbered 0) is partly contained in the exon. Each following *miratge* subunit is shown in black capital and gray lower-case letters for the L- and S-alleles, respectively. The exon 6 of the recombinant allele 6 cannot be ascribed to any of the L- or S- original alleles. (B) Sequence alignment of the exon 6-intron 6 GT-*miratge* array of the L- and S- alleles. The segment corresponding to the coding region (capital letters over a dark gray box) and intron sequences (lower-case letters) are shown. The GT-splice site (bold low case) and its reiteration are depicted on a light-gray background. Nucleotide substitutions in the repeated sequences are shown, identities are represented by dots, and deletions by dashes.

occasion internal short tandem repeats (microsatellites) and one putative mobile element in intron 9 (Table 2). Remarkably, among the VNTRs a novel type of minisatellite was found whose distinct feature was that the repeat subunit spanned the exon-intron boundary (Fig. 3). We have named these novel minisatellites "*miratges*", the Catalan word for mirages, prompted by the fact that during sequence analysis, when a new exon was apparently reached, it unexpectedly vanished.

The *miratges* were from 10 to 72 nt imperfect direct tandem repeats, arranged in stretches of 2 to more than 15 subunits (Table 2). The fragment overlapping the coding region varied from 3 to 32 nt, while the non-coding sequence ranged from 5 to 46 nt. Fourteen of the 19 amphioxus *Adh* introns exhibited *miratges*, some of which spanned the donor splice site (GT-*miratges*) and were uniformly spread throughout the gene, whereas others (AG-*miratges*) were only found in introns 1 and 2 of both amphioxus species. Even though the putative splice sites (AG/GT) were conserved in 153 of the 157 (97.5%) repetitive subunits, no aberrant splicing variants were detected by RT-PCR (data not shown). Finally, although some tandem repeat sequences showed poor resemblance to the exon, they still maintained detectable homology ("cryptic" *miratges*).

Dynamics of the repeats

Instability of tandemly repeated DNA sequences has been frequently associated with germline-specific processes (1, 22). To elucidate the source of amphioxus repeat variability, DNA was prepared from male gonadal and somatic tissues from single

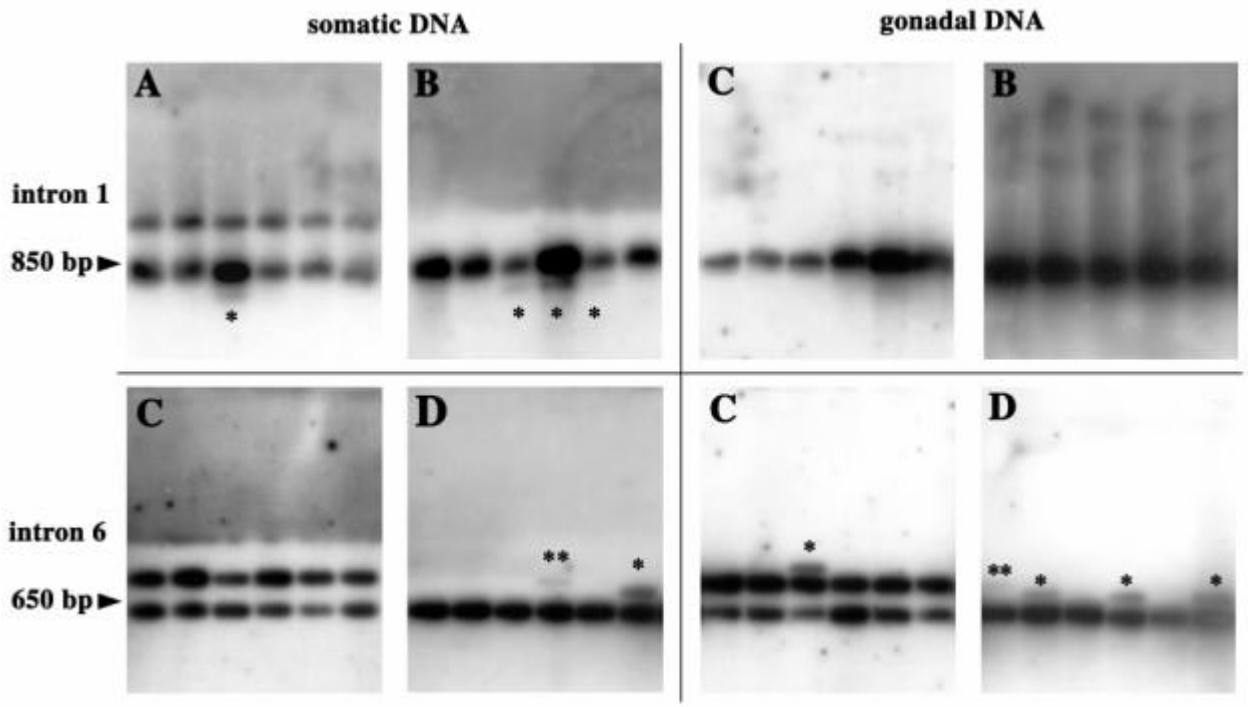


Figure 4. Detection of allelic length variants of *BlAdh3* in 4 single animals (A, B, C and D) by SP-PCR. Examples of size mutants corresponding to somatic (left) and gonadal-enriched DNA fractions (right), of intron 1 (top) and intron 6 (bottom). Animals were homozygous (B and C for intron 1; D for intron 6) or heterozygous (A for intron 1 and C for intron 6) with respect allele size. PCR products were electrophoretically resolved and abnormal length mutant molecules (* one subunit or ** two subunit variation) were detected by blotting and hybridization.

individuals. Gonadal preparations were not devoid of somatic tissue, but enriched with DNA from cells that had entered meiosis. SP-PCRs were performed with primers flanking two *BlAdh3 miratge*-containing introns, 1 and 6 (Table 2). Four animals, either homozygous or heterozygous for intron lengths, were analyzed. All the PCR reactions showed an intense common signal, which corresponded to the original alleles, and occasionally, some PCR products of abnormal length (Fig. 4).

Of 335 PCR amplifications a total of 40 new variants out were detected. Assuming the most conservative scenario in which each new variant arose from a single template molecule, this value represents 1 new variant per 670 haploid genomes (0.15%) (Table 3). The frequency of the new forms was not higher in gonadal-enriched DNA preparations than in the somatic fraction (0.10% and 0.19%, respectively) (χ^2 = 6.96; d.f. = 3; P-value = 0.07). Although gonadal preparations were not devoid of somatic tissue, if the repeat instability would be linked to germline processes, the proportion of new variants detected in the gonadal fraction should be higher than in the somatic one. Therefore, from our data, repeat instability was not associated with germline processes, but rather supported intron-length variability leading to somatic mosaicism. Moreover, the four individuals showed up to 10-fold differences, from 0.02% to 0.27% (χ^2 = 16.14; d.f. = 3; P-value = 0.001) in their level of somatic mosaicism. Nevertheless, no statistical differences were observed between homozygous or heterozygous animals (0.16% and 0.13%, respectively), nor between the two introns (0.13% for intron 1 and 0.16% for intron 6) (χ^2 = 4.37; d.f. = 3; P-value = 0.223). Interestingly, while the new variants of intron 6 increased in size, those from intron 1 decreased.

SP-PCR control experiments were performed with cloned DNA from two intron 6 alleles as mutant-free DNA (see Material and Methods) in order to check that no artifactual bands were generated in the assays. None of the PCR reactions (0/64, Fisher's exact test P-value = 0.001) produced a new variant of abnormal length, indicating that the new length variants were genuine mutated alleles.

Table 2. Repeat elements in *BlAdh3* and *BfAdh3* genes. *Miratges* (gray cells), ag/gt splice sites (bold) and the sequence corresponding to the coding (capital letters) or non-coding (lower-case letters) are indicated

	B. lanceolatum					***B. floridae***				
Intron	Repeat	Subunit sequence	Subunit length (Cod+Non Cod)	Number of repeated subunits	subunit GT/AG conservation	Repeat	Subunit sequence	Subunit length (Cod+NonCod)	Number of repeated subunits	subunit GT/AG conservation
1	GT-*mirage*	GGAAAG**gt**accgggg	15 (6+9)	1	1/1	Cryptic AG-*mirage*	ttgggaattaac**ag**CCCA TCTCGTG	25 (11+14)	>2	2/2
	AG-*mirage*	ttgtttgtttgtttac **ag**CCCATCACGTG	28 (10+18)	2-8	9/10					
2	GT-*mirage*	AAG**gt**gagtcgagctg ttatcacaaatagg	30 (3+27)	2	2/2	VNTR	GACATGACTTTGGAATGG GATCAAATTACA	30	3	-
						VNTR	GGTGCTGGCTACAGGAGT CTGCCACACACAAAGCGC TAACATTTTCCCGTTGTA TCGTCA	60	>10	-
	AG-*mirage*	ggaatgccacacagca gctggtttaaacggta cattttcctgc**ag**GT ACTGGCTACAGGAGTC TGCCACAC	72 (26+46)	>10	10/10	AG-*mirage*	tctaaacgctacattttc ccgttgtatcggc**ag**GTG CTGGCTACAGGAGTGTGC CACACG	60 (27+33)	>7	7/7
3	Cryptic GT-*mirage*	AACCAACCTGTGCCAG AAAATCAG**gt**aagaac	33 (24+9)	1	1/1	GT-*mirage*	AGAAAATCAG**gt**gagaac cgatctacctgtgtt	33 (10+23)	>15	15/15
4	Microsatellite	TA	2	5	-	GT-*mirage*	CAAG**gt**cagg	10 (4+6)	12	10/12
5	GT-*mirage*	TCTCAACACAGCTAAG **gt**tag	21 (16+5)	1	1/1	GT-*mirage*	CTGGCTATGGAGCTGCTC TCAACACAGCTAAG**gt**a gt	38 (32+6)	>15	15/15
	Repeated motif	GCATGCTGGGT	11	3	-					
6	GT-*mirage*	CATCAACCCTGACAAG GCGGAAATAG**gt**acgt at	34 (26+8)	8-13	29/30	VNTR	CTGTGTGTGTGTGTTTCT GTGTGTTTGTGTTT	32	7	-
	Repeated motif	CCTGTGTGATAA	12	2	-	microsatellite	GT	2	>100	-
7	GT-*mirage*	GGCAACGTGCACACCA TG**gt**gagtaaccat	30 (18+12)	6-17	34/34	GT-*mirage*	CACTATG**gt**gagtaacct gacac	23 (7+16)	>14	14/14
	VNTR	TGACTGTCCCAGAGTT AAACCCAGTTTTG	29	>13	-	VNTR	TTGAAAGAATTCTTGTTC TCCTGTT	25	3	-
8	GT-*mirage*	GCAGCAGGGCAGG**gt**a agtccattttatgtca gggtgggg	40 (13+27)	2	2/2	Cryptic GT-*mirage*	CAGG**gt**gggggt	12 (4+8)	2	2/2
9	VNTR	CCCCTGTATAACCTGT GCTGTA	22	13+*AmphiEXE* +5	-	-	-	-	-	-
	Cryptic VNTR	ATGTATAACC	10	>17	-					

-73-

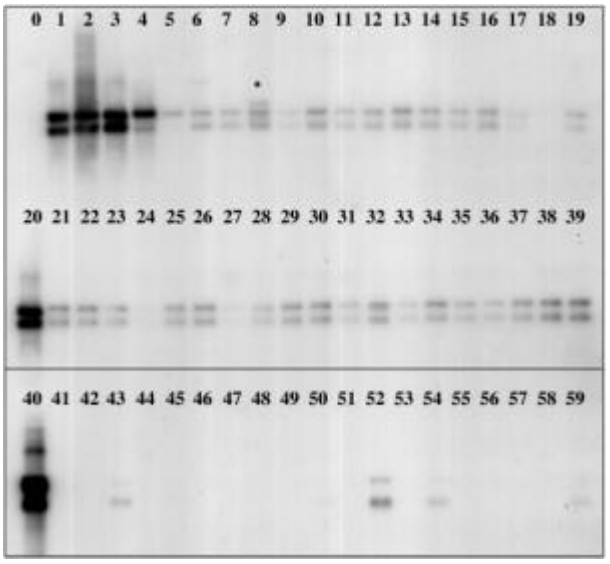


Figure 5. Recombinant alleles of *BlAdh3*. Genomic DNA from an S- and L-allele heterozygous animal was *Hae*III digested and recombinant alleles were selectively amplified with the F_6 and R_{7L} primers, and detected with the *Pst*I-*Xba*I probe (Fig. 1A) (lanes 6-19 and 21-39). Control PCR (lanes 1- 5) correspond to the samples shown in Fig. 1B after blotting and hybridization with the same probe. Notice that the control assays over cloned variants yielded a single band on ethidium bromide stained gels (Fig. 1B) but produced a double banded pattern after Southern analysis (lanes 1-5 present figure), likely due to a running effect of single-stranded DNA, poorly stained with ethidium bromide, generated after long PCR assays. In lane 8, an abnormal length recombinant is shown (*). Lanes 41-59. The extent of *Hae*III digestion on the initially restricted DNA was assayed by a second PCR with the F_3 and R_5 primers of *BfAdh3* (Fig. 1A) followed by blotting and hybridization with the *Pst*I-*Xba*I probe. Lane 40: Control

Detection of recombinant alleles.

To examine the mechanism underlying intron variability an experiment was performed to detect recombinant alleles in total genomic DNA. Allele-specific PCR amplification after genomic *Hae*III-digestion (Fig. 1) showed recombinant alleles in 37 out of 38 (97%) amplifications (Fig. 5) whereas only 27% of the digestion-control samples was amplified, which corresponded to non-digested templates. Artifactual chimeric molecules caused by jumping PCR were observed in less than 10% of the reactions. Overall, our data support recombinant *Adh* alleles in the amphioxus genome. In one PCR, an additional expanded band was detected, which could represent a recombinant allele whose size had been increased (Fig. 5, lane 8). Finally, to assess recombinant alleles, PCR amplified bands involving *Adh* length-polymorphic intron 6 from a single animal were cloned and sequenced. When several clones were analyzed, two major sequences corresponding to the original alleles (L and S of 679 and 646 bp, respectively) (Fig. 3) were detected. These sequences differed by a single GT-*miratge* (33 bp) and several single-nucleotide substitutions. Additionally, other clones were characterized whose structure could be explained as recombinant variants (Fig. 3A). Five may have arisen from a single crossing-over event in the repeated segment (1, 2 and 3) and two in the non-repeated fragment (4 and 5). Nevertheless, the origin of sequences 6 and 7 was not obvious. Variant 6 could have arisen either by deletion of the "A" *miratge* subunit in the L allele or by an unequal crossover between the S and L alleles. Finally, the structure of variant 7 was surprisingly complex, with intra-allelic duplications and deletions (BCBD) plus an inter-allelic recombination. Artifactual recombinant alleles were never found in control experiments (0/15) in which the PCR products obtained by PCR co-amplification of the two isolated L and S-alleles were cloned and sequenced.

AmphiEXE mobile element.

Other repetitive sequences were found in the *BlAdh3* genomic region that contributed to the genome polymorphisms observed in the amphioxus population. A 1.3 kb element was identified in intron 9 of *BlAdh3*, named *AmphiEXE*, after the two *Eco*RI and one *Xba*I restriction sites contained in the element following the order E-X-E (Fig. 2 and 6A). Supporting evidence of *AmphiEXE* being a mobile element were: i) its presence in the intron 9 of *BlAdh3* in only some of the isolated library phages, disrupting a tandem repeat array; ii) the pattern observed on Southern blots of genomic DNA restricted at external *Pst*I (multiple-band ladder typical from moderately repetitive DNA) or at the conserved internal *Eco*RI sites (870-bp strong single band) (Fig. 6B); and iii) the

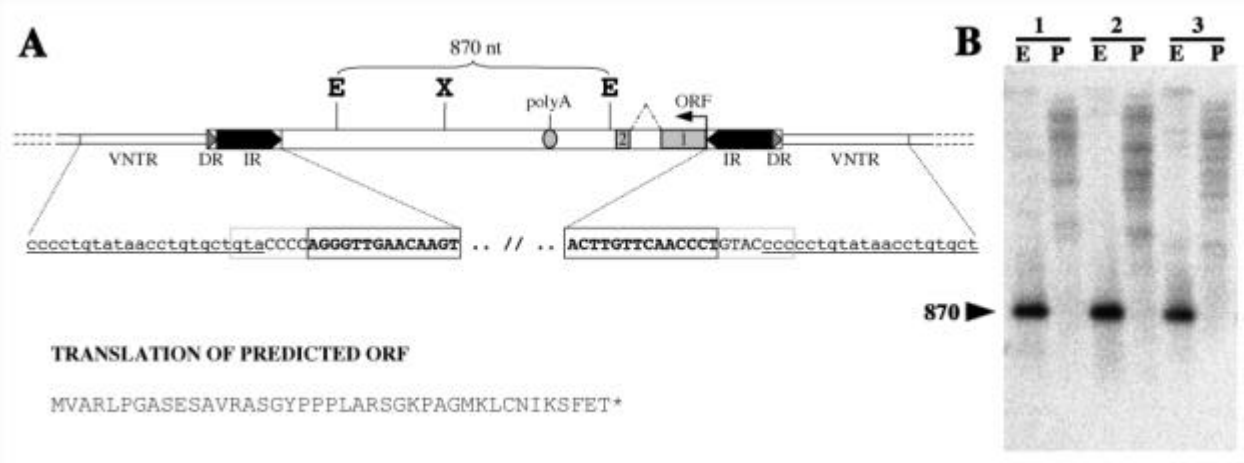


Figure 6. The *AmphiEXE* mobile element. (A) Structure of *AmphiEXE* located in intron 9 of *BlAdh3* (Fig. 2): direct target (DR) and inverted terminal repeats (IR), *Eco*RI and *Xba*I restriction sites, the predicted two-exon ORF (gray boxes), the transcription direction (arrow) and the polyadenylation site are depicted. The sequence of the 7-bp DR (gray-lined box), 14-bp IR (black-lined box) and the flaking VNTR sequences (underlined) are shown. The predicted 41 amino acid sequence is indicated. (B) Southern analysis of genomic DNA of 3 single animals (1, 2 and 3) restricted at the internal *Eco*RI (E) and the external *Pst*I sites (P), probed with the 526-bp *Eco*RI-*Xba*I fragment of *AmphiEXE*. Arrow indicates the predicted 870 bp

structural hallmarks characterized by sequence analysis, 14-bp inverted repeats (IR) at the boundaries of *AmphiEXE*, and the 7-bp direct repeats flanking this sequence. The FGENES program predicted an ORF of 41 amino acids split into two exons, and a putative polyadenylation signal . However, no significant homology with any reported sequence was found when *AmphiEXE* was compared with databanks. Our data suggest that the *AmphiEXE* is a highly conserved non-autonomous mobile element.

DISCUSSION

Amphioxus intragenic polymorphism.

The cephalochordate amphioxus is believed to predate the large-scale gene duplications that marked early vertebrate evolution and to this end, estimates of amphioxus gene number have been crucial. However, data on DNA variability at the coding regions and the observation of variable banding patterns after Southern analysis has led to the widespread assumption that the genome of lancelets is highly polymorphic (6 - 18). and this polymorphism can hamper the gene-copy number assessments.

Previous studies showed that *Adh3* was a single-copy gene in amphioxus (16). Our data on the Southern analysis of several individuals corroborate the single-copy status, although the inter-individual multiple banding pattern revealed that the lancelet *Adh3* is highly polymorphic. This high level of polymorphism produced complex Southern patterns, which were substantially simplified when reduced-size probes on single animal DNA were used (Fig. 2). Sequence analysis indicates that most of the differences in the banding patterns could be ascribed to intragenic repeated DNA such as VNTRs and mobile elements**.** Additional allelic variation could also be generated by small insertions/deletions and nucleotide substitutions. This overall polymorphism could be further increased during the animal lifetime, as repeat instability leads to mosaicism. Different individuals showed different degrees of mosaicism (Table 3), probably because of the generation rate of new forms, and the timing and cell target of the mutational event. Hence, if early in development or in a highly proliferative cell lineage, more than two alleles can be characterized from a single animal, and even are detectable as faint bands by Southern blot, as was the case for the *BfAdh3* (Fig. 2D-F, lane 7). Overall, to avoid misleading estimations, the high level of polymorphism together with the mosaic status of the amphioxus genome should be taken into account when evaluating gene copy number.

The amount of intragenic variability could be affected by factors such as chromosomal location, methylation status, level of transcription, proximity to telomeres,

Table 3. Abnormal length variants detected by SP-PCR assays

		Abnormal length variants/ SP-PCR reactions	Percentage of abnormal length variants[1]
Introns	I	14/134	0.13%
	VI	23/201	0.16%
Heterozygosity	Homozygotes	27/213	0.16%
	Heterozygotes	13/122	0.13%
Individuals	A	13/115	0.10%
	B	4/67	0.07%
	C	22/102	0.27%
	D	1/51	0.02%
DNAs	Somatic	28/189	0.19%
	Gonadal	12/146	0.10%
Total		40/335	0.15%

[1]The percentage has been calculated assuming 80 haploid genomes per SP-PCR reaction.

centromeres, and also to recombination hot-spots, which affect the organization of repetitive DNA in eukaryotic genomes (23 - 25). Hence, it can be assumed that not all amphioxus genes behave similarly.

Minisatellite instability

Previous evidence showing that the length variants detected by SP-PCR represent genuine mutant alleles rather than PCR artifacts (2) were confirmed in our study because: i) length variants were never detected under our SP-PCR conditions using cloned alleles as a mutant-free DNA control, and ii) the frequency of new length variants differed among animals, contrary to what would be expected if they were PCR artifacts.

SP-PCR analysis on genomic DNA from dissected single animals revealed that amphioxus repeat instability was related to somatic polymorphism rather than germline processes. We therefore conclude that amphioxus are mosaic animals. The detected length variants suggested a stepwise mutation fashion, compatible with one subunit variation in each step. Longer variants were only observed in the animal with the highest mosaicism level, in which subsequent steps would have taken place (Fig. 4, Table 3). Our data suggest that mutational bias (2) could drive the repeat dynamics because gains or losses did not occur with equal frequency: intron 6 exclusively showed expansions, whereas intron 1, only reductions.

The *Hae*III-SP-PCR assays together with the sequence analysis of cloned PCR amplified intron 6 showed recombinant alleles in the amphioxus genome. Most of the intron 6 cloned segments could be explained by single equal crossover events without altering in the repeat copy number, whereas in others unequal crossover and/or intra-allelic duplications and deletions have to be assumed. Therefore, although we cannot exclude replication slippage as the cause of minisatellite variability, the most likely mechanism for somatic instability is inter-allelic unequal exchange and less frequently intra-allelic or intramolecular recombination. Finally, the detection of one recombinant allele that increased in size provides the link between variation of the repeat number and the generation of recombinant alleles (Fig. 5, lane 8).

In summary, somatic polymorphism in amphioxus is supported by three independent approaches: i) Southern analysis, ii) SP-PCR experiments and iii) characterization of recombinant alleles. Our data validate cephalochordate amphioxus as a model with which monitor repeat dynamics in natural populations and to provide valuable data about the genome structure of the common ancestor of vertebrates.

"Miratges", a novel type of minisatellites.

We here describe a new type of minisatellite sequence, termed *miratge*, whose distinctive feature

with respect the conventional tandem repeats is that the repeat subunit spans the exon-intron boundaries and generates duplications of exon edges. *Miratges* are abundant in the *Adh* gene of the two amphioxus species (*B. lanceolatum* and *B. floridae*), 75% of the introns. However, they are probably not restricted to this genomic region, nor to this organism. In humans, repetitive elements with a similar structure have been reported in relation to genes associated with human disorders, in exon 4 of the platelet-derived growth factor A (*PDGFA*) gene (26), in exon 13 of the cystathionine β-synthase (*CBS*) gene (27) and in exon 7 of the factor VII (*F7*) gene (28).

Although none of the *miratge* subunit was identical to the exon-intron sequence from which it was generated, 97.5% of *miratge* subunits retained the AG/GT splice site. This similarity did not appear to compromise function, as no aberrant cDNA variants were detected. Concerning their origin, a link between the splicing machinery and the origin of *miratges* cannot be discarded. Moreover, their expansion could increase the similarity of the regions surrounding the coding exons, and affect gene conversion.

ACKNOWLEDGMENTS

We thank G. Marfany, A. Jeffreys and A. Carracedo for helpful discussion, and R. Rycroft for revising the English version of this manuscript. R.G. received a grant from DGICYT (Ministerio de Educación y Cultura, Spain, BMC2000-0536) and C.C. was the recipient of an FPI fellowship from the CIRIT (Generalitat de Catalunya, 1997FI00007).

REFERENCES

1. Jeffreys, A. J., Royle, N. J., Wilson, V. and Wong, Z. (1988) Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature*, **332,** 278-281.

2. Jeffreys, A. J., Tamaki, K., MacLeod, A., Monckton, D. G., Neil, D. L. and Armour, J. A. (1994) Complex gene conversion events in germline mutation at human minisatellites. *Nat.Genet.*, **6,** 136-145.

3. Jeffreys, A. J. and Neumann, R. (1997) Somatic mutation processes at a human minisatellite. *Hum. Mol. Genet.*, **6,** 129-136.

4. Garcia-Fernàndez, J. and Holland, P. W. (1994) Archetypal organization of the amphioxus Hox gene cluster. *Nature*, **370,** 563-566.

5. Holland, P. W., Garcia-Fernàndez, J., Williams, N. A. and Sidow, A. (1994) Gene duplications and the origins of vertebrate development. *Dev. Suppl.*, **43,** 125-133.

6. Holland, P. W., Koschorz, B., Holland, L. Z. and Herrmann, B. G. (1995) Conservation of Brachyury (T) genes in amphioxus and vertebrates: developmental and evolutionary implications. *Development*, **121,** 4283-4291.

7. Holland, L. Z., Kene, M., Williams, N. A. and Holland, N. D. (1997) Sequence and embryonic expression of the amphioxus engrailed gene (*AmphiEn*): the metameric pattern of transcription resembles that of its segment-polarity homolog in *Drosophila*. *Development*, **124,** 1723-1732.

8. Shimeld, S. M. (1997) Characterisation of amphioxus HNF-3 genes: conserved expression in the notochord and floor plate. *Dev. Biol.*, **183,** 74-85.

9. Tweedie, S., Charlton, J., Clark, V. and Bird, A. (1997) Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Mol. Cell. Biol.*, **17,** 1469-1475.

10. Karabinos, A. and Riemer, D. (1997) The single calmodulin gene of the cephalochordate *Branchiostoma*. *Gene*, **195,** 229-233.

11. Kusakabe, R., Kusakabe, T., Satoh, N., Holland, N. D. and Holland, L. Z. (1997) Differential gene expression and intracellular mRNA localization of amphioxus actin isoforms throughout development: Implications for conserved mechanisms of chordate development. *Dev. Genes Evol.*, **207,** 203-215.

12. Williams, N. A. and Holland, P. W. (1998) Gene and domain duplication in the chordate *Otx* gene family: insights from amphioxus *Otx*. *Mol. Biol. Evol.*, **15,** 600-607.

13. Kusakabe, R., Satoh, N., Holland, L. Z. and Kusakabe, T. (1999) Genomic organization and evolution of actin genes in the amphioxus *Branchiostoma belcheri* and *Branchiostoma floridae*. *Gene*, **227,** 1-10.

14. Shimeld, S. M. (1999) The evolution of the hedgehog gene family in chordates: insights from amphioxus hedgehog. *Dev. Genes Evol.*, **209,** 40-47.

15. Kozmik, Z., Holland, N. D., Kalousova, A., Paces, J., Schubert, M. and Holland, L. Z. (1999) Characterization of an amphioxus paired box gene, *AmphiPax2/5/8*: developmental expression patterns in optic support cells, nephridium, thyroid-like structures and pharyngeal gill slits, but not in the midbrain-hindbrain boundary region. *Development*, **126,** 1295-1304.

16. Cañestro, C., Hjelmqvist, L., Albalat, R., Garcia-Fernàndez, J., Gonzàlez-Duarte, R. and Jörnvall, H. (2000) Amphioxus alcohol dehydrogenase is a class 3 form of

single type and of structural conservation but with unique developmental expression. *Eur. J. Biochem.*, **267,** 6511-6518.
17. Yasui, K., Zhang, S., Uemura, M. and Saiga, H. (2000) Left-right asymmetric expression of *BbPtx*, a *Ptx*-related gene, in a lancelet species and the developmental left-sidedness in deuterostomes. *Development*, **127,** 187-195.
18. Dalfó, D., Cañestro, C., Albalat, R. and Gonzàlez-Duarte, R. (2001) Characterization of microsomal retinol dehydrogenase gene from amphioxus: retinoid metabolism before vertebrates. *Chem-Biol. Interact.*, **130-132,** 359-370.
19. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*, **18,** 5294-5299.
20. Garcia-Fernàndez, J., Baguñà, J. and Saló, E. (1993) Genomic organization and expression of the planarian homeobox genes *Dth-1* and *Dth-2*. *Development*, **118,** 241-253.
21. Jansen, R. and Ledley, F. D. (1990) Disruption of phase during PCR amplification and cloning of heterozygous target sequences. *Nucleic Acids Res.*, **18,** 5153-5156.
22. Buard, J., Shone, A. C. and Jeffreys, A. J. (2000) Meiotic recombination and flanking marker exchange at the highly unstable human minisatellite CEB1 (D2S90). *Am. J. Hum. Genet.*, **67,** 333-344.
23. Hancock, J. M. (1996) Simple sequences and the expanding genome. *Bioessays*, **18,** 421-425.
24. Charlesworth, B., Sniegowski, P. and Stephan, W. (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature*, **371,** 215-220.
25. Colot, V. and Rossignol, J. L. (1999) Eukaryotic DNA methylation as an evolutionary device. *Bioessays*, **21,** 402-411.
26. Bonthron, D. T., Smith, S. J. and Campbell, R. (1999) Complex patterns of intragenic polymorphism at the *PDGFA* locus. Hum. Genet., **105,** 452-459.
27. Yang, F., Hanson, N. Q., Schwichtenberg, K. and Tsai, M. Y. (2000) Variable number tandem repeat in exon/intron border of the cystathionine beta-synthase gene: a single nucleotide substitution in the second repeat prevents multiple alternate splicing. *Am. J. Med. Genet.*, **95,** 385-390.
28. Marchetti, G., Patracchini, P., Gemmati, D., DeRosa, V., Pinotti, M., Rodorigo, G., Casonato, A., Girolami, A. and Bernardi, F. (1992) Detection of two missense mutations and characterization of a repeat polymorphism in the factor VII gene (F7). *Hum. Genet.*, 89, 497-502.

DISCUSSIÓ

L'estudi de l'Evolució de les funcions: "l'EVO-FUN"

El constant increment en el nombre de gens caracteritzats i recollits en les bases de dades i la imminent seqüenciació d'un elevat nombre de genomes, anuncia la fi del període dels "caçadors de gens" per obrir les portes a l'era de "la recerca de funcions". Malgrat les comparacions entre organismes afavoreix la comprensió de les diferències genòmiques responsables de la gran diversitat biològica actual, i s'accepta que l'increment en el nombre de gens ha estat un dels motors de les grans transicions en les formes de vida durant l'evolució (Holland, 1998), encara es desconeix com els gens duplicats han anant adquirint noves funcions i com aquestes han afavorit la complexa organització dels éssers vius. L'estudi comparatiu de les funcions dels membres paràlegs d'una família gènica amb els ortòlegs de diferents organismes, permet arribar a establir hipòtesis sobre l'evolució funcional dels gens. Per fer aquest tipus d'estudi, és necessari partir de camps molt diversos: genètica, biologia molecular, bioquímica, zoologia, embriologia, fisiologia, evolució, filogènia i bioinformàtica. El caràcter multidisciplinari d'aquest tipus de recerca fa que sigui difícil de catalogar. L'EVO-DEVO, o estudi de les estratègies de desenvolupament per entendre l'evolució, seria l'enfoc multidisciplinari més proper. Però quan la investigació no està directament relacionada amb el desenvolupament en sentit estricte, sinó que fa referència a qualsevol procés biològic, s'hauria de parlar de l'Evolució de les Funcions (EVO-FUN) com una disciplina que en sentit ampli englobaria qualsevol tipus d'anàlisi dirigit a entendre l'Evolució a partir de l'estudi de les adquisicions i/o modificacions de les funcions dels gens al llarg de la història de la vida.

Estudiar la funció d'un gen és una tasca àrdua i difícil, que la major part de vegades passa per la generació de mutants i l'anàlisi dels seus fenotips. Aquests tipus d'investigacions de vegades són infructuoses, ja sigui degut a la manca d'efectes aparents dels mutants, o bé per la impossibilitat de generar-los. A diferència de molts gens reguladors del desenvolupament, que sovint presenten una elevada variabilitat i dominis conservats molt restringits, l'anàlisi de famílies d'enzims, normalment molt més conservades, ofereix l'avantatge de poder combinar els estudis dels patrons d'expressió amb la quantificació de l'activitat catalítica, per discernir algunes de les possibles funcions biològiques. La família de les ADHs ha esdevingut al llarg d'aquest treball un exemple paradigmàtic per fer aquest tipus d'estudi.

1. L'Evo-Fun a la família de les ADHs als cordats

L'expansió de la família de les ADHs va tenir lloc fa uns 500 MA, després de la separació cefalocordats-vertebrats (700 MA), probablement després de la separació dels gnatòstoms (550 MA) però abans de l'aparició dels tetràpodes (400 MA). Es creu que en aquest període hauria tingut lloc una "explosió d'isoformes" que hauria generat nous gens com a fruit de nombroses duplicacions a petita escala abans de l'aparició dels àgnats (Suga *et al.,* 1999). Així doncs, cap de les dues rondes de duplicacions genòmiques haurien estat associades a l'aparició de membres funcionals de la família de les ADHs. En aquest sentit les innovacions funcionals lligades a l'expansió de la família de les ADHs no haurien contribuït de forma directa a l'aparició dels vertebrats, sinó a la seva radiació abans de l'aparició dels tetràpodes.

Un cop originades les noves còpies, les forces evolutives varen actuar a dos nivells per generar noves funcions: (*i*) a nivell de la regió codificant del gen, modificant les propietats físico-químiques de la proteïna codificada, i (*ii*) a nivell de la regulació gènica, alterant el lloc i moment de l'expressió.

1.1 Evolució de la regió codificant

1.1.1 Quina era l'activitat ADH ancestral abans de l'expansió?

El fet que l'activitat formaldehid deshidrogenasa dependent de glutatió sigui l'única detectada en els organismes prevertebrats suggereix fermament que abans de l'aparició dels vertebrats, l'ADH ancestral precursora de les altres classes, ja presentava l'activitat d'una ADH de classe 3. Per aquest motiu, i malgrat filogenèticament el gen de l'*Adh* dels procordats s'hauria de denominar *Adh1/2/3/4/5/6/7* o simplement *Adh*, doncs és el pro-

ortòleg de tota la família d'*Adh*s a vertebrats (Sharman, 1999), seguint la nomenclatura proposada (Duester *et al.,* 1999) s'ha anomenat *Adh3* (*BfAdh3*, *BlAdh3* i *CiAdh3*), ja que codifica de forma inequívoca per una ADH amb l'activitat de la classe 3 dels vertebrats.

1.1.2 Processos evolutius implicats en el destí dels gens duplicats

Donada la manca de multifuncionalitat de l'ADH en els procordats s'ha de descartar la subfuncionalització com el procés evolutiu pel qual s'han generat les diferents classes d'ADHs en els vertebrats. Les noves funcions, enteses com a noves activitats enzimàtiques, provindrien d'un procés de neofuncionalització. Per mutagènesi dirigida s'ha demostrat que amb un nombre reduït de substitucions aminoacídiques l'ADH3 adquireix característiques típiques de les altres classes, com per exemple activitat front l'etanol, o es torna sensible a inhibidors específics de la classe 1 (Estonius *et al.,* 1994, Hedberg *et al.,* 1998).

Aquestes observacions reforcen la idea que les noves activitats, entre d'altres l'oxidació del retinol per l'ADH4, haurien aparegut per neofuncionalització durant l'evolució dels vertebrats. En aquest marc l'amfiox, sense l'elevada redundància gènica observada a vertebrats, és un organisme model ideal per estudiar el metabolisme del retinol en la via de síntesi de l'àcid retinoic. La recent caracterització de la retinol deshidrogenasa de cefalocordats apunta a aquest membre de la superfamília de les SDR com un possible protagonista en aquesta via metabòlica en prevertebrats (Dalfó *et al.,* 2001), i futures investigacions permetran dilucidar quines innovacions ha aportat la incorporació de l'ADH4 a aquesta via.

Un fet interessant que reforçaria el procés de neofuncionalització per l'evolució de les ADHs, és la descripció en els peixos d'una ADH de classe 1 amb característiques mixtes entre les classes 3 i la classe 1 de mamífers (Danielsson *et al.,* 1992), que podria representar una forma transitòria en l'evolució de les classes. No obstant, al considerar aquesta hipòtesi, cal tenir en compte que en el llinatge dels peixos s'han donat més duplicacions genòmiques que a la resta de vertebrats (Aparicio, 2000), i per tant poden tenir classes d'ADHs sense paràlegs a mamífers. Tant sols una anàlisi bioquímica i genòmica exhaustiva resoldrà si l'ADH1 de peixos és un veritable ortòleg a l'ADH1 de mamífers, o bé si es tracta d'un trans-homòleg amb convergència funcional.

1.1.3 La "densitat funcional" com eina per inferir l'Evo-Fun

Un concepte a tenir en compte en estudiar el procés de l'evolució de les noves funcions és la "densitat funcional", introduïda de forma teòrica per Zuckerkandl (1976) per quantificar la fracció d'una proteïna involucrada en la funció. Una elevada densitat funcional determinaria que tan sols una reduïda fracció d'aminoàcids podrien variar sense comprometre la seva funció, establint-se una relació inversament proporcional entre densitat funcional i taxa evolutiva. A més, si la funció d'una proteïna no varia entre diferents organismes, la densitat funcional, i per tant la taxa evolutiva, romanen constants. En un procés de duplicació gènica, si els dos gens seguissin un procés de degeneració i subfuncionalització, les densitats funcionals es veurien alterades produint canvis en les seves taxes evolutives. Per contra, si una de les còpies retingués la funció ancestral i l'altre n'adquirís una de nova, la taxa evolutiva de la primera romandria constant i la segona la modificaria. El fet que la taxa d'evolució de l'ADH3 s'hagi mantingut constant dins dels cordats, mentre que altres membres de la família (ADH1) hagin incrementat quasi 4 vegades la seva taxa d'evolució, aporta una base real a les prediccions teòriques de Zuckerkandl i obre la possibilitat de que a partir d'un paràmetre mesurable com la taxa d'evolució, es pugui inferir el procés evolutiu seguit, neofuncionalització o subfuncionalització, en l'adquisició de les noves funcions.

1.1.4 Evo-Fun de les proteïnes que formen homodímers

Les proteïnes que actuen com a homodímers, com és el cas de les ADHs, estan sotmeses a restriccions funcionals addicionals que determinen el seu procés evolutiu. Així, malgrat una de les còpies pot preservar la funció ancestral, l'altra no és lliure d'evolucionar mentre pugui interferir formant heterodímers no funcionals. Aquesta interferència determina que la nova còpia no obté complerta *llibertat mutacional* fins que assoleixi o bé la seva *independència estructural* per mutacions que impedeixin l'heterodimerització , o bé la seva *independència espaial* mitjançant canvis en el seus patrons d'expressió que condueixin a nivells de solapament tolerables o inexistents. Aquests dos processos poden donar-se per separat o de forma simultània, però mentre no s'assoleixi la *independència*, la pressió de la selecció mantindrà les dues còpies amb redundància funcional. Així doncs, després d'una duplicació d'una proteïna que dimeritza hi hauria una fase inicial on la selecció mantindria una redundància funcional frenant la desfuncionalització d'una de les còpies. En conseqüència, s'allargaria el temps en el que es poden adquirir noves funcions per neofuncionalització directe o a través d'un procés de refinament funcional i posterior subfuncionalització (fig. 1).

Totes les classes d'ADHs analitzades presenten *independència estructural*. No poden heterodimeritzar, degut a la redistribució de

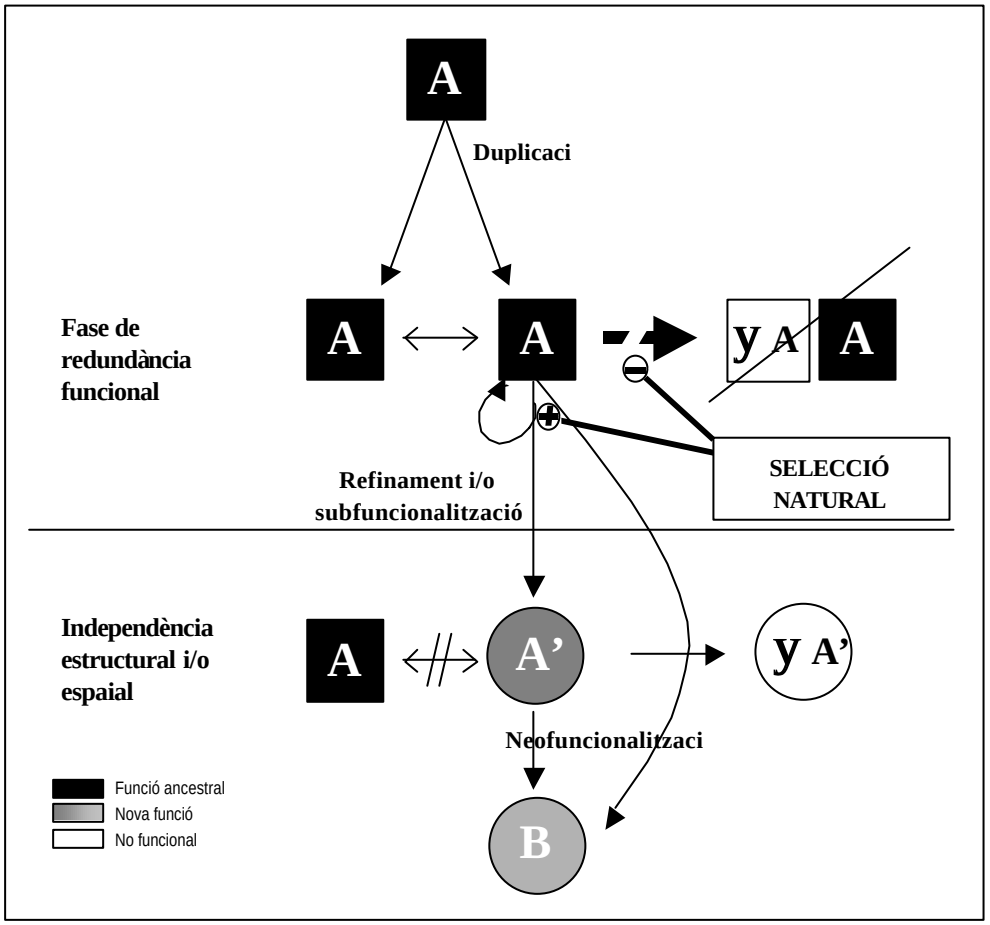


Figura 1. Model d'evolució funcional per gens duplicats que codifiquen per proteïnes que dimeritzen. En la part superior es representa la fase de redundància funcional mantinguda per selecció, degut a l'efecte deleteri de la formació d'heterodímers no funcionals. En la part inferior es mostra la fase d'independència estructural i/o espaial on el nou gen ha assolit la llibertat mutacional i podria adquirir noves funcions per neofuncionalització directe (B) o per refinament i subfuncionalització (A'). ψ: pseudogèn. ↔: dimerització, ⇿: no dimerització

residus i algunes substitucions en posicions d'interacció entre els monòmers (Danielsson *et al.,* 1994b) (Danielsson *et al.,* 1994a) (Cañestro *et al.,* 2000). Això torna a reflectir una redistribució de la densitat funcional com un pas clau en l'adquisició de noves funcions. D'altra banda, el canvi de patró d'expressió de l'*Adh3*, probablement lligat a les duplicacions gèniques que expandiren la família de les *Adh*s, pot haver influït en el seu destí. Malgrat no es coneix com s'elimina el formaldehid en organismes prevertebrats, a partir de l'expressió de l'*Adh3* es podria pensar que té lloc majoritàriament en teixits digestius. Així doncs, una possible *independència espaial* podria haver-se produït quan en els vertebrats l'*Adh3* va perdre l'especificitat tissular i passà a expressar-se en molts teixits on no s'expressaven els seus duplicats. Els efectes deleteris de la formació d'heterodímers no funcionals haurien estat menors i les còpies gèniques haurien pogut evolucionar de forma més independent.

1.2 Evolució del patró d'expressió dels membres de la família ADH

El segon nivell sobre el qual l'evolució pot actuar per determinar el destí funcional dels gens duplicats és la regulació de l'expressió. L'*Adh3* a prevertebrats (*Drosophila*, ascidis i amfioxs) presenta una expressió restringida a determinats teixits, els quals directa o indirectament tenen relació amb funcions digestives. Per contra, els vertebrats (peix zebra, ratolí i home) presenten un patró d'expressió generalitzat (fig. 2). En base a aquestes dades, sembla probable que el patró específic correspongués a l'estat ancestral, i que l'expressió ubiqua dels vertebrats fos una innovació d'aquest llinatge, produïda per guanys o pèrdues d'elements reguladors en les regions de control.

És interessant destacar la coincidència entre els patrons d'expressió de l'*Adh3* a organismes prevertebrats amb els d'altres membres de la família de les *Adh*s de vertebrats, com l'*Adh*4 a l'estómac (Moreno & Parés, 1991, Moreno *et al.,* 1996) i l'*Adh*1 al fetge i l'intestí (Smith *et al.,* 1971). Aquests resultats plantegen un escenari en el que després de les duplicacions gèniques una còpia hauria preservat les seves propietats bioquímiques, però hauria perdut l'especificitat tissular, mentre que les altres haurien adquirit la capacitat de metabolitzar nous substrats (etanol, retinol, hormones esteroidees) mantenint total o parcialment els elements reguladors de l'especificitat tissular (fig. 3). Aquests resultats ens han fet iniciar una nova línia d'investigació per estudiar l'Evo-Fun de les regions promotores de diferents ADHs a animals models situats filogenèticament abans i després de l'expansió (*Drosophila*, ascidi, amfiox i peix zebra). Anàlisis de predicció d'elements reguladors sobre les regions promotores de l'*Adh3* de prevertebrats, suggereixen l'existència de caixes reconegudes per factors de transcripció com C/EBPs i AEF-1, que a mamífers regulen l'expressió de molts gens en el teixit hepàtic. A més, experiments de retard en gel i protecció a la DNAsa, mostren com les C/EBPs que regulen l'expressió hepàtica de l'ADH1 murina, poden també reconèixer els llocs d'unió de les C/EBPs de l'Adh-SDR a cos gras de *Drosophila*. Aquest *puzzle* és un clar exemple de *bricolatge* molecular, on les mateixes peces són utilitzades per gens diferents que han convergit a nivell de funció (ADH1-MDR i ADH-SDR), o per organismes diferents que han d'expressar els seus gens en òrgans anàlegs i funcionalment equivalents (cos gras i fetge).

Aquestes dades podrien fins i tot contribuir a aportar nova llum sobre l'origen evolutiu del fetge: a l'amfiox existeix un òrgan amb funcions equivalents que es coneix com a cec hepàtic que es desenvolupa durant la metamorfosi a partir de la regió que uneix l'esòfag i l'intestí anterior. Si es considerés l'expressió de l'*Adh3* com un marcador posicional de la regió homòloga entre vertebrats i cefalocordats, les nostres dades recolzarien una relació d'analogia i no d'homologia entre fetge i cec hepàtic, ja que l'*Adh3* s'expressa en l'intestí mitjà i posterior, lluny del lloc on es desenvoluparà el cec hepàtic. A més, aquests resultats conjuntament amb la recent detecció d'activitat β-galactosidasa endògena (Cañestro, 2001), revelen una diferenciació funcional al llarg del tub digestiu de l'amfiox malgrat l'aparent manca de diferenciació morfològica, i aporten dades d'interès sobre l'origen evolutiu del fetge i el sistema digestiu.

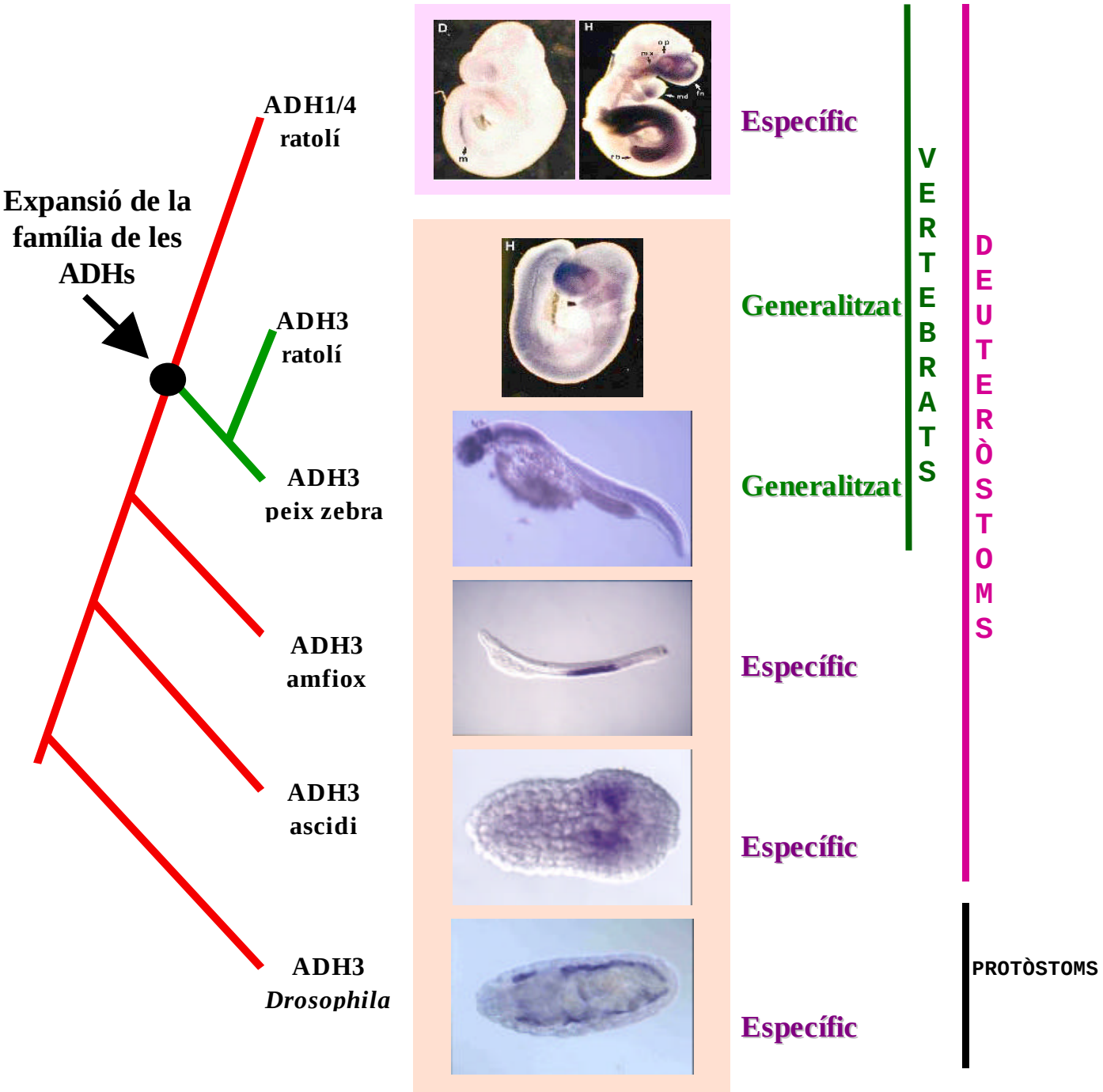


Figura 2. Representació gràfica de l'evolució de les funcions dels membres de la família de les *Adh*s en relació a la seva expansió. A l'esquerra es representa la relació filogenètica de les ADHs i es mostra la història evolutiva dels patrons d'expressió: específica de teixit (vermell) i generalitzada (verd). En el centre, s'agrupen les diferents classes d'ADHs segons les seves activitats catalítiques: ADH3 (fons salmó) i ADH1/4 (fons lila).

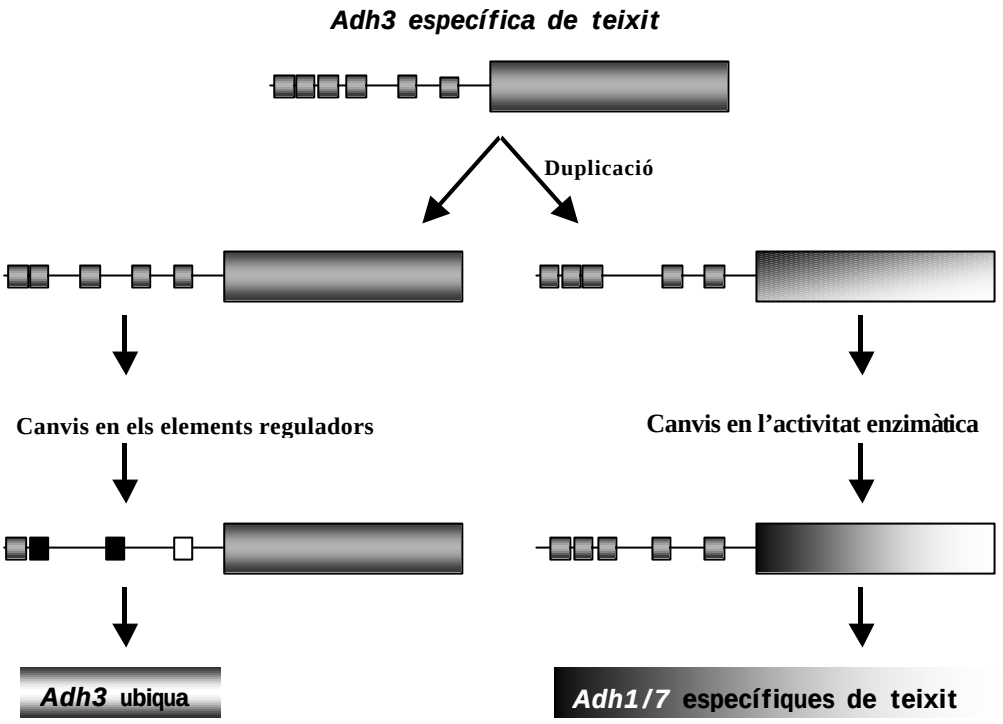


Figura 3. Representació esquemàtica dels canvis produïts en les regions codificants i en les regions reguladores que haurien determinat l'evolució funcional de la família de les *Adh*s. L'Adh3 ha preservat la seva activitat mentre que ha modificat el seu patró d'expressió específic de teixit a ubic. Per contra, les noves Adhs haurien adquirit la capacitat de metabolitzar nous substrats però haurien mantingut en gran part la seva especificitat tissular.

1.3 La família ADH: un cas singular

Una idea predominant en els estudis de l'evolució de les funcions és que els gens reguladors tenen més possibilitats d'evolucionar que aquells que realitzen funcions enzimàtiques (Cooke *et al.,* 1997). En els enzims s'haurien de donar substitucions múltiples en la regió codificant que generessin una nova funció, i donada la baixa freqüència d'aquest fenomen molts gens es perdrien abans d'adquirir noves funcions (Patton *et al.,* 1998). A més, la major part d'exemples d'evolució per complementació i subfuncionalització hipotetitzen que un patró general canvia a un de més restringit (Force *et al.,* 1999). La família de les ADHs és una clara excepció d'aquestes dues premisses: són enzims que han modificat les seves característiques bioquímiques, reconeixen nous substrats i s'han donat canvis d'un patró d'expressió restringit a general.

2. Constància evolutiva de l'ADH3

L'alineament de l'ADH3 dels procordats amb les d'altres organismes tant distants com bacteris, llevats i plantes, revela la seva "constància" al llarg de tota l'escala evolutiva com un dels seus trets característics (Danielsson *et al.,* 1994a). Un terç de les posicions estan totalment conservades, i quasi la meitat de la molècula presenta una variabilitat molt limitada. Tanmateix, el patró de conservació no és uniforme al llarg de la molècula, havent-hi segments molt conservats i d'altres més variables. La caracterització de l'estructura terciària de l'ADH3 ha mostrat com aquestes regions més constants es troben a la part interna i corresponen als dominis funcionals (centre catalític, unió al substrat, unió al cofactor, unió als àtoms de zenc, i regions de dimerització) mentre que les menys conservades ocupen regions més superficials (Yang *et al.,* 1997).

La baixa saturació mutacional de l'ADH3 en el regne animal juntament amb la seva constància evolutiva, permet considerar que l'ADH3 és una bona eina per estimar els temps de divergència entre diferents grups d'organismes. Les estimes fetes a partir de l'ADH3 en el decurs d'aquest treball indiquen que:

i) les separacions entre deuteròstoms–protòstoms i cefalocordats-vertebrats varen tenir lloc 731-744 MA i 682-694 MA, respectivament, estimes molt similars a d'altres fetes amb molècules diferents (Nikoh *et al.,* 1997, Gu, 1998, Doolittle *et al.,* 1996). Cal destacar el petit interval de temps (50-70 MYA) entre les dues separacions, que recolzen una veritable "explosió Càmbrica", entesa com l'aparició abrupta de la majoria de grups dels metazous durant el Neoproterozoic tardà (Ayala & Rzhetsky, 1998). Moltes filogènies moleculars, entre elles les basades en l'ADH3 (fig. 4), mostren incongruències en les relacions entre organismes que divergiren en aquest curt període, degut a la impossibilitat d'acumular diferències moleculars suficients com per reproduir arbre filogenètics amb embrancaments robustes (Philippe *et al.,* 1994).

ii) la separació entre les dues espècies d'amfioxs, *B. floridae* i *B. lanceolatum*, es va produir fa 190 MA. Aquesta és una de les primeres estimes moleculars de l'especiació entre dos cefalocordats. Malgrat podria semblar massa gran per dos organismes que morfològicament són tan semblants, és congruent amb la distribució geogràfica actual de ambdues espècies, als dos costats de l'oceà Atlàntic, i es relacionaria amb l'inici del seu aïllament geogràfic produït pel trencament de la Pangea i l'obertura del Atlàntic fa uns 200 MA. Així, es podria argumentar que la gran similitud entre les dues espècies no es deguda a que divergiren fa poc temps, sinó a la absència de grans canvis morfològics durant els darrers 200 MA. Recentment, la separació entre *B. floridae* i *B. belcherei*, dues espècies separades per l'oceà Pacífic, reforcen els nostres resultats (Nohara *et al.,* 2001). De fet, aquesta estabilitat morfològica recolzaria la condició d'organisme "arquetípic" postulada per aquest subfílum (Owen, 1848, Garcia-Fernàndez & Holland, 1994).

L'ADH3 probablement ha mantingut constant la seva taxa d'evolució des de temps remots, com reflecteix la reduïda variabilitat de les seves constants cinètiques a bacteris, fongs, plantes i animals. Aquesta darrera hipòtesi no ha pogut ser comprovada amb els mètodes de filogènia actuals, donat que encara subestimen el nombre real de substitucions entre dues molècules quan el temps de divergència és massa gran, com és el cas de les separacions animals-plantes-fongs i eucariotes-procariotes. Les estimes dels temps de divergència basades en l'ADH3 malgrat possiblement no són acurades, donen valors molt similars a d'altres obtinguts amb altres molècules (Nikoh *et al.,* 1997, Doolittle *et al.,* 1996), 1120-1490 MA entre animals-plantes-fongs i 1630 MA entre eucariotes-procariotes. En el decurs de l'anàlisi filogenètic de totes les ADHs-MDR dels éssers vius, i a partir del fet que l'ADH3 no ha estat mai identificada en cap arqueobacteri ni en cap eubacteri gram positiu, essent present en tots els eucariotes exceptuant els protistes, sempre s'ha contemplat la possibilitat d'una transferència horitzontal molt antiga. Així, l'estima de 1630 MA no indicaria la divergència entre eucariotes i procariotes sinó el moment de la transferència horitzontal. Aquesta possibilitat concordaria la hipòtesi recent de l'existència de dues grans onades de transferències horitzontals, una d'elles fa uns 1800 MA (Hedges *et al.,* 2001), i amb l'anàlisi filogenètica de totes les ADHs al llarg de tota l'escala evolutiva (Philippe & Lopez, 2001).

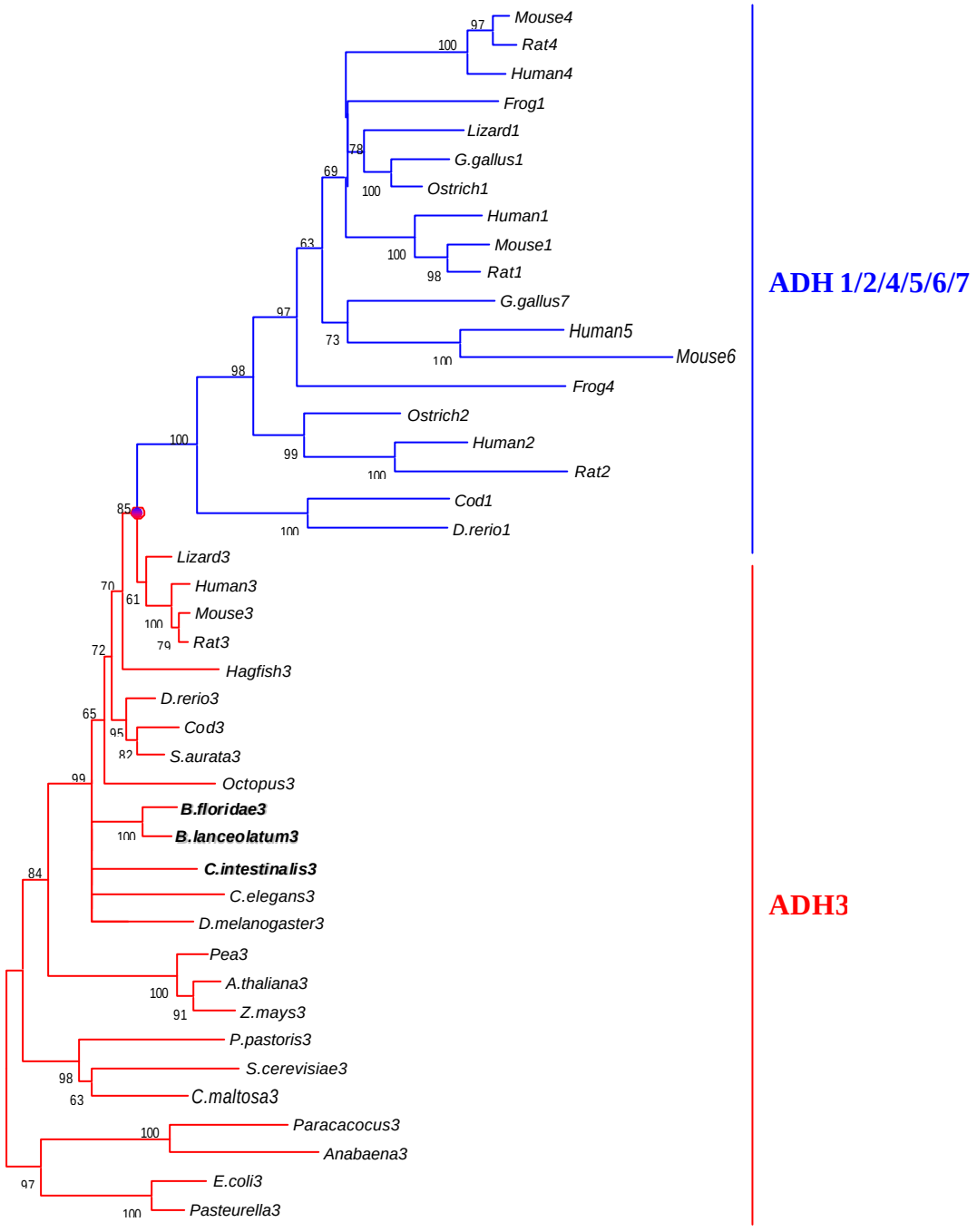


Figura 4. Relacions filogenètiques entre les diferents classes d'ADHs (classe 3 en vermell i la resta de classes en blau) al llarg de tota l'escala evolutiva. L'arbre filogenètic ha estat construït per *neighbor-joining* amb una distribució gamma (α=0.8) i fent servir les ADHs bacterianes com a grup extern. Els números indiquen els valors *bootstrap* (n=1000) per cada node i aquells que presentaven valors inferiors a 50 ha estat col·lapsats. El punt indicaria l'expansió de la família dins del llinatge dels vertebrats.

3. L'ADH3, un enzim *housekeeping*?

L'elevat grau de conservació estructural i bioquímic que ha mostrat l'ADH3 dels procordats, estaria d'acord amb que la seva funció enzimàtica s'ha mantingut al llarg de l'evolució dels animals. No obstant, els canvis en el patró d'expressió, obre nous interrogants sobre la seva funció(ns) fisiològica en els vertebrats.

Una mateixa proteïna pot estar realitzant tasques diferents en llocs o organismes diferents. S'han descrit gens que desenvolupen més d'un rol: *i*) existeixen proteïnes multifuncionals que degut a que contenen diferents dominis poden dur a terme diferents funcions (Breitling *et al.,* 2001), *ii*) hi ha gens que malgrat la seva funció molecular és única, el fet que intervinguin en diferents xarxes gèniques implica que fan tasques molt diferents (Wilkinson & Shyu, 2001, Ladomery, 1997), i per últim, *iii*) el fenomen conegut com a "*gene sharing*", segons el qual un gen pot adquirir i mantenir una segona funció sense perdre la funció primària com per exemple enzims com la lactat deshidrogenasa, argininosuccinat liasa o aldehid deshidrogenasa, a més de dur a terme la seva funció enzimàtica, formen part de l'estructura del cristal·lí de l'ull participant en la refracció i absorció de la llum (Borras & Rodokanaki, 1989; Piatigorsky *et al.,* 1988, Gonzalez *et al.,* 1994).

Tradicionalment, l'ADH3 ha estat considerada un enzim involucrat en el manteniment cel·lular basal eliminant el formaldehid. Es defineix un gen com *housekeeping* quan té una tasca de manteniment cel·lular, una expressió ubiqua, generalment baixa, molts cops localitzat en regions genòmiques no metilades amb nivells de transcripció basals poc regulats, i expressió uniforme en la majoria de teixits. Sota aquest concepte, l'*Adh3* no s'hauria de classificar en aquesta categoria perquè: 1) en la major part de prevertebrats no presenta una expressió ubiqua, 2) es troben diferències importants entre diferents teixits, 3) l'estudi del seu promotor sembla revelar que es un gen finament regulat amb elements reguladors específics de teixit, i 4) es troba en regions genòmiques metilades. Aquest seria el primer pas per obrir noves perspectives en l'estudi funcional de l'ADH3 i tractar de relacionar-la a d'altres possibles funcions: generador de poder reductor en forma de $NADH^+$, implicació en el manteniment del potencial redox en situacions metabòliques particulars: elevada proliferació, síntesi de DNA, la fecundació, etc.

4. L'amfiox com a model animal: un organisme arquetípic amb un genoma ... peculiar.

Durant la darrera dècada, els amfioxs han tornat a ocupar el centre de l'escenari de la recerca orientada a estudiar l'origen dels vertebrats (Gee, 1994). Des de que Owen (1848) va descriure el seu pla corporal com arquetípic, s'han intentat identificar les homologies amb els diferents òrgans dels vertebrats. En aquest sentit, l'estratègia de comparar els patrons d'expressió de gens d'amfiox amb els homòlegs a vertebrats ha servit per identificar algunes homologies corporals i per a distingir les funcions ancestrals de les innovacions evolutives dels vertebrats. Aquests treballs s'han basat majoritàriament en la caracterització de gens a nivell de cDNA i de fet, des dels treballs sobre la reassociació de DNA genòmic de Schmidtke (1979), pocs són els casos on s'han estudiat les estructures gèniques, la composició de les regions no codificants i l'arquitectura del genoma. La caracterització de les regions genòmiques de l'*Adh* a *B. floridae* i *B. lanceolatum* ha servit per a revelar propietats dels gens dels cefalocordats, que posteriorment s'ha vist eren compartides per altres gens, fent-se així extensives al genoma d'aquests organismes.

4.1 Dinàmica dels introns durant l'evolució dels cordats.

Els guanys o pèrdues d'introns com a canvis genòmics rars (RGCs) han estat molt útils com a marcadors cladístics en estudis filogenètics de peixos (Venkatesh *et al.,* 1999) o d'insectes (Rokas & Holland, 2000) i per estudiar expansions de famílies gèniques (Cañestro *et al.,* 2002). Però, a més a més l'anàlisi del numero, mida i posició dels introns, també aporta informació sobre l'evolució del genoma. Així, els nostres resultats i els d'altres autors indiquen que les posicions dels introns tendeixen a estar altament conservades en tots els cordats i el seu nombre a ésser lleugerament més elevat en el procordats. La mida mitjana dels introns difereix significativament entre els diferents subfila, des dels extremadament petits dels urocordats (pocs centenars de bp), als moderadament petits dels cefalocordats (poques kb) i als enormes

introns dels vertebrats (fins a centenars de kb). D'altra banda, els protòstoms solen tenir un baix número d'introns i de mides petites, però la manca d'un nombre de dades suficients no permet caracteritzar l'estructura genòmica ancestral dels deuteròstoms i protòstoms, alimentant el clàssic debat sobre l'origen primerenc o tardà dels introns. Malgrat tot, l'anàlisi d'introns de les *Adh*s encaixaria millor amb la presència d'introns primerencs, els quals s'haurien perdut majoritàriament en els protòstoms, possiblement lligat a la compactació del seu genoma en un procés d'adaptació a un ràpid desenvolupament embrionari.

4.2 El polimorfisme genètic de l'amfiox com a eina per l'anàlisi funcional

Una de les característiques més notòries del genoma de l'amfiox és el seu elevat grau de polimorfisme genètic tant a nivell de regions codificants com no codificants. La comparació dels abundants polimorfismes a nivell inter- i intraespecífics en el cas de l'ADH, mostra com la majoria dels canvis no sinònims es produeixen en les posicions variables entre les dues espècies d'amfioxs. Aquestes posicions probablement deuen alterar poc la funció bàsica de l'ADH3 de l'amfiox i per tant tindrien poc pes en la densitat funcional de l'enzim. Per altra banda, les regions no codificants (regions reguladores, introns i regions transcrites no traduïdes) presenten un elevat nombre de substitucions i petites insercions-delecions. També en aquest cas, l'anàlisi dels polimorfismes en aquestes regions pot permetre identificar regions invariables com a candidates a tenir constriccions funcionals participant en el control de la regulació gènica.

4.3 Polimorfisme intrònic de l'amfiox

L'anàlisi Southern d'un gen de còpia única en cefalocordats presenta freqüentment patrons d'hibridació complexos amb l'aparició de múltiples bandes de mida variable entre diferents individus, i ocasionalment bandes addicionals tènues. En el cas del gen *Adh*, l'anàlisi de les regions genòmiques corresponents ha revelat una quantitat important de DNA repetit amb un gran polimorfisme al·lèlic que genera una elevada variabilitat en la mida dels introns. Aquest DNA repetit està format majoritàriament per repeticions imperfectes en tàndem del tipus minisatèl·lit (VNTRs) i en menor grau de microsatèl·lits i elements genètics mòbils.

En vertebrats, la inestabilitat dels minisatèl·lits va generalment lligada a la línia germinal. Per contra, en cefalocordats sembla ser abundant a nivell somàtic i en aquest sentit l'amfiox seria un organisme mosaic. La detecció d'al·lels recombinants en el genoma d'amfiox, algun dels quals havia variat la seva mida, planteja la possibilitat que l'entrecreuament desigual sigui el mecanisme generador de la inestabilitat d'aquestes repeticions. El grau de mosaïcisme per un locus determinat varia entre diferents individus, probablement depenent del moment del desenvolupament o del tipus cel·lular en què es va donar el canvi.

Aquest elevat grau de polimorfisme genètic s'ha de tenir en compte en fer les estimes del nombre de còpies d'un gen en el genoma d'amfiox, doncs podria conduir a valoracions errònies. Per tant, seria recomanable en anàlisis Southern utilitzar sondes curtes, sense regions intròniques, i sobre DNAs de diferents individus per separat per evitar patrons d'hibridació complexos. A més, s'ha de tenir en compte que l'aparició de bandes d'hibridació tènues podrien ser variants al·lèliques fruit del polimorfisme somàtic, enlloc d'altres gens relacionats. Les variants fins i tot podrien ser aïllades en cribratges genòmics on apareixerien més de dos al·lels per un gen de còpia única i individu. Finalment, aquest elevat polimorfisme genètic hauria d'ésser considerat a l'hora de dissenyar el possible projecte de seqüenciació del genoma d'aquest organisme.

4.4 Miratges: un nou tipus de minisatèl·lits

Cal esmentar l'abundant presència d'un nou tipus de minisatèl·lits anomenats *miratges*, que es caracteritzen perquè la seva subunitat de repetició inclou part de l'extrem de l'exó i part de la seqüència intrònica adjacent. La seva abundància en el genoma d'amfiox apunta a que la seva estructura no seria atzarosa, i es generarien i es processarien per mecanismes específics, que farien dels *miratges* una nova entitat genètica. Malgrat no hi ha cap evidència al respecte, l'elevat grau de conservació dels senyals donadors-acceptors de *splicings* en cadascuna de les subunitats, insinuaria que l'origen i la dinàmica dels *miratges* podria estar lligada d'alguna manera a la maquinària de *splicing*. Tanmateix, que s'hagin descrit repeticions similars en vertebrats planteja dos possibles escenaris: a) els *miratges* ja estaven presents en l'ancestre dels vertebrats i cefalocordats i han estat

gairebé eliminats en el genoma de vertebrats, mentre que s'han preservat en els cefalocordats, recolzant la visió de l'amfiox com un organisme arquetípic; b) la seva presència en vertebrats és una mera coincidència sense cap mena de lligam evolutiu amb els *miratges* dels cefalocordats.

S'ha demostrat que entrecreuaments i reordenacions dins dels *miratges* provoquen la formació d'al·lels recombinants. Així doncs, l'existència de *miratges* podria afavorir l'aparició de regions amb alta freqüència de recombinació desigual i per tant contribuir de forma important a l'augment de la diversitat al·lèlica d'aquests organismes. Donat que la freqüència d'entrecreuaments desiguals és proporcional a la quantitat de seqüències similars, en els casos de gens duplicats, l'existència de miratges flanquejant els exons incrementaria la freqüència de recombinacions entre ells i d'aquesta manera promouria fenòmens de conversió gènica. S'han descrit casos de "resurreccions" de pseudogens per conversió gènica mitjançant la transferència d'informació d'un gen funcional de la mateixa família (Trabesinger-Ruef *et al.,* 1996). Així doncs, l'homogenització dels gens duplicats per entrecreuament desigual i conversió gènica afavorida per la presència de *miratges*, retardaria la desfuncionalització i augmentaria la probabilitat de neofuncionalització d'una de les còpies. Finalment, en un suposat escenari evolutiu en què els *miratges* fossin ancestrals, la seva presència podria haver influenciat el destí dels gens resultants de les duplicacions a gran escala postulades durant l'evolució primerenca dels vertebrats.

4.5 Origen i manteniment del polimorfisme genètic de l'amfiox

L'elevat polimorfisme genètic observat en les poblacions de cefalocordats és el reflex de l'estat transitori resultant de l'equilibri dinàmic entre la creació i desaparició de variants al·lèliques al llarg del temps. Les poblacions grans i ben establertes poden mantenir una gran riquesa al·lèlica com a fruit del baix efecte de la selecció sobre caràcters quasi neutres i la poca influència de la deriva gènica. Així doncs, el polimorfisme dels amfioxs podria estar revelant la seva estructura poblacional, amb una profunda estabilitat sense colls d'ampolles recents. De fet, són les poblacions marginals amb un nombre reduït d'individus, ja sigui degut a extincions massives o aïllament geogràfic, les més susceptibles a l'especiació i a patir canvis dràstics en les seva biologia. L'ecologia dels cefalocordats, amb poblacions grans en ambients estables, podria explicar el baix nombre d'espècies existent en aquest subfílum i hauria contribuït a mantenir l'estatus arquetípic del seu disseny corporal.

4.6 L'amfiox com a organisme model

L'amfiox és l'animal viu més interessant per l'estudi de l'evolució i origen dels vertebrats, degut a la seva privilegiada posició filogenètica, la seva baixa redundància gènica i el seu disseny corporal "arquetípic". No obstant, aquest organisme presenta una sèrie de característiques que dificulten el seu establiment definitiu entre els organismes models més emprats. Tot i la seva distribució cosmopolita, la seva accessibilitat és bastant reduïda. Malgrat s'estan fent progressos en l'estabulació de l'espècie *B. belcherei* (Watanabe *et al.,* 1999), encara és llunyana la possibilitat de mantenir-los en un laboratori i controlar el seu cicle biològic. L'espècie *B. floridae* és una de les més estudiades de les quals es coneix més la seva biologia. No obstant, la particular forma de reproduir-se, una nit cada dues setmanes durant l'estiu a les costes de Florida, a més de suposar un elevat cost, limita molt les condicions de treball i els tipus d'experiments que es poden realitzar. De l'espècie *B. lanceolatum*, la més abundant en les costes europees, no es coneix amb exactitud el seu cicle reproductiu, i el fet que sol viure a fondàries de més de 3 m dificulta la seva captura requerint l'ús de dragues.

D'altra banda, els embrions d'amfiox són molt fràgils, i fins ara no s'ha establert cap protocol d'introducció de DNA per electroporació ni microinjecció. L'existència de certa fluorescència corporal i el descobriment que l'amfiox té activitat β-galactosidasa endògena (Cañestro *et al.,* 2001), també dificulten l'ús dels gens *GFP* i *LacZ* com a gens reporters en experiments de transgènesi. Finalment, el fet que molts gens hagin tingut la seva particular història evolutiva en el llinatge dels amfiox (Dalfó *et al.*, 2001, Ferrier *et al.*, 2001, Garcia-Fernàndez *et al*., 2001) l'elevat polimorfisme del genoma de l'amfiox i l'existència notòria de DNA repetit (minisatèl·lits i elements mòbils) dificulten en gran mesura la recerca i caracterització de gens i els possibles projectes de seqüenciació genòmica. Aquestes dades, juntament amb la presència no tan esporàdica de duplicacions gèniques pròpies del

llinatge dels cefalocordats, fan que l'amfiox no sigui un organisme tan senzill d'estudiar com s'esperava *a priori*.

CONCLUSIONS

CONCLUSIONS

1. L'expansió de la família gènica de les *Adh*s es va produir fa uns 500 MA, després de la separació entre els cefalocordats i els vertebrats (690 MA), probablement per duplicacions gèniques en tàndem i independentment de les 2 rondes de duplicacions genòmiques postulades durant l'evolució primerenca dels vertebrats. Es tracte d'una família complexa i heterogènia formada per més de set classes diferents. L'ADH3 és l'únic membre de la família present en els procordats

1. The *Adh* gene family expansion took place 500 MA, after the cephalochordate-vertebrate split (690 MA), probably by gene tandem duplications, rather than linked to the 2-Rounds of genome duplications postulated in the early vertebrate evolution. This is a complex and highly heterogeneous family that contains a minimum of seven classes. ADH3 is the only family member present in pro-chordates

---

2. L'*Adh3* codifica per una proteïna que ha mantingut les propietats bioquímiques i ha canviat el patró d'expressió al llarg de l'evolució dels cordats. Contràriament, els gens duplicats han adquirit noves funcions - per neofuncionalització- i han mantingut l'especificitat tissular ancestral.

2. The *Adh3* encodes a protein that has retained the biochemical features but its expression pattern has changed during the evolution of chordates. On the other hand, the duplicated genes have acquired novel functions -by neofunctionalization- and have maintained the original tissue specificity.

---

3. L'*Adh3* a invertebrats presenta una expressió específica de teixit. Això suggereix diferències funcionals respecte al rol de manteniment cel·lular basal de l'ADH3 postulat a vertebrats.

3. The *Adh3* tissue-specificity shown in invertebrates suggests functional differences with the housekeeping role ascribed to the vertebrate ADH3.

---

4. L'abundància de DNA repetit –minisatèl·lits i elements mòbils- en els introns de l'*Adh3* de *B. floridae* i *B. lanceolatum* és la causa de l'elevat polimorfisme descrit en aquesta regió genòmica.

4. The abundance of repetitive DNA –minisatellites and mobile elements- found in intron sequences of the *B. floridae* and *B. lanceolatum Adh3* is the cause of the high level of polymorphism described in this genomic region.

---

5. L'existència d'inestabilitat genètica a nivell somàtic, probablement generada per fenòmens d'entrecreuaments interal·lèlics, fa de l'amfiox un organisme mosaic.

5. The genetic instability described at the somatic level, probably resulting from inter-allelic crossover events, confers to the amphioxus animal a mosaic status.

---

6. S'ha descrit un nou tipus de minisatèl·lits -*miratges* – característics del genoma dels cefalocordats, que es caracteritzen per la seva peculiar subunitat de repetició que inclou part de l'exó adjacent.

6. A novel kind of minisatellites has been described –*miratges*- as an extended feature through the cephalochordate genome. They are characterized by an odd repeat subunit, which spans over the exon-intron boundaries.

---

7. Finalment, s'ha descobert que els amfioxs tenen activitat β–galactosidasa endògena, i la seva detecció aporta un nou marcador histoquímic del seu sistema digestiu útil per l'estudi de la seva diferenciació funcional i morfològica.

7. Finally, endogenous β–galactosidase activity has been detected in amphioxus, which provides a new histochemical marker for the digestive system, and is useful to monitor its functional and morphological differentiation.

REFERÈNCIES

Referències

- Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., *et al.* (2000) The genome sequence of Drosophila melanogaster. *Science* 287:2185-95.
- Alvarez-Valin, F., Jabbari, K., Bernardi, G. (1998) Synonymous and nonsynonymous substitutions in mammalian genes: intragenic correlations. *J Mol Evol* 46:37-44.
- Ang, H., Deltour, L., Hayamizu, T. F., Zgombic Knight, M., Duester, G. (1996a) Retinoic acid synthesis in mouse embryos during gastrulation and craniofacial development linked to class IV alcohol dehydrogenase gene expression. *Journal of Biological Chemistry* 271:9526-9534
- Ang, L., Deltour, L., Zgombic Knight, M., Wagner, M. A., Duester, G. (1996b) Expression patterns of class I and class IV alcohol dehydrogenase genes in developing epithelia suggest a role for alcohol dehydrogenase in local retinoic acid synthesis. *Alcoholism Clinical and Experimental Research* 20:1050-1064
- Aparicio, S. (2000) Vertebrate evolution: recent perspectives from fish. *Trends Genet* 16:54-6.
- Ayala, F. J., Rzhetsky, A. (1998) Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. *Proc Natl Acad Sci U S A* 95:606-11.
- Beisswenger, T. B., Holmquist, B., Vallee, B. L. (1985) chi-ADH is the sole alcohol dehydrogenase isozyme of mammalian brains: implications and inferences. *Proc Natl Acad Sci U S A* 82:8369-73.
- Berget, S. M., Moore, C., Sharp, P. A. (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A* 74:3171-5.
- Bilanchone, V., Duester, G., Edwards, Y., Smith, M. (1986) Multiple mRNAs for human alcohol dehydrogenase (ADH): developmental and tissue specific differences. *Nucleic Acids Res* 14:3911-26.
- Bird, A. P., Taggart, M. H. (1980) Variable patterns of total DNA and rDNA methylation in animals. *Nucleic Acids Res* 8:1485-97.
- Boleda, M. D., Julià, P., Moreno, A., Parés, X. (1989) Role of extrahepatic alcohol dehydrogenase in rat ethanol metabolism. *Archives of Biochemistry and Biophysics* 274:74-81
- Boleda, M. D., Saubi, N., Farrés, J., Parés, X. (1993) Physiological substrates for rat alcohol dehydrogenase classes: Aldehydes of lipid peroxidation, omega-hydroxyfatty acids, and retinoids. *Archives of Biochemistry and Biophysics* 307:85-90
- Bonnichsen, R. K., Wassén, A. M. (1948) Crystalline alcohol deshydrogenase from horse liver. *Arch. Biochem Biophys* 18:361-363
- Borras, T., Rodokanaki, A. (1989) Primary structure of zeta-crystallin protein from guinea pig. Its similarity to the enzyme alcohol dehydrogenase. *Lens Eye Toxic Res* 6:795-805
- Braunitzer, G., Gehring-Muller, R., Hilschmann, N., Hilse, K., Hobom, G., *et al.* (1961) Die konstitution des normalen adulten hämoglobins. *Z Physiol Chem* 325:283-286
- Breitling, R., Marijanovic, Z., Perovic, D., Adamski, J. (2001) Evolution of 17beta-HSD type 4, a multifunctional protein of beta-oxidation. *Mol Cell Endocrinol* 171:205-10.
- Bridges, C. (1936) The Bar gene, a duplication. *Science* 83:210-211
- Brock, T., Madigan, M., Martinko, J., Parker, J. 1994. *Biology of microorganisms.*
- Buard, J., Collick, A., Brown, J., Jeffreys, A. J. (2000) Somatic versus germline mutation processes at minisatellite CEB1 (D2S90) in humans and transgenic mice. *Genomics* 65:95-103.
- *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode C. elegans: a platform for investigating biology.. *Science* 282:2012-8.
- Cañestro, C., Albalat, R., Escrivà, H., Gonzàlez-Duarte, R. (2001) Endogenous beta-galactosidase activity in amphioxus: a useful histochemical marker for the digestive system. *Dev Genes Evol* 211:154-6.
- Cañestro, C., Albalat, R., Hjelmqvist, L., Godoy, L., Jörnvall, H., Gonzàlez-Duarte, R. (2002) Ascidian and Amphioxus Adh genes correlate functional and molecular features of the ADH Family expansion during vertebrate evolution. *J Mol Evol* 54:81-89
- Cañestro, C., Hjelmqvist, L., Albalat, R., Garcia-Fernàndez, J., Gonzàlez-Duarte, R., Jörnvall, H. (2000) Amphioxus alcohol dehydrogenase is a class 3 form of single type and of structural

conservation but with unique developmental expression. *Eur J Biochem* 267:6511-8.

- Casanova, M., Bell, D. A., Heck, H. D. (1997) Dichloromethane metabolism to formaldehyde and reaction of formaldehyde with nucleic acids in hepatocytes of rodents and humans with and without glutathione S-transferase T1 and M1 genes. *Fundam Appl Toxicol* 37:168-80.
- Cederlund, E., Peralba, J. M., Parés, X., Jörnvall, H. (1991) Amphibian alcohol dehydrogenase, the major frog liver enzyme: Relationships to other forms and assessment of an early gene duplication separating vertebrate class I and class III alcohol dehydrogenases. *Biochemistry* 30:2811-2816
- Charlesworth, B., Sniegowski, P., Stephan, W. (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371:215-20.
- Chen, C. S., Yoshida, A. (1991) Enzymatic properties of the protein encoded by newly cloned human alcohol dehydrogenase ADH6 gene. *Biochem Biophys Res Commun* 181:743-7.
- Cheung, C., Smith Camilla, K., Höög Jan, O., Hotchkiss Sharon, A. M. (1999) Expression and localization of human alcohol and aldehyde dehydrogenase enzymes in skin. *Biochemical-and-Biophysical-Research-Communications. July 22, 1999;* 261:100-107.
- Claverie, J. M. (2001) Gene number. What if there are only 30,000 human genes? *Science* 291:1255-7.
- Colot, V., Maloisel, L., Rossignol, J. L. (1999) [DNA repeats and homologous recombination: a probable role for DNA methylation in genome stability of eukaryotic cells]. *J Soc Biol* 193:29-34
- Colot, V., Rossignol, J. L. (1999) Eukaryotic DNA methylation as an evolutionary device. *Bioessays* 21:402-11.
- Comeron, J. M., Kreitman, M. (2000) The correlation between intron length and recombination in drosophila. Dynamic equilibrium between mutational and selective forces. *Genetics* 156:1175-90.
- Cooke, J., Nowak, M. A., Boerlijst, M., Maynard-Smith, J. (1997) Evolutionary origins and maintenance of redundant gene expression during metazoan development. *Trends Genet* 13:360-4.
- Dalfó, D., Cañestro, C., Albalat, R., Gonzàlez-Duarte, R. (2001) Characterization of a microsomal retinol dehydrogenase gene from amphioxus: retinoid metabolism before vertebrates. *Chem Biol Interact* 130-132:359-370
- Danielsson, O., Atrian, S., Luque, T., Hjelmqvist, L., Gonzàlez Duarte, R., Jörnvall, H. (1994a) Fundamental molecular differences between alcohol dehydrogenase classes. *Proceedings of the National Academy of Sciences of the United States of America* 91:4980-4984
- Danielsson, O., Eklund, H., Jörnvall, H. (1992) The major piscine liver alcohol dehydrogenase has class-mixed properties in relation to mammalian alcohol dehydrogenases of classes I and III. *Biochemistry* 31:3751-3759
- Danielsson, O., Shafqat, J., Estonius, M., El Ahmad, M., Jörnvall, H. (1996) Isozyme multiplicity with anomalous dimer patterns in a class III alcohol dehydrogenase: Effects on the activity and quaternary structure of residue exchanges at "nonfunctional" sites in a native protein. *Biochemistry* 35:14561-14568
- Danielsson, O., Shafqat, J., Estonius, M., Jörnvall, H. (1994b) Alcohol dehydrogenase class III contrasted to class I: Characterization of the cyclostome enzyme, the existence of multiple forms as for the human enzyme, and distant cross-species hybridization. *European Journal of Biochemistry* 225:1081-1088
- Debrauwere, H., Gendrel, C. G., Lechat, S., Dutreix, M. (1997) Differences and similarities between various tandem repeat sequences: minisatellites and microsatellites. *Biochimie* 79:577-86.
- Deltour, L., Haselbeck, R. J., Ang, H. L., Duester, G. (1997) Localization of class I and class IV alcohol dehydrogenases in mouse testis and epididymis: potential retinol dehydrogenases for endogenous retinoic acid synthesis. *Biol Reprod* 56:102-9.
- Dietrich, C. J., Richards, I. S., Bernard, T. E., Hammad, Y. Y. (1996) Human stress protein response to formaldehyde exposure. *Exp Toxicol Pathol* 48:518-9.
- Doolittle, R. F., Feng, D. F., Tsang, S., Cho, G., Little, E. (1996) Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* 271:470-7.
- Duester, G. (2000) Families of retinoid dehydrogenases regulating vitamin A function: production of visual pigment and retinoic acid. *Eur J Biochem* 267:4315-24.

- Duester, G., Farrés, J., Felder, M. R., Holmes, R. S., Höög, J. O., *et al*. (1999) Recommended nomenclature for the vertebrate alcohol dehydrogenase gene family. *Biochem Pharmacol* 58:389-95.
- Edenberg, H. J. (2000) Regulation of the mammalian alcohol dehydrogenase genes. *Prog Nucleic Acid Res Mol Biol* 64:295-341
- Edenberg, H. J., Ho, W. H., Hur, M. W. (1995) Promoters of the mammalian class III alcohol dehydrogenase genes. *Adv Exp Med Biol* 372:295-300
- Engeland, K., Höög Jan, O., Holmquist, B., Estonius, M., Jörnvall, H., Vallee Bert, L. (1993) Mutation of Arg-115 of human class III alcohol dehydrogenase: A binding site required for formaldehyde dehydrogenase activity and fatty acid activation. *Proceedings of the National Academy of Sciences of the United States of America* 90:2491-2494
- Estonius, M., Danielsson, O., Karlsson, C., Persson, H., Jörnvall, H., Höög Jan, O. (1993) Distribution of alcohol and sorbitol dehydrogenases: Assessment of mRNA species in mammalian tissues. *European Journal of Biochemistry* 215:497-503
- Estonius, M., Höög Jan, O., Danielsson, O., Jörnvall, H. (1994) Residues specific for class III alcohol dehydrogenase. Site-directed mutagenesis of the human enzyme. *Biochemistry* 33:15080-15085
- Estonius, M., Svensson, S., Höög Jan, O. (1996) Alcohol dehydrogenase in human tissues: Localisation of transcripts coding for five classes of the enzyme. *FEBS Letters* 397:338-342
- Felsenstein, J. (1988) Phylogenies from molecular sequences: inference and reliability. *Annu Rev Genet* 22:521-65
- Felsenstein, J. (1996) Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol* 266:418-27
- Fernàndez, M. R., Jörnvall, H., Moreno, A., Kaiser, R., Parés, X. (1993) Cephalopod alcohol dehydrogenase: Purification and enzymatic characterization. *FEBS Federation of European Biochemical Societies Letters* 328:235-238
- Ferrier, D. E., Minguillón, C., Cebrián, C., Garcia-Fernàndez, J. (2001) Amphioxus Evx genes: implications for the evolution of the Midbrain-Hindbrain Boundary and the chordate tailbud. *Dev Biol* 237:270-81.
- Foglio, M., Duester, G. (1996) Characterization of the functional gene encoding mouse class III alcohol dehydrogenase (glutathione-dependent formaldehyde dehydrogenase) and an unexpressed processed pseudogene with an intact open reading frame. *European Journal of Biochemistry* 237:496-504
- Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., Postlethwait, J. (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151:1531-45.
- Garcia-Fernàndez, J., Holland, P. W. (1994) Archetypal organization of the amphioxus Hox gene cluster. *Nature* 370:563-6.
- Garcia-Fernàndez, J., Ferrier, D. E., Minguillón, C., Cebrián, C. (2001) The amphioxus genome in Evo-Devo: archetype or "cul de sac". *Int J Dev Biol* 45 (S1):S137-S138
- Gee, H. (1994) Vertebrate morphology. Return of the amphioxus. *Nature* 370:504-5.
- Gee, H. (1996) Before the backbone. View on the origin of the vertebrates. Chapman & Hall London
- Giri, P. R., Linnoila, M., O'Neill, J. B., Goldman, D. (1989) Distribution and possible metabolic role of class III alcohol dehydrogenase in the human brain. *Brain Research* 481:131-141
- González, P., Hernández-Calzadilla, C., Rao, P. V., Rodríguez, I. R., Zigler, J. S., Jr., Borrás, T. (1994) Comparative analysis of the zeta-crystallin/quinone reductase gene in guinea pig and mouse. *Mol Biol Evol* 11:305-15.
- Gotoh, Y., Sumimoto, H., Minakami, S. (1990) Formation of 20-oxoleukotriene B4 by an alcohol dehydrogenase isolated from human neutrophils. *Biochim Biophys Acta* 1043:52-6.
- Graham, A. (2000) The evolution of the vertebrates--genes and development. *Curr Opin Genet Dev* 10:624-8.
- Graur, D., Li, W. 2000. *Fundamentals of molecular evolution*. 2nd edition ed. Sunderland: Sinauer Associates, Inc.
- Grey, M., Schmidt, M., Brendel, M. (1996) Overexpression of ADH1 confers hyper-resistance to formaldehyde in Saccharomyces cerevisiae. *Current Genetics* 29:437-440
- Gu, X. (1998) Early metazoan divergence was about 830 million years ago. *J Mol Evol* 47:369-71.
- Guigo, R., Muchnik, I., Smith, T. F. (1996) Reconstruction of ancient molecular phylogeny. *Mol Phylogenet Evol* 6:189-213.

- Gutheil, W. G., Holmquist, B., Vallee, B. L. (1992) Purification, characterization, and partial sequence of the glutathione-dependent formaldehyde dehydrogenase from Escherichia coli: A class III alcohol dehydrogenase. *Biochemistry* 31:475-481
- Haldane, J. 1932. *The Causes of Evolution* London: Longmans&Green
- Han, C. L., Liao, C. S., Wu, C. W., Hwong, C. L., Lee, A. R., Yin, S. J. (1998) Contribution to first-pass metabolism of ethanol and inhibition by ethanol for retinol oxidation in human alcohol dehydrogenase family - Implications for etiology of fetal alcohol syndrome and alcohol-related diseases. *European Journal Of Biochemistry. May* 254:25-31
- Hanke, J., Brett, D., Zastrow, I., Aydin, A., Delbruck, S., *et al.* (1999) Alternative splicing of human genes: more the rule than the exception? *Trends Genet* 15:389-90.
- Haselbeck, R., Duester, G. (1997) Regional restriction of alcohol/retinol dehydrogenases along the mouse gastrointestinal epithelium. *Alcoholism Clinical and Experimental Research. Nov.* 21:1484-1490
- Haselbeck, R. J., Ang, H. L., Deltour, L., Duester, G. (1997) Retinoic acid and alcohol/retinol dehydrogenase in the mouse adrenal gland: a potential endocrine source of retinoic acid during development. *Endocrinology* 138:3035-41.
- Hedberg, J. J., Stromberg, P., Höög, J. O. (1998) An attempt to transform class characteristics within the alcohol dehydrogenase family. *Febs Letters. Sep* 436:67-70
- Hedges, S. B., Chen, H., Kumar, S., Wang, D. Y., Thompson, A. S., Watanabe, H. (2001) A genomic timescale for the origin of eukaryotes. *BMC Evol Biol* 1:4
- Hjelmqvist, L., Shafqat, J., Siddiqi Abdur, R., Jörnvall, H. (1995) Alcohol dehydrogenase of class III: Consistent patterns of structural and functional conservation in relation to class I and other proteins. *FEBS Letters* 373:212-216
- Hoffmann, I., Ang, H. L., Duester, G. (1998) Alcohol dehydrogenases in Xenopus development: Conserved expression of ADH1 and ADH4 in epithelial retinoid target tissues. *Developmental Dynamics. Nov* 213:261-270
- Hogenesch, J. B., Ching, K. A., Batalov, S., Su, A. I., Walker, J. R., *et al.* (2001) A comparison of the Celera and Ensembl predicted gene sets reveals little overlap in novel genes. *Cell* 106:413-5.
- Holland, N. D., Chen, J. (2001) Origin and early evolution of the vertebrates: new insights from advances in molecular biology, anatomy, and palaeontology. *Bioessays* 23:142-51.
- Holland, N. D., Holland, L. Z. (1999) Amphioxus and the utility of molecular genetic data for hypothesizing body part homologies between distantly related animals. *Amer Zool* 39:630-640
- Holland, P. W. (1999) Gene duplication: past, present and future. *Semin Cell Dev Biol* 10:541-7.
- Holland, P. W., Garcia-Fernàndez, J., Williams, N. A., Sidow, A. (1994) Gene duplications and the origins of vertebrate development. *Dev Suppl* 43:125-33.
- Holland, P. W. H. (1998) Major transitions in animal evolution: A developmental genetic perspective. *American Zoologist* 38:829-842
- Holmquist, B., Moulis Jean, M., Engeland, K., Vallee Bert, L. (1993) Role of arginine-115 in fatty acid activation and formaldehyde dehydrogenase activity of human class III alcohol dehydrogenase. *Biochemistry* 32:5139-5144
- Hoog, J. O., Brandt, M. (1995) Mammalian class VI alcohol dehydrogenase. Novel types of the rodent enzymes. *Adv Exp Med Biol* 372:355-64
- Höög, J. O., Estonius, M., Danielsson, O. (1994) Site-directed mutagenesis and enzyme properties of mammalian alcohol dehydrogenases correlated with their tissue distribution. *Exs* 71:301-9.
- Hughes, A. L. (1998) Phylogenetic tests of the hypothesis of block duplication of homologous genes on human chromosomes 6, 9, and 1. *Mol Biol Evol* 15:854-70.
- Hughes, A. L. (1999) Phylogenies of developmentally important proteins do not support the hypothesis of two rounds of genome duplication early in vertebrate history. *J Mol Evol* 48:565-76.
- Hughes, M. K., Hughes, A. L. (1993) Evolution of duplicate genes in a tetraploid animal, Xenopus laevis. *Mol Biol Evol* 10:1360-9.
- Hur, M. W., Edenberg, H. J. (1992) Cloning and characterization of the ADH5 gene encoding human alcohol dehydrogenase 5, formaldehyde dehydrogenase. *Gene* 121:305-11.
- Hur, M. W., Ho, W. H., Brown, C. J., Goldman, D., Edenberg, H. J. (1992) Molecular cloning of mouse alcohol dehydrogenase-

B2 cDNA: nucleotide sequences of the class III ADH genes evolve slowly even for silent substitutions. *DNA Seq* 3:167-75

- Iborra, F. J., Renau Piqueras, J., Portoles, M., Boleda, M. D., Guerri, C., Parés, X. (1992) Immunocytochemical and biochemical demonstration of formaldehyde dehydrogenase (class III alcohol dehydrogenase) in the nucleus. *Journal of Histochemistry and Cytochemistry* 40:1865-1878
- Itano, H. (1957) The human hemoglobins: their properties and genetic control. *Adv Prot Chem* 12:216-268
- Jeffreys, A. J., Neumann, R. (1997) Somatic mutation processes at a human minisatellite. *Hum Mol Genet* 6:129-32; 134-6.
- Jeffreys, A. J., Royle, N. J., Wilson, V., Wong, Z. (1988) Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature* 332:278-81.
- Jeffreys, A. J., Tamaki, K., MacLeod, A., Monckton, D. G., Neil, D. L., Armour, J. A. (1994) Complex gene conversion events in germline mutation at human minisatellites. *Nat Genet* 6:136-45.
- Jörnvall, H., Danielsson, O., Eklund, H., Hjelmqvist, L., Höög, J. O., *et al*. (1993) Enzyme and isozyme developments within the medium-chain alcohol dehydrogenase family. *Adv Exp Med Biol* 328:533-44
- Jörnvall, H., Höög, J. O., Persson, B. (1999) SDR and MDR: completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS Lett* 445:261-4.
- Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzàlez-Duarte, R., *et al*. (1995) Short-chain dehydrogenases/reductases (SDR). *Biochemistry* 34:6003-13.
- Jörnvall, H., Persson, M., Jeffery, J. (1981) Alcohol and polyol dehydrogenases are both divided into two protein types, and structural properties cross-relate the different enzyme activities within each type. *Proc Natl Acad Sci U S A* 78:4226-30.
- Julià, P., Boleda, M. D., Farrés, J., Parés, X. (1987) Mammalian alcohol dehydrogenase: characteristics of class III isoenzymes. *Alcohol Alcohol* Suppl:169-73.
- Kedishvili, N., Gough, W., Chernoff, E., Hurley, T., Stone, C., *et al*. (1997) CDNA sequence and catalytic properties of a chick embryo alcohol dehydrogenase that oxidizes retinol and 3-beta,5-alpha-hydroxysteroids. *Journal of Biological Chemistry* 272:7494-7500
- Koivusalo, M., Baumann, M., Uotila, L. (1989) Evidence for the identity of glutathione-dependent formaldehyde dehydrogenase and class III alcohol dehydrogenase. *FEBS* Letters 257:105-109
- Kricker, M. C., Drake, J W., Radman, M. (1992) Duplication-targeted DNA methylation and mutagenesis in the evolution of eukaryotic chromosomes. *Proc Natl Acad Sci U S A* 89:1075-9.
- Kuemmerle, N., Feucht Heinz, H., Kaulfers Paul, M. (1996) Plasmid-mediated formaldehyde resistance in Escherichia coli: Characterization of resistance gene. *Antimicrobial Agents and Chemotherapy* 40:2276-2279
- Ladomery, M. (1997) Multifunctional proteins suggest connections between transcriptional and post-transcriptional processes. *Bioessays* 19:903-9.
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., *et al*. (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860-921.
- Lewis, E. B. (1978) A gene complex controlling segmentation in Drosophila. *Nature* 276:565-70.
- Li, W. 1983. *Evolution of duplicated genes* Suderland: Sinauer
- Logsdon, J. M., Jr. (1998) The recent origins of spliceosomal introns revisited. *Curr Opin Genet Dev* 8:637-48.
- Lundin, L. G. (1993) Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. *Genomics* 16:1-19.
- Luque, T., Atrian, S., Danielsson, O., Jörnvall, H., Gonzàlez-Duarte, R. (1994) Structure of the Drosophila melanogaster glutathione-dependent formaldehyde dehydrogenase/octanol dehydrogenase gene (class III alcohol dehydrogenase). Evolutionary pathway of the alcohol dehydrogenase genes. *Eur J Biochem* 225:985-93.
- Lyko, F. (2001) DNA methylation learns to fly. *Trends Genet* 17:169-72.
- Lynch, M., Conery, J. S. (2000) The evolutionary fate and consequences of duplicate genes. *Science* 290:1151-5.
- Martin, A. (2001) Is tetralogy true? Lack of support for the "one-to-four rule". *Mol Biol Evol* 18:89-93.
- Martínez, M. C., Achkor, H., Persson, B., Fernàndez, M. R., Shafqat, J., *et al*. (1996) Arabidopsis formaldehyde dehydrogenase. Molecular properties of plant class III alcohol

dehydrogenase provide further insights into the origins, structure and function of plant class P and liver class I alcohol dehydrogenases. *European Journal of Biochemistry* 241:849-857

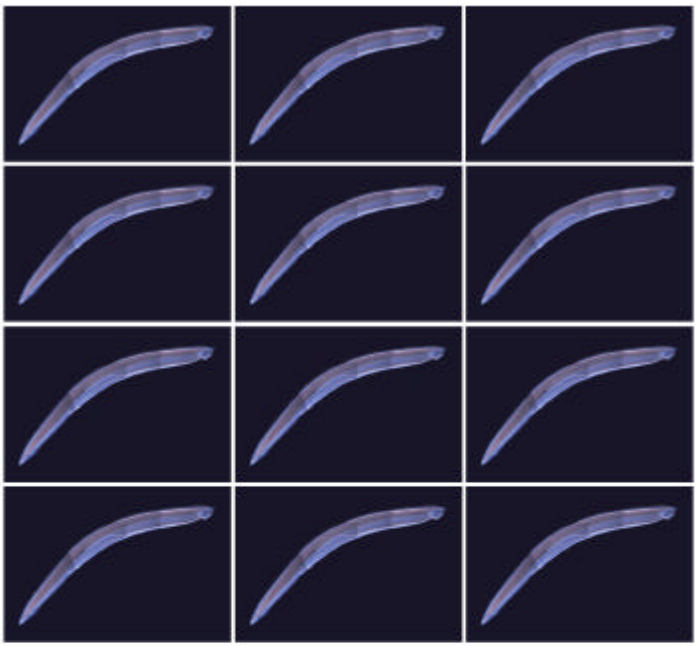
- Martinez, S. E., Vaglenova, J., Sabria, J., Martinez, M. C., Farres, J., Pares, X. (2001) Distribution of alcohol dehydrogenase mRNA in the rat central nervous system Consequences for brain ethanol and retinoid metabolism. *Eur J Biochem* 268:5045-56.
- McEvily, A. J., Holmquist, B., Auld, D. S., Vallee, B. L. (1988) 3 beta-Hydroxy-5 beta-steroid dehydrogenase activity of human liver alcohol dehydrogenase is specific to gamma-subunits. *Biochemistry* 27:4284-8.
- Meyer, A., Schartl, M. (1999) Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol* 11:699-704.
- Moreno, A., Farrés, J., Parés, X., Jörnvall, H., Persson, B. (1996) Molecular modelling of human gastric alcohol dehydrogenase (class IV) and substrate docking: differences towards the classical liver enzyme (class I). *FEBS Lett* 395:99-102.
- Moreno, A., Parés, X. (1991) Purification and characterization of a new alcohol dehydrogenase from human stomach. *Journal of Biological Chemistry* 266:1128-1133
- Mori, O., Haseba, T., Kameyama, K., Shimizu, H., Kudoh, M., *et al.* (2000) Histological distribution of class III alcohol dehydrogenase in human brain. *Brain Research. Jan* 852:186-190
- Muller, H. (1935) The origination of chromatin deficiencies as minut deletions subject to insertation elsewhere. *Genetics* 17:237-252
- Nadeau, J. H., Sankoff, D. (1997) Comparable rates of gene loss and functional divergence after genome duplications early in vertebrate evolution. *Genetics* 147:1259-66.
- Nikoh, N., Iwabe, N., Kuma, K., Ohno, M., Sugiyama, T., *et al.* (1997) An estimate of divergence time of Parazoa and Eumetazoa and that of Cephalochordata and Vertebrata by aldolase and triose phosphate isomerase clocks. *J Mol Evol* 45:97-106.
- Nohara, M., Nishida, M., Yamaguchi, T., Tomiyama, M., Zhang, P., *et al.* (2001) Mitochondrial DNA variation and genetic relationships of Branchiostoma Species from the Pacific and Atlantic Oceans. *Zoological Sciece* 18:447-464
- Ohno, S. 1970. *Evolution by gene duplication* Heidelberg: Springer-Verlag
- Owen, R. 1848. *On the archetype and homologies of the vertebrate skeleton* London: John van Voorst
- Parés, X., Vallee, B. L. (1981) New human liver alcohol dehydrogenase forms with unique kinetic characteristics. *Biochem Biophys Res Commun* 98:122-30.
- Patton, S., Luke, G., Holland, P. (1998) Complex history of a chromosomal paralogy regions: insights from amphioxus aromatic amino acid hydroxylase genes and insulin-related genes. *Mol Biol Evol* 15:1373-1380
- Peralba, J. M., Cederlund, E., Crosas, B., Moreno, A., Julià, P., *et al.* (1999a) Structural and enzymatic properties of a gastric NADP(H)- dependent and retinal-active alcohol dehydrogenase. *J Biol Chem* 274:26021-6.
- Peralba, J. M., Crosas, B., Martínez Susana, E., Julià, P., Farrés, J., Parés, X. (1999b) Amphibian alcohol dehydrogenase: Purification and characterization of classes I and III from Rana perezi.
- Persson, B., Zigler, J. S., Jr., Jörnvall, H. (1994) A super-family of medium-chain dehydrogenases/reductases (MDR): Sub-lines including zeta-crystallin, alcohol and polyol dehydrogenases, quinone oxidoreductases, enoyl reductases, VAT-1 and other proteins. *European Journal of Biochemistry* 226:15-22
- Petrov, D. A. (2001) Evolution of genome size: new approaches to an old problem. *Trends Genet* 17:23-28.
- Philippe, H., Chenuil, A., Adoutte, A. (1994) Can the cambrian explosion be inferred through molecular phylogeny? *Development* 120 (sup):15-25
- Philippe, H., Laurent, J. (1998) How good are deep phylogenetic trees? *Curr Opin Genet Dev* 8:616-23.
- Philippe, H., Lopez, P. (2001) On the conservation of protein sequences in evolution. *Trends Biochem Sci* 26:414-16
- Piatigorsky, J., O'Brien, W. E., Norman, B. L., Kalumuck, K., Wistow, G. J., *et al.* (1988) Gene sharing by delta-crystallin and argininosuccinate lyase. *Proc Natl Acad Sci U S A* 85:3479-83.

- Poss, S. G., Boschung, H. T. (1996) Lancelets (Cephalochordata: branchiostomatidae): how many species are valid? *Israel Journal of Zoology* 42:S-13-S-66
- Ras, J., Van Ophem Peter, W., Reijnders Willem, N. M., Van Spanning Rob, J. M., Duine Johannis, A., *et al.* (1995) Isolation, Sequencing, and Mutagenesis of the Gene Encoding NAD-and Glutathione-Dependent Formaldehyde Dehydrogenase (GD-FALDH) from Paracoccus denitrificans, in Which GD-FALDH Is Essential for Methylotrophic Growth. *Journal of Bacteriology* 177:247-251
- Rhinesmith, H., Schroeder, W., Martin, N. (1958) The N-terminal sequence of the beta chains of normal adult hemoglobin. *J Am Chem Soc* 80:3358-3361
- Rokas, A., Holland, P. W. (2000) Rare genomic changes as a tool for phylogenetics. *Trends in Ecology and Evolution* 15:454-459.
- Sanghani, P. C., Stone, C. L., Ray, B. D., Pindel, E. V., Hurley, T. D., Bosron, W. F. (2000) Kinetic mechanism of human glutathione-dependent formaldehyde dehydrogenase. *Biochemistry . Sep* 39:10720-10729
- Schmidtke, J., Epplen, J. T., Engel, W. (1979) Genome analysis of Amphioxus and speculation as to the origin of contrasting vertebrate genome organization patterns. *Comp Biochem Physiol B* 63:455-8.
- Scopes, R. K. (1983) An iron-activated alcohol dehydrogenase. *FEBS Lett* 156:303-6.
- Shafqat, J., El Ahmad, M., Danielsson, O., Martínez, M. C., Persson, B., *et al.* (1996) Pea formaldehyde-active class III alcohol dehydrogenase: Common derivation of the plant and animal forms but not of the corresponding ethanol-active forms (class I and P). *Proceedings of the National Academy of Sciences of the United States of America* 93:5595-5599
- Sharma, C. P., Fox, E. A., Holmquist, B., Jörnvall, H., Vallee, B. L. (1989) cDNA sequence of human class II alcohol dehydrogenase. *Biochem Biophys Res Commun* 164:631-7.
- Sharman, A., Holland, P. (1996) Conservation, duplication, and divergence of developmental genes during chordate evolution. *Neth J Zool* 46:47-67
- Sharman, A. C. (1999) Some new terms for duplicated genes. *Semin Cell Dev Biol* 10:561-3.
- Shimeld, S. M. (1999) Gene function, gene networks and the fate of duplicated genes. *Semin Cell Dev Biol* 10:549-53.
- Shimeld, S. M., Holland, P. W. (2000) Vertebrate innovations. *Proc Natl Acad Sci U S A* 97:4449-52.
- Sidow, A. (1996) Gen(om)e duplications in the evolution of early vertebrates. *Curr Opin Genet Dev* 6:715-22.
- Skrabanek, L., Wolfe, K. H. (1998) Eukaryote genome duplication - where's the evidence? *Curr Opin Genet Dev* 8:694-700.
- Smit, A. F. (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr Opin Genet Dev* 9:657-63.
- Smith, M., Hopkinson, D. A., Harris, H. (1971) Developmental changes and polymorphism in human alcohol dehydrogenase. *Ann Hum Genet* 34:251-71.
- Smith, N. G., Knight, R., Hurst, L. D. (1999) Vertebrate genome evolution: a slow shuffle or a big bang? *Bioessays* 21:697-703.
- Spring, J. (1997) Vertebrate evolution by interspecific hybridisation--are we polyploid? *FEBS Lett* 400:2-8.
- Strittmatter, P., Ball, E. (1955) Formaldehyde dehydrogenase, a glutathione-dependent enzyme system. *J. Biol. Chem* 213:445-461
- Stromberg, P., Höög, J. O. (2000) Human class V alcohol dehydrogenase (ADH5): A complex transcription unit generates C-terminal multiplicity. *Biochem Biophys Res Commun* 278:544-9.
- Sturtevant, A. (1925) The effects of unequal crossing over at the Bar locus in Drosophila. *Genetics* 10:117-147
- Suga, H., Hoshiyama, D., Kuraku, S., Katoh, K., Kubokawa, K., Miyata, T. (1999) Protein tyrosine kinase cDNAs from amphioxus, hagfish, and lamprey: isoform duplications around the divergence of cyclostomes and gnathostomes. *J Mol Evol* 49:601-8.
- Svensson, S., Hedberg, J. J., Höög, J. O. (1998) Structural and functional divergence of class II alcohol dehydrogenase - Cloning and characterisation of rabbit liver isoforms of the enzyme. *European Journal Of Biochemistry. Jan* 251:236-243
- Svensson, S., Stromberg, P., Höög, J. O. (1999) A novel subtype of class II alcohol dehydrogenase in rodents. Unique Pro(47) and Ser(182) modulates hydride transfer in the mouse enzyme. *J Biol Chem* 274:29712-9.
- Trabesinger-Ruef, N., Jermann, T., Zankel, T., Durrant, B., Frank, G., Benner, S. A. (1996) Pseudogenes in ribonuclease

evolution: a source of new biomacromolecular function? *FEBS Lett* 382:319-22.

- Tweedie, S., Charlton, J., Clark, V., Bird, A. (1997) Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Mol Cell Biol* 17:1469-75.
- Uotila, L., Koivusalo, M. (1974) Formaldehyde dehydrogenase from human liver. Purification, properties, and evidence for the formation of glutathione thiol esters by the enzyme. *J Biol Chem* 249:7653-63.
- Uotila, L., Koivusalo, M. 1989. *Coenzymes and Cofactors. Glutathione. Chemical, Biochemical and Medical Aspects* New York: John Wiley & Sons
- Uotila, L, Koivusalo, M. (1997) Expression of formaldehyde dehydrogenase and S-formylglutathione hydrolase activities in different rat tissues. *Adv Exp Med Biol* 414:365-71
- Venkatesh, B., Ning, Y., Brenner, S. (1999) Late changes in spliceosomal introns define clades in vertebrate evolution. *Proc Natl Acad Sci U S A* 96:10267-71.
- Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., *et al*. (2001) The sequence of the human genome. *Science* 291:1304-51.
- Vinogradov, A. E. (2001) Intron length and codon usage. *J Mol Evol* 52:2-5.
- Vonesch, J. L., Nakshatri, H., Philippe, M., Chambon, P., Dolle, P. (1994) Stage and tissue-specific expression of the alcohol dehydrogenase 1 (Adh-1) gene during mouse development. *Dev Dyn* 199:199-213.
- Watanabe, T., Yoshida, M., Shirai, H. (1999) Effect of light on spawning time in the amphioxus, Branchiostoma belcheri Tsingtauense. *Contributions from the Ushimado Marine Laboratory* 37:1-7
- Watson, J., Crick, F. (1953) Genetical implications of the structure of deoxyribonucleic acid. *Nature* 171:964-967
- Wilkinson, M. F., Shyu, A. B. (2001) Multifunctional regulatory proteins that control gene expression in both the nucleus and the cytoplasm. *Bioessays* 23:775-87.
- Williamson, V. M., Paquin, C. E. (1987) Homology of Saccharomyces cerevisiae ADH4 to an iron-activated alcohol dehydrogenase from Zymomonas mobilis. *Mol Gen Genet* 209:374-81.
- Wolfe, K. H. (2001) Yesterday's polyploids and the mystery of diploidization. *Nat Rev Genet* 2:333-41.
- Xia, L., Liu, J., Sage, C., Trexler, E. B., Andrews, M. T., Maxwell, E. S. (1995) Intronic U14 snoRNAs of Xenopus laevis are located in two different parent genes and can be processed from their introns during early oogenesis. *Nucleic Acids Res* 23:4844-9.
- Yang, Z., Davis, G., Hurley, T., Stone, C., Li, T., Bosron, W. (1994) Catalytic efficiency of human alcohol dehydrogenases for retinol oxidation and retinal reduction. *Alcoholism Clinical and Experimental Research* 18:587-591
- Yang, Z. N., Bosron, W. F., Hurley, T. D. (1997) Structure of human chi chi alcohol dehydrogenase: a glutathione-dependent formaldehyde dehydrogenase. *J-Mol-Biol. 1997 Jan 24;* 265:330-43
- Yasunami, M., Chen, C. S., Yoshida, A. (1991) A human alcohol dehydrogenase gene (ADH6) encoding an additional class of isozyme. *Proc Natl Acad Sci U S A* 88:7610-4.
- Zgombic-Knight, M., Ang Hwee, L., Foglio Mario, H., Duester, G. (1995) Cloning of the mouse class IV alcohol dehydrogenase (retinol dehydrogenase) cDNA and tissue-specific expression patterns of the murine ADH gene family. *Journal of Biological Chemistry* 270:10868-10877
- Zheng, Y. W., Bey, M., Liu, H., Felder, M. R. (1993) Molecular basis of the alcohol dehydrogenase-negative deer mouse. Evidence for deletion of the gene for class I enzyme and identification of a possible new enzyme class. *J Biol Chem* 268:24933-9

.



Departament de Genètica
Facultat de Biologia
Universitat de Barcelona