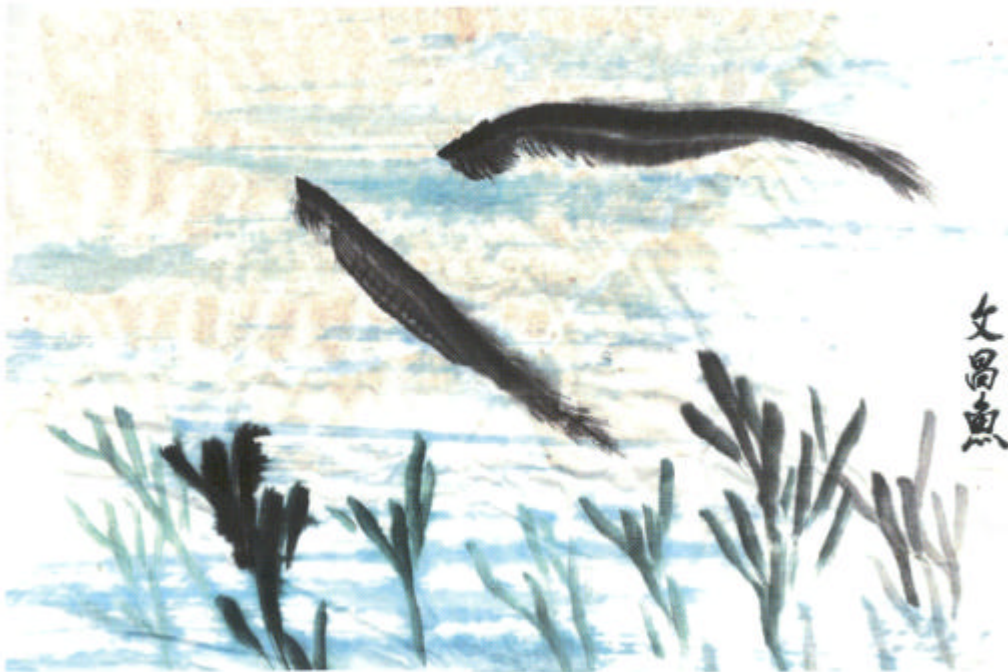


TESI DOCTORAL

**Caracterització molecular i evolutiva
de la família de les alcohol deshidrogenases
de cadena mitjana als cordats**



Wen Chang Yu
peintre local de Qingdao (Xina)
Imatge cedita d'Illbrucki Wala

Cristian Cañestro García
2001



Departament de Genètica
Facultat de Biologia
UNIVERSITAT DE BARCELONA

**Caracterització molecular i evolutiva
de la família de les alcohol deshidrogenases
de cadena mitjana als cordats**

Cristian Cañestro García

2001

Memòria presentada per
Cristian Cañestro García

per optar al grau de
Doctor en Ciències Biològiques

Tesi realitzada sota la direcció de
la Dra. Roser Gonzàlez-Duarte i del Dr. Ricard Albalat Rodríguez
al Departament de Genètica de la Facultat de Biologia
Universitat de Barcelona
Programa de Genètica (Bienni 1996-1998)

Roser Gonzàlez-Duarte

Ricard Albalat

Cristian Cañestro

Barcelona, Octubre de 2001

A mis abuelos ... por luchar tanto en esta vida,
por vuestra entrega, por vuestro cariño ...
sin vosotros esto no habría llegado ni a empezar.

A mi padres, ... por ser como sois,
haberme enseñado a vivir
y sentir que estoy vivo.

A l'Adriana, sense tu res tindria sentit.

Agraïments

Després de quasi cinc anys envoltat de tubs i centrífugues, de molts i molts experiments infructuosos (que ajudaven a omplir la llista de "trofeius") i uns pocs amb final feliç que m'han permès escriure aquesta memòria, allò que enriquirà el meu record seran les experiències personals que he pogut compartir amb tots vosaltres. Durant tot aquest temps moltes són les coses que m'anava dient a mi mateix per citar en aquest moment. Són moltíssimes les vivències que he anat acumulant, i ara que arriba el moment, què difícil és posar-les totes en un paper. El viatge ha estat intens i molts sou els protagonistes que n'heu format part. El laboratori si per alguna cosa s'ha caracteritzat durant aquests anys, és per la quantitat de gent que ens hem fet companyia... però, de vegades, quanta més gent t'envolta és quan més sol pots arribar a sentir-te ... i us vull donar les gràcies a tots per no permetre que això mai passés. La llista de gent és molt llarga, i espero no deixar-me ningú.

Si ens remuntem en el temps, primer em vaig trobar amb la generació dels "veterans", la majoria dels quals estan coneixent món o fins i tot ja estan de tornada. Gràcies, a la Teresa per descobrir-me el "meravellós!!!! *puzzle* de les ADHs"; al Marc, per tenir tanta paciència i guiar-me en les primeres passes pel laboratori; a la Mònica i el Bru, per tenir sempre un moment disponible ple de bons consells; al Salva i al "Pijo-Gonzalone", per mostrar-me les profunditats dels frondosos boscos de la filogènia; a l'Eugènia-et-al., mai oblidaré aquelles pràctiques *estudio-line*; a tots els Evos: Sebas alias "Sport-man", David de Lorenzo, per explicar-nos tantes bones històries de les profunditats dels USA, Agustí, Àurea ... gràcies per tot; a Estela, al Francesc Cebrià i a la Berta per la seva valentia a ultramar; i al Taulé, ara d'aventures pel Pentàgon, al Dave (el catalan-english-man) i al Jose (dels autèntics i genuïns de nou barris) que entre partits de futbol, cerveses i esquiniques sempre m'han aconsellat en tot (*in situs*, RTs, ...). Aquí aprofito per agrair a tots els Okazaki Fragments, ... amb l'Andreu, el Gerard, el Pau i molts altres ... per totes les vesprades de bon futbol, tantes pilotes perdudes, tants partits "guanyats" i sobretot tantes "copes"!!!

En aquest grup he d'incloure dos "veterans senyors", que llavors quan vaig entrar encara s'estaven més temps a la poïata que al despatx. Gràcies, a la Gemma, per ser font d'inspiracions i de respostes, d'interpretacions i de nous interrogants ... ahh! i pels sopars a casa seva, el menjar grec, i sobretot els pastissos de formatge amb "frambuesa"!! Moltíssimes són les coses que m'has ensenyat (i jo sempre tan incrèdul!!), moltes de les quals anaven més enllà de lo estrictament experimental. Gràcies per obrir-me tantes portes, confiar en mi i recolzar-me en tot moment. I al Jordi, sense ell aquesta tesi no hagués estat possible. La seva duresa crítica em va fer veure la cara oculta dels experiments. Després de tantes coses junts i tanta tabarra, gràcies per aguantar-me.

Gràcies a les tres germanes: a n'Amàlia, per ensenyar-me les possibilitats de la PCR i les mil-i-una punyetes associades (encara obro les neveres amb el dit "menyique"!!). Mai oblidaré aquelles reunions maratonianes plenes de fags diferents, bandes polimòrfiques, patrons de Southernns incomprendibles... i al final de tot se sentia una veu dolça que deia: *no sé si m'he perdut, però ... si fesses sa PCR amb s'oligo des braç curt i s'altre específic de s'al-lel, seguida d'un nested amb s'oligos degenerats ...* (jo crec que n'Amàlia té la seva pròpia versió de la frase d'Arquimedes: Dóna'm un oligo i et mouré el món!!). A la Neus, l'esperit jove de MOL: sempre, sempre, sempre i sempre fent costat, donant ànims, aportant confiança carregada d'energia i amb el seu positivisme inesgotable (- *Va home! ... ja veuràs com si que surt!!!*). A la Dra. Valery, ... bé que difícil es parlar de sa majestat!!! Ella ha estat clau en aquesta pel·lícula, ... sempre a prop, de vegades molt a prop, atenta, vigilant des de la foscor... guiant-me pel bon camí (pel camí recte, mooolt recte!!!), ... confiant en mi, ajudant-me, compartint molts bons moments, recomanant-me el millor de la cartellera, pujant el volum de la ràdio per fer-nos còmplices d'aquella cançó que sabem ens posaria la pell de gallina i sobretot mostrant-me el seu *carinyo!* *Coleguita*, no t'ha sigut fàcil interpretar alhora a *Harry el sucio* i a *Bambi* ... una barreja explosiva de sentiments ... gràcies per ser la Reina.

Gràcies Mari, moltíssimes gràcies per tots aquests anys colze a colze, aguantant-me, ficant-te amb mi, comprenent-me i sobretot per preocupar-te i cuidar-me tant. La teva entrega i la teva companyia ha estat crucial per ajudar-me a orientar molts dies que s'iniciaven amb un rumb incert. També a l'Anna Sauqué i l'Amparo, per facilitar-nos tant i tant la feina.

A la Sara, ella sola ha estat molt més que un grup. Cara a cara, fent-nos de mirall i reflectint les nostres alegries i les nostres penes. Sempre tan cuca, ... ja saps ... la *chica Reality bytes*, sempre davant ... infal·lible ... "poco ruido y muchas nueces!!" Gràcies.

Als biotecs (o "secta" o "evofunks" o com vulguin dir-nos) ... perquè per damunt de tot hem estat una família. A la Diana, la *Lady*, la germaneta gran ... gràcies per mostrar-me la teva rebel·lia més tendre, perquè després de tantes "discussions" hem après que les negociacions volen dir que sí i que la realitat va molt més enllà de lo tangible. Les paraules i els fets se'ls endú el vent i l'únic que

queda són els sentiments. Gràcies pel famós café a casa teva ... i no t'oblidis de *Todo sobre mi madre*, que jo encara m'enrecordo que et dec una cervesa. Per compartir tantes coses ... que poc a poc, granet a granet ... han anat consolidant la nostra amistat ... gràcies. A la Raquel, que en la seva estada fugaç va treballar de valent. Al "*pequenyin*" del grup, ara ja Mister Master Jon... ha estat un any que ha donat molt de sí, hem tingut temps per fer moltes coses ... des de maxipreps a les tres de la matinada ... fins cerveses i *bocates* encara més tard!! Gràcies per estar sempre tan atent i fer-te'n càrrec de la situació que t'envoltava. A la Godoy, l'alegria de la huerta (del "políngono"), sempre somrient... per entrar amb nous aires i aportar tanta alegria. Moltes són les manies i punyetes que t'ha tocat aguantar-me ... i les que et falten!! Espero que continuï's divertint-te amb aquest trencaclosques d'ADHs tal i com hem fet fins ara.

Als metalos, Laura quanta raó tens, la dimensió temporal pateix una gran deformació en l'espai que ocupa el nostre laboratori, i com tu vas dir (i la majoria dels que m'envolten ho han pogut comprovar) *en ciència deu minuts poden ser ben bé una hora*[®]; Jordi (ueo ... ueo ...holaaaa!), per tots aquells moments "alternatius" i frases memorables dignes de ser recollides en una enciclopèdia. A la Sílvia, gràcies per totes les batalles que has mantingut contra la rebel·lió de les màquines (hardware, software i el petit "duende" que les desconfigura cada nit) i sobretot per estar sempre disposada a atendre qualsevol dubte sobre les nostres amigues les deshidrogenases.

A l'Olga, gran veterana omnipresent. Sempre fent caliu i animant l'ambient barallant-se per les PCRs, fent que MOL funcioni,... i tots aquells colorits "modelitos" que incitaven comentaris ja des de primera hora del matí, aguantant el tipus ja ben entrada la nit i sempre disposada a tota mena d'interpretacions de les converses menys transcendents. A totes les/(els) humanes, ara separades per l'alçada, però que encara es retroben en el psiquiàtric (per dinar i fer el cafetillo juntes com sempre), a la Kika, "l'eixiríquirikà" companya infatigable de viatge, a l'Anna "ziuziu", l'única que "riu" els meus acudits, a la Magda, la reina dels baculovirus, al Raül, per suportar estoicament entre tanta femina, a la Nats, *enchanté* ... la reina loca de la noche, a la Roser, a la Judit, per tants viatges des de Sant JD, al Miquel, el "piolín" i petit artista de MOL, i a l'Anna Bosch, el darrer fitxatge. A la Paloma, bé ... si hi ha hagut algú que em coneix és ella. Ha plogut molt des de que em passaves els apunts a la carrera,... la més matinerà, la més endreçada, la més constant, la meva amiga... què et podria dir que encara no sàpigues... MOLTES GRÀCIES. A tots els "titis" d'humana, a la Lluïsa, al Dani i a la Susanna, primer per despertar la meva curiositat per la Genètica Molecular durant la carrera, i després per tota l'ajuda i interès mostrat durant aquests anys, plens de bons consells en cada seminari.

Gràcies a tots als Epis, companys i veïns inseparables: a la Mingui, la còmplice d'aquesta amphi-aventura poc arquetípica i molt descefalocordada, a la "J", que també sap que una mica de Jarabe de Palo sempre va bé, a la Mette, amb aquells ulls que irradien un univers de tranquil·litat i felicitat, al David, per tantes sortides alpines que hem compartit, a l'Anna-"Manelillu" supervivent del Bronx de Manlleu, per la seva alegria inesgotable, a la Maria la "modernis" pel seu salero (ooleeeee!), a la Cristina, al Javi, a l'Èlia, al Josep, l'Elisa i tot un seguit de gent jove que tot just ara comença i em fan adonar que els anys no passen en va. A l'Emili (i a la Nati), per tants bons caps de setmana per la costa Brava en busca i captura de *Discoceles tigrina*. Al Jaume Bagunyà, sempre obert a qualsevol pregunta o comentari, i per atendre'm sempre. Als Epis de la saga filogeno-forestal, Iñaki, Paps, Mercé i especialment a la Marta Riutort, per orientar-me entre tantes branques, distàncies i màximes parsimònies. Al grup del Rafa, a la Juani, a la Susana, al Victor don pollo, al Miquel, a l'Albert i al David Bueno per totes les seves ensenyances sobre la reproducció dels peixos zebra i mostra-me que els pistatxos no són realment verds. Al Vela, per la loteria de les pedres, "la que sempre toca"! Als Drosophileros d'Epi, a la Montse Amorós, la primera a assolir la fita, al John "the Scottish-man", a na Marta per ser tan especial, al Manel, el terror de les nenes, al Jordi i al Sergi-windows-advanced. Gràcies al Flo i a la Montse, per venir als seminaris i interessar-se sempre per la feina. Als Drosophileros d'Evo, a l'Eli i la seva ajuda amb els portàtils, al Francesc Mestres, per mantenir les cambres de les mosques impol·lutes i infal·lible reporter del Barça-bulletin, i al Lluís Serra, un gran professor i principal responsable de que m'agradés la Genètica i m'encuriosis per l'Origen de la Vida. Gràcies a la Marta Pascual, al Joan Balanyà i a la Montse Papacèit, per ajudar-me tant amb les pràctiques de Genètica. A tothom d'Evo, a Umberto, per fer-nos companyia a altes hores de la nit, i a tots els titulars, al Julio, pel seu suport informàtic, a l'Elvira, a la Carme, i especialment a la Montse Aguadé, per interessar-se per la nostra feina i venir sempre a llegir tots els pòsters en els congressos que hem coincidit. A l'Amaya i al Ramon, per posar els milions d'As Tes Ces i Ges damunt del pic que tocava, sempre disposats a solucionar els nostres problemes i sobretot per la seva infinita paciència

A les secretàries, a la Isabel, al Xavi, al Sergi, a la Núria i a la Irene, que han fet possible que tots els tràmits i paperassa arribin a bon port.

A la Sra. Ramona i totes les seves companyes, per sempre aconseguir mantenir net tot aquest caos.

He d'agraïr a tots els amics que habiten aquell sorprenent món fora del departament, que molts cops oblidó que existeix. A la Feli, la primera doctora que va traçar-nos el camí a seguir i ànima del grup, a Frau Itziar, la doctora aventurera *trotamons*, a la Sandra, al Pol, a la Núria, al Ramon, a la Mònica, a en Toni, a en Ferran i a l'Ester, per rescatar-me tants caps de setmana i organitzar barbacoes, sortides, esquíades, viatges i sobretot per la vostra amistat. A la Natàlia, per tota la preocupació i amistat que m'ha ofert. A la Mar i al Terro, sempre animant des de les gèlides terres de Finlàndia. Al Toni i al Néstor, per tantes escalades. Al Tomás, més que un amic,... sota l'aigua hem sigut companys de viatges ingràvids plens de colors i emocions indescriptibles, i a terra ferma ... quasi un germà. Gràcies a tots.

De forma més formal voldria agrair a totes aquelles persones i institucions que amb el seu suport tècnic o econòmic han permès que aquest treball es pugui realitzar.

A la Generalitat de Catalunya, per concedir-me una beca de formació d'investigadors CIRIT durant quatre anys.

Al grans equips del servei de seqüències i del servei de microscopia dels Serveis Científic-Tècnics de la Universitat de Barcelona.

A la Imma Rafeques i el seu equip de la unitat de Radioactivitat, per mantenir les instal·lacions sempre a punt i mirar per la nostra salut.

Al Sergi González-Crespo del departament de Biologia Molecular i Cel·lular del Consell Superior d'investigacions Científiques, CID, per la seva inestimable ajuda amb les *in situs* de *Drosophila*.

Als companys del departament de Zoologia, especialment al Xavier Turon i a la Isabel, pel seu assessorament amb els tunicats.

I would like to extend my acknowledgements to all the people abroad that have contributed to this work. To the Swedish team, to Hans Jörnvall, for all his dedication and support, Lars, for all the nice time we enjoyed together and for every time you did something "just for fun". To the English team in Reading, Peter Holland, Seb Shimeld, to Hiroshi Wada, Dave Ferrier, Nee, Nic, Simes, Graham, Rob and Françoise ... for making me stay in the lab so enjoyable and fruitful, and accepting me as a partner the amphioxus' world. To the American team, to Linda and Nick Holland, for your warm hospitality and sharing with us all your knowledge about the pallid anxovy, Bill Jackman, Michael Schubert and Sky, for all those scientific conversations among pools and beers in Iborg-City. To the French team, to Patrick Lemaire for his attention and wise advices, to Claire Hudson, Sebastian, Sylvaine, Laurent, Yass and Michelle, for sharing with me their invaluable knowledge on the ascidians, and take me out to Marseille to taste the delicious bouillabaisse. A Jean Claude Roca i el seu equip de mariners del Observatoire Océanologiques de Banyuls-sur-Mer, per facilitar-nos els espècimens de *B. lanceolatum*. A l'Hèctor, que tant a França com als USA sempre hem compartit bons moments... discussions sobre la veritat de l'ésser, els mecanismes intrínsecs de l'*escombra* i molt menjar mexicà.

Bé, i pel final d'aquesta llarga llista d'agraïments he deixat a aquelles dues persones que li han donat sentit a tot aquest temps i a les quals sempre estaré profundament agraït.

A la Roser, ... els que hem tingut la sort de tenir-la com a directora de tesi, sabem que ella és molt més que això. No solament s'ha encarregat de proporcionar-nos tot allò que ens ha fet falta per realitzar els nostres experiments i arribar a ser doctors, sinó que ha aconseguit el què sobrehumàment ha estat al seu abast. Ha dirigit i reconduït tots els nostres esforços perquè arribin finalment ha tenir un cos, una forma ... una tesi. Ens ha mostrat amb el seu exemple, la seva entrega i la seva dedicació que la humilitat i la constància són les bases de la feina ben feta. Gràcies Roser, per tenir tanta paciència, per no desesperar-te i aguantar la meua tossudesesa, per confiar en mi i recolzar-me en tot moment, per sempre estar al davant i mai perdre el rumb, per ésser la nostra mare aquí al laboratori.

Al Ricard, ... què puc dir-te,... per on començo? No hi ha paraules que puguin agrair-te tot el que has fet per mi. Són àntes els moments que hem compartit, tantes vivències, tantes coses que podria explicar, ... no acabaria mai. Perdó Anna, per segrestar-te'l tants dies fins tan tard, i gràcies per cuidar d'aquest home, que a part de ser el meu director, per damunt de tot és un amic. Gràcies a tots dos per la vostra amistat.

Ricard, et trobaré a faltar!!

*Nothing in Biology Makes Sense
Except in the Light of Evolution*

Theodosius Dobzhansky (1900-1975)

Presentació

Aquesta tesi va sorgir de la convergència entre la llarga història científica de les Alcohol Deshidrogenases en el grup de Molecular del departament de Genètica, i la llavors recent introducció de l'amfiox en el grup d'Epigenètica. La gran diversitat d'alcohol deshidrogenases existents en els vertebrats i la posició estratègica d'aquest organisme en l'evolució dels cordats, ens brindava l'oportunitat d'estudiar l'origen d'aquesta família gènica i els possibles mecanismes evolutius implicats en l'adquisició de les noves funcions.

La memòria d'aquest treball ha estat organitzada de la següent manera:

En la introducció es presenten els diferents àmbits que sorgeixen en el decurs d'aquesta tesi. Primer, la importància de la duplicació gènica en l'evolució, especialment en la dels vertebrats, es descriuen algunes de les bases teòriques que governen el destí dels gens duplicats i es consideren alguns dels paràmetres que relacionen l'evolució del genoma i les expansions de famílies gèniques. A continuació, s'introdueixen les característiques dels dos grups germans dels vertebrats - els cefalocordats i els urocordats - fent especial èmfasi en els avantatges que proporciona el seu estudi. Finalment, es descriu la família de les alcohol deshidrogenases (ADH), fent una anàlisi comparativa dels principals trets de l'ADH de classe 3 envers la resta de classes.

La secció de resultats s'estructura en cinc capítols, cadascun dels quals agrupa un conjunt de resultats en format de publicació científica. Cada capítol ve precedit d'un breu resum on s'exposa de forma concisa les dades més rellevants. Els tres primers recullen de forma cronològica les dades referents a l'estudi pròpiament dit de la família ADH: des de la caracterització de l'ADH a cefalocordats, l'estudi de l'evolució molecular de la família, fins l'estudi comparatiu dels seus patrons d'expressió al llarg del regne animal. En el tercer capítol també es recullen els resultats preliminars de la caracterització i l'evolució de les regions reguladores d'aquesta família mitjançant experiments de transgènesi. El capítol quart és fruit dels experiments duts a terme per establir un protocol d'electroporació en l'amfiox, on es revela l'existència d'activitat β -galactosidasa endògena i es descriu la seva utilitat com a marcador histoquímic del sistema digestiu. Finalment, el capítol cinquè recull una sèrie de dades obtingudes al llarg de tota la tesi, que revelen algunes de les característiques típiques dels gens dels cefalocordats que, si menys no, confereixen al seu genoma certes peculiaritats.

La discussió d'aquesta tesi s'estructura en dos nivells ben diferenciats. La discussió més severa i rigorosa de les dades s'inclou en els diferents articles de la secció de resultats. D'altra banda, en la secció de discussió general, s'ha intentat fer un exercici d'abstracció, en un to distès, on algunes de les idees que apareixen tenen un elevat grau d'especulació. Aquestes consideracions intenten reflectir algunes de les hipòtesis que han anat sorgint en el decurs d'aquest treball, algunes de les quals no tenen encara evidències experimentals suficients que les sustentin, i que de fet han inspirat algunes de les línies de treball que s'estan desenvolupant en el nostre grup.

Thesis Synopsis

My research has been focused on the "Functional and Evolutionary implications of the Alcohol dehydrogenase (*Adh*) gene family expansion during chordate evolution". A single *Adh* class had been reported in invertebrates, whereas at least 7 classes were described in mammals. In this scenario, three research projects have been developed:

i) Evolutionary analysis of the *Adh* gene family expansion, which main goal was to fit this event with the large-scale duplications postulated during the early vertebrate evolution.

ii) Molecular study of the mechanisms underlying the high polymorphism ascribed to cephalochordate genes, using the *Adh* genomic region as a paradigmatic example of this common feature of amphioxus populations.

iii) "Evo-Functional" project, which main objective was to elucidate how the new duplicated genes acquired novel functions by a comparative analysis of the biochemical activities and the expression patterns among different animals (flies, ascidians, cephalochordates, fishes and mammals).

The most significant contributions resulted from this work can be summarized in the next points.

We have characterized the *Adh3* gene (genomic structure and cDNA) of two cephalochordates (*B. floridae* and *B. lanceolatum*) and one urochordate (*C. intestinalis*), as the only *Adh* member present in lower chordates. Our data from protein and gene structures, exon-intron distribution, evolutionary rates, phylogenetical analysis, chromosomal mapping, biochemical activities and expression patterns strongly support that the ADH family expansion occurred 500 million years ago, after the cephalochordate/vertebrate split, probably concurrent with the extensive isoform burst observed before the fish/tetrapode split, rather than through the two rounds of large-scale genome duplications postulated in early vertebrate evolution. Moreover, comparison between *B. floridae* and *B. lanceolatum* ADHs provided the first estimate for a cephalochordate speciation, 190 million years ago, probably concomitant with the opening of the Atlantic Ocean (Cañestro *et al.*, 2002 or chapter 2).

Our data from Southern, small-pool PCR and sequence analysis in the amphioxus *Adh* locus have revealed that the

high level of polymorphism was related to the vast existence of repetitive DNA in their introns (minisatellites and mobile elements). We have shown that repeat instability appeared during the amphioxus lifetime and generated mosaicism. The characterization of recombinant alleles suggested inter-allelic crossover as the main cause of this somatic instability. Finally, most introns contained a novel class of minisatellites, whose distinctive feature was that the repeat subunit spanned over the exon-intron boundaries and generates duplications of AG/GT splice sites. However, these extra splice sites did not seem to compromise functionality, as no aberrant mRNA variants could be detected (chapter 5 and Martínez-Mir *et al.*, 2002). Significant efforts are being performed in the amphioxus, considered the closest living relative of vertebrates and placed in a key position before the large-scale duplications during early vertebrate evolution, in order to identify the orthologue gene of vertebrate gene families. Our findings on the molecular bases underlying the high level of polymorphism together with the mosaic status of the amphioxus genome should be taken into account when trying to evaluate gene copy number to avoid misleading estimations. Moreover, the huge abundance of intragenic repetitive DNA in cephalochordate genes, the description of a novel kind of minisatellites and the first approach to their dynamics, could provide new perspectives for the understanding of the evolution of vertebrate genomes.

The purification and biochemical characterization of the amphioxus ADH let us to confirm that the ADH3 activity was the ancestral situation before the gene family expansion. However, the amphioxus *Adh* showed an expression pattern restricted to the intestine, in contrast with the ubiquitous expression reported in mammals. In other hand, the amphioxus *Adh* expression pattern invoked similarities with other vertebrate ADH classes, which also showed digestive tissue-specific expression domains (liver, stomach, small intestine and pancreas) (Cañestro *et al.*, 2000 or chapter 1). This analysis was extended to *Drosophila*, *Ciona intestinalis* (both with specific expressions) and Zebra-fish (ubiquitous expression), which corroborated our previous results. From all these data we may conclude that after the gene

family expansion the *Adh3* has retained the same biochemical properties whereas its expression has changed to a more general pattern. However, the new duplicated genes have acquired novel biochemical activities by changes in the coding region, whereas their expressions have maintained the “digestive tissue-specificity” (chapter 3).

Based on these results, we are developing a research plan to study the evolution of the of the *Adh* promoters through the family expansion. Introduction of transgenic DNA into *Ciona intestinalis* embryos, a chordate animal model, is allowing us to characterize the cis-regulatory elements that drive the specificity of the pre-vertebrate *Adh* (chapter 3). These elements could be traced in all the new duplicated *Adh* genes, and will help to understand the complex pattern of expression (with a high level of biochemical and expression redundancy) of those vertebrate genes, some of them playing roles as synthesis of retinoic acid (crucial for animal development) (Dalfó *et al.*, 2001), metabolism of alcohols from beverages and synthesis of steroid hormones.

Finally, we have shown that the *in situ* detection of endogenous β galactosidase activity in amphioxus, as a very useful histochemical marker to monitor morphological and functional differentiation of the digestive system, and the posteriorization of the endodermal structures in retinoic-acid treated embryos. Moreover, the discovery of this endogenous activity can help to develop future gene transfer assays, since it

could interfere with the expression of the *lac-Z* reporter gene (Cañestro *et al.*, 2001 or chapter 4).

References

- Cañestro, C., Hjelmqvist, L., Albalat, R., Garcia-Fernández, J., González -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.
- Cañestro, C., Albalat, R., Escrivà, H., González -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., González -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., González -Duarte, R., Albalat, R. (submitted) Minisatellite instability at the *Adh* locus reveals somatic polymorphism in amphioxus.
- Dalfó, D., Cañestro, C., Albalat, R., González -Duarte, R. (2001) Characterization of a microsomal retinol dehydrogenase gene from amphioxus: retinoid metabolism before vertebrates. *Chem Biol Interact* 130-132:359-370
- Martínez-Mir, A., Cañestro, C., González -Duarte, R., Albalat, R. (2002) Characterization of the amphioxus presenilin gene in a high gene-density genomic region illustrates duplication during the vertebrate lineage. *Gene* in press

Abreviacions

Adh o ADH – gen o proteïna de l'alcohol deshidrogenasa

Bf: *Branchiostoma floridae*

Bl: *Branchiostoma lanceolatum*

bp: parells de bases

Bra: gen Brachyury

Ci: *Ciona intestinalis*

Dm: *Drosophila melanogaster*

Dr: *Danio rerio*

DDC: duplicació-degeneració i complementació

h: hores

Hs: *Homo sapiens*

FA: formaldehid

MA: milions d'anys

MDR: deshidrogenasas-reductasas de cadena mitjana

RA: àcid retinòic

RGC: canvis genètics rars

SDR: deshidrogenasas-reductasas de cadena curta

ÍNDEX

Índex

Presentació.....	<i>i</i>
Thesis synopsis.....	<i>iii</i>
Abreviacions.....	<i>v</i>
Índex	<i>vii</i>

INTRODUCCIÓ

1. Duplicacions gèniques: element clau en l'evolució dels vertebrats.....	1
2. El destí dels gens duplicats	4
3. Paràmetres relacionats amb l'expansió de famílies gèniques durant l'evolució del genoma dels cordats	6
3.1 DNA molt repetit.....	6
3.1.1 Repeticions en tàndem: DNA satèl·lit, minisatèl·lits i microsatèl·lits.....	6
3.1.2 Repeticions disperses: elements genètics mòbils.....	7
3.2 Introns.....	7
3.3 Estat de metilació	8
4. Evolució del filum cordats.....	10
4.1 Subfilum Cefalocordats.....	10
4.2 Subfilum Urocordats	14
4.3 Models animals invertebrats per inferir les funcions dels gens	15
5. La família gènica de les alcohol deshidrogenases.....	16
5.1 L'ADH classe 3.....	17
5.2 Altres classes d'ADH	20

<u>OBJECTIUS</u>	23
-------------------------------	----

RESULTATS

CAPÍTOL I	25
Amphioxus alcohol dehydrogenase is a class 3 form of single type and of structural conservation but with unique developmental expression	
CAPÍTOL II	35
Ascidian and Amphioxus <i>Adh</i> genes correlate functional and molecular evolution of the ADH family expansion during vertebrate evolution	
CAPÍTOL III	47
Anàlisi evolutiu dels patrons d'expressió de la família gènica de les alcohol deshidrogenases de cadena mitjana	
CAPÍTOL IV	59
Endogenous beta-galactosidase activity in amphioxus: a useful histochemical marker for the digestive system	
CAPÍTOL V	65
Minisatellite instability at the <i>Adh</i> locus reveals somatic polymorphism in amphioxus	

DISCUSSIÓ GENERAL

L'estudi de l'Evolució de les funcions: "l'EVO-FUN"	79
1. L'EVO-FUN a la família de les ADHs als cordats	79
1.1 Evolució de la regió codificant.....	79
1.1.1 Quina era l'activitat ADH ancestral abans de l'expansió?	79
1.1.2 Processos evolutius implicats en el destí dels gens duplicats	80
1.1.3 La "densitat funcional" com eina per inferir l'Evo-Fun.....	80
1.1.4 Evo-Fun de les proteïnes que formen homodímers.....	81
1.2 Evolució del patró d'expressió dels membres de la família ADH.....	82
1.3 La família ADH: un cas singular.....	84
2. Constància evolutiva de l'ADH3	85
3. L'ADH3, un enzim housekeeping?.....	87
4. L'amfiox com a model animal: un organisme arquetipic amb un genoma ... peculiar	87

<u>CONCLUSIONS</u>	91
---------------------------------	----

<u>REFERÈNCIES</u>	93
---------------------------------	----

INTRODUCCIÓ

1. Duplicacions gèniques: element clau en l'evolució dels vertebrats

Moltes fases de l'evolució dels animals s'han dut a terme mitjançant canvis radicals en els mecanismes de desenvolupament. S'han proposat sis grans canvis com els més importants fins la radiació dels vertebrats: l'origen 1) de la multicel·lularitat, 2) de les dues capes germinals i la simetria radial, 3) de la tercera capa germinal i la simetria bilateral, 4) la inversió de l'eix dorsoventral, i l'origen 5) dels vertebrats i dins d'aquests 6) l'aparició dels gnatòstoms (fig. 1) (Holland, 1998). Moltes d'aquestes transicions, especialment les dues darreres, han anat acompanyades, i possiblement afavorides, per duplicacions gèniques, grans increments en el nombre de gens i adquisicions de noves regions reguladores en cis.

El concepte de duplicació com a mecanisme generador de trets fenotípics va sorgir molt abans de l'era de la genètica molecular, i fins i tot abans del 1953, quan Watson i Crick proposaren l'estructura de la doble hèlix del DNA. Haldane (1932) i Muller (1935) van ser els primers en esmentar la importància evolutiva de la duplicació gènica, suggerint que mitjançant la incorporació de mutacions divergents, eventualment els gens duplicats esdevindrien nous gens. Sturtevant (1925) i Bridges (1936) van ser els primers en descriure duplicacions gèniques al locus *Bar* de *Drosophila* mentre estudiaven entrecruaments desiguals i bandes duplicades dels cromosomes politéncics. Noves evidències a finals dels 50s reafirmarien el concepte anterior. La seqüenciació de proteïnes va permetre reconèixer el parentiu estructural de l'hemoglobina i la mioglobina i proposar per ambdues un origen comú per duplicació (Itano, 1957, Rhinesmith *et al.*, 1958, Braunitzer *et al.*, 1961).

Definitivament als anys 70, estudis d'anàlisi genètic i de l'estructura del genoma varen enfortir l'evidència de la duplicació gènica en el centre de l'escenari de l'evolució animal. Per un costat, Lewis (1978) proposà el mecanisme de duplicació gènica com a generador del Complex Bithorax de *Drosophila*, definit com un conjunt de gens homeòtics relacionats a nivell tan estructural com funcional, que especifiquen la identitat dels segments al llarg de l'eix antero-posterior de la mosca. Lewis

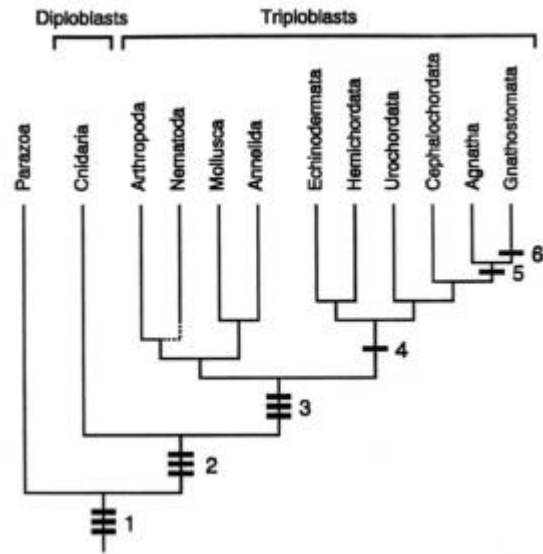


Figura 1. Possibles relacions filogenètiques entre els majors grups d'animals, mostrant les sis grans transicions més importants per l'evolució animal. Transició 1= origen de la multicel·lularitat; 2 = simetria, dues capes germinals, neurones; 3= simetria bilateral, tres capes germinals, cordó nerviós axial, budell buit; 4= inversió eix dorsoventral; 5= cresta neural, nous tipus cel·lulars; 6= mesoderm migratori, apèndixs parells, mandíbules (Holland, 1998).

proposà explícitament la duplicació gènica com a mecanisme evolutiu que proporciona als organismes la possibilitat d'incrementar l'especialització segmental mitjançant l'expansió de l'agrupació homeòtica. Paral·lelament, els treballs de Ohno sobre la mida dels genomes d'organismes molt diversos (Ohno, 1970) semblaven reafirmar la idea que la duplicació gènica era el principal mecanisme per explicar l'evolució dels organismes més complexes. Després de la duplicació, una de les còpies es deslliuraria dels efectes purificadors de la selecció natural i podria acumular canvis prèviament inviables, i eventualment adquirir noves funcions. A diferència de Lewis que limitava la duplicació a gens individuals, Ohno anà més enllà i proposà duplicacions gèniques "en massa" per poliploiditzacions, entre d'altres vies, com a procés més efectiu per adquirir noves funcions, postulant almenys l'existència de dues duplicacions genòmiques durant l'evolució primerenca dels cordats.

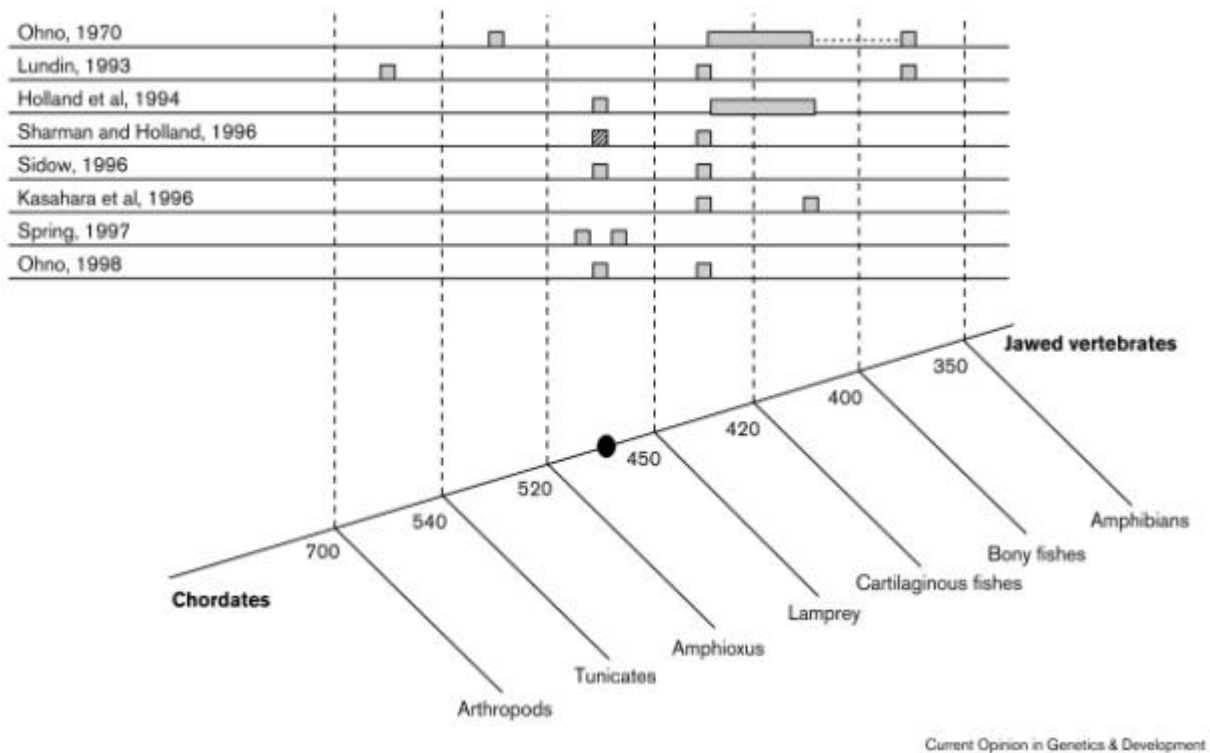


Figura 2. Representació gràfica de les diferents hipòtesis proposades per les duplicacions a gran escala al llarg de l'evolució dels cordats. Les divergències de les espècies es representen esquemàticament. La situació en el temps (MA) de cadascuna de les duplicacions gèniques s'indica amb caixes ombrejades. La caixa ratllada indica una onada de duplicacions gèniques en tàndem. El cercle a 500 MA correspon a l'origen dels vertebrats (Skrabaneck, 1998).

Als anys 90s, es va dur a terme l'anàlisi molecular d'un gran nombre de famílies gèniques en els vertebrats. La cerca d'homòlegs en els prevertebrats va revelar l'existència d'una única còpia en moltes d'aquestes famílies. Així, per molts gens de còpia única a *Drosophila*, es trobaren quatre gens paral·lels a vertebrats (relació 1:4) (Spring, 1997, Sidow, 1996, Meyer & Schartl, 1999), d'acord amb les dues rondes de duplicacions gèniques en el llinatge dels vertebrats. La integració de tots aquests estudis, especialment els realitzats a urocordats, cefalocordats i àgnats, ha permès revisar alguns dels conceptes exposats per Ohno, i resituar les expansions de les famílies gèniques al llarg de la història dels vertebrats (fig. 2). Diversos autors han proposat dues fases d'expansió en el nombre de gens: una primera de tetraploidització, posterior a la divergència dels cefalocordats, i una segona d'extensives duplicacions gèniques en tàndem, posterior a la separació dels àgnats (Holland *et al.*, 1994, Sharman & Holland, 1996, Skrabaneck & Wolfe, 1998).

Un dels punts més controvertits en relació a les duplicacions és si els increments en el nombre de gens s'haurien produït per duplicacions que abarcarien la major part del genoma, o per contra, per múltiples duplicacions de curt abast distribuïdes independentment (Smith *et al.*, 1999). Una de les dades més rellevants en aquest debat és l'existència de regions cromosòmiques paral·leles dins dels genomes de vertebrats (Lundin, 1993), on certs cromosomes contenen els mateixos grups de gens lligats. S'ha suggerit que els quatre grups de regions cromosòmiques paral·leles de l'home i ratolí, representarien reminiscències conservades de les dues o tres rondes de tetraploidització proposades per Ohno (Lundin, 1993). Altres hipòtesis mantenen que aquest lligament es deuria a associacions posteriors de gens per la seva importància funcional, mantingut per la selecció natural degut al seu valor adaptatiu. En aquest cas, les associacions serien el fruit de duplicacions gèniques independents i posteriors translocacions cromosòmiques (Hughes, 1998). Amb les dades actuals es fa

diffícil descartar totalment una de les dues hipòtesis. No totes les famílies gèniques que comparteixen grups de paralogia mostren una història idèntica respecte el temps i ordre de duplicació (Holland, 1999). En contra de la hipòtesi de duplicacions en bloc s'ha argumentat que la majoria de les topologies dels arbres filogenètics de famílies gèniques de vertebrats amb relacions 1:4 no són les esperades d'un procés de dues rondes de duplicacions genòmiques: "(invert, [(vert1,vert2), (vert3,vert4)])" (Hughes, 1999, Martin, 2001). Malgrat tot, la caracterització de pèrdues gèniques aleatòries com un fet prevalent després de les duplicacions (Nadeau & Sankoff, 1997, Wolfe, 2001) i les limitacions inherents als mètodes de filogenia molecular, sensibles a diferències en les taxes evolutives, "long-branch attraction", saturació mutacional i evolució concertada (Felsenstein, 1988; Felsenstein, 1996; Guigo *et al.*, 1996; Philippe & Laurent, 1998, Graur & Li, 2000), expliquen que encara no hi hagi evidències suficients per a rebutjar o afirmar aquesta hipòtesi de les dues rondes de duplicacions genòmiques.

La seqüenciació del genoma humà (Venter *et al.*, 2001, Lander *et al.*, 2001), un dels esdeveniments més importants de la nostra història científica recent, ha obert nous horitzons a la comprensió de l'evolució del genoma envers el nombre total de gens i regions duplicades. Existeixen 1077 blocs duplicats que comprenen 3522 gens diferents. Aquests blocs estan distribuïts aleatòriament pel genoma, i la gran mida d'alguns d'ells juntament amb el gran nombre de gens implicats, semblaria indicar que provindrien de duplicacions en massa, seguides de reordenacions i pèrdues gèniques. Per exemple, un dels segments duplicats més gran correspon a una regió de 20 Mb del cromosoma 2 i una regió de 63 Mb del cromosoma 14, amb un total de 97 i 332 gens, respectivament, dels quals 33 són compartits. La probabilitat que la distribució d'aquesta associació hagi estat a l'atzar és de 2.3×10^{-68} . La majoria de blocs

duplicats presenten sintènia amb subregions cromosòmiques de ratolí, el que implica que el seu origen hauria precedit la divergència entre primats i rosegadors. Futures comparacions amb altres genomes (aus, peixos, àgnats, cefalocordats ...) ajudaran a precisar quan varen tenir lloc aquests esdeveniments.

Finalment, una de les dades més importants del projecte genoma és el nombre de gens que el componen. Inicialment aquest va ésser d'uns 30.000. No obstant, recentment s'ha posat en evidència el poc solapament existent entre els resultats obtinguts pels projectes de Celera i el Consorci Internacional (Hogenesch *et al.*, 2001). Malgrat el nombre final pugui incrementar-se, encara quedarà lluny dels 80.000 a 100.000 gens que s'estimaven suposant dues rondes de duplicacions genòmiques. Per tant el nombre romandrà "relativament" proper als 14.000 de *Drosophila*, o 20.000 de *C. elegans* (*C. elegans* Sequencing Consortium 1998; Adams *et al.*, 2000). Aquest fet genera la nova paradoxa del "valor N", on es proposa que ni la quantitat de DNA (valor C) ni el contingut en gens (valor M) estarien directament relacionats amb la complexitat dels organismes. Aquests arguments ens allunyen de la perspectiva genètica reduccionista al voltant dels gens i ens condueixen cap a una visió global i integradora de les xarxes gèniques com a principal font de complexitat biològica (Claverie, 2001).

En resum, la idea essencial del model d'Ohno de l'existència de grans duplicacions gèniques durant l'evolució dels vertebrats ha estat confirmada, malgrat resten qüestions per resoldre. Entre aquestes, cal destacar els mecanismes moleculars responsables de l'increment del nombre de gens, que explicarien la divergència dels gens duplicats per adquirir noves funcions o ésser silenciats, la formació i manteniment d'agrupacions de gens i la contribució d'aquest increment en l'evolució dels animals, especialment en la radiació dels vertebrats.

2. El destí dels gens duplicats

Les duplicacions gèniques tenen lloc en un individu dins d'una població i quan es produeixen generen dues còpies idèntiques totalment redundants. Els primers efectes incideixen en la dosi gènica. Posteriorment, la divergència entre ambdues còpies serà deguda a: *i*) mutacions en les regions codificants, que alteraran les propietats de les proteïnes sintetitzades, i *ii*) mutacions en les regions reguladores que modificaran els patrons d'expressió. El model clàssic prediu que una de les còpies romandrà a la població amb la funció original, mentre que l'altra probablement desapareixerà per acumulació de mutacions deletèries (procés de "desfuncionalització", *nonfunctionalization*). Ocasionalment la segona còpia quedarà fixada per l'adquisició de mutacions favorables que puguin ser seleccionades positivament (procés de "neofuncionalització") (Ohno, 1970, Li, 1983). Hi ha però dades que semblen indicar que la fracció de gens duplicats que

es mantenen, amb redundància funcional o amb noves funcions, pot arribar a un 50% (Nadeau & Sankoff, 1997, Hughes & Hughes, 1993). Aquest elevat índex de fixació dels gens duplicats discrepa de la baixa freqüència suposada pel model clàssic de desfuncionalització i neofuncionalització. Force i col·laboradors han suggerit un nou procés de duplicació-degeneració-complementació (DDC) que enriqueix el model anterior (fig. 3). Fins llavors, les mutacions deletèries en les regions reguladores havien estat considerades únicament com a part del procés de desfuncionalització. Però si aquest procés degeneratiu es donés de forma complementària entre les dues còpies duplicades, la probabilitat de preservació augmentaria enlloc de disminuir ja que totes dues serien indispensables. El procés de DDC incorpora al model clàssic una primera fase de "subfuncionalització" on la suma de les funcions dels gens duplicats equival a l'ancestral. Per exemple, petits canvis en les regions reguladores podrien produir canvis en els patrons d'expressió de cadascun dels nous gens, de forma que la seva

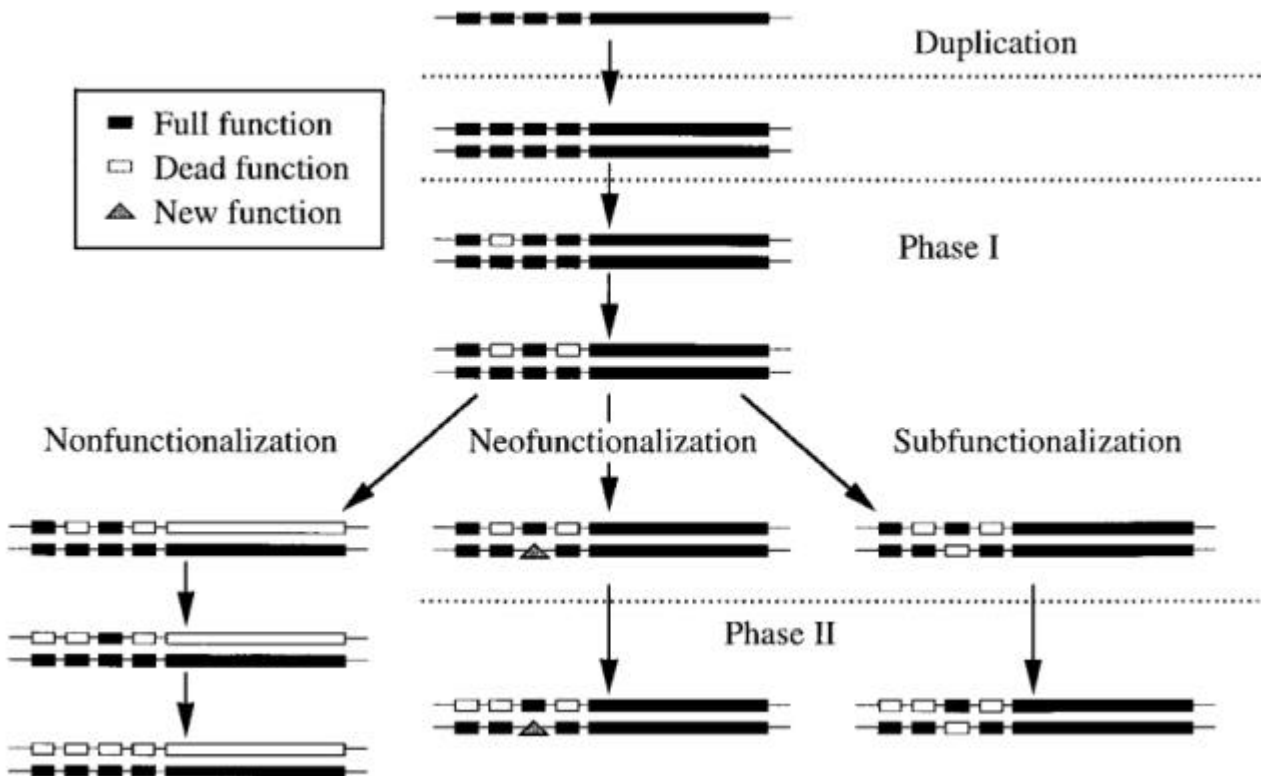


Figura 3. Model d'adquisició de noves funcions dels gens duplicats mitjançant un procés de duplicació-degeneració i complementació (DDC) (Force *et al.*, 1999).

complementarietat equivalgués a l'original. El model DDC incorpora una segona fase durant la qual els gens ja preservats poden refinar i/o derivar de forma aleatòria en noves funcions. Els gens *engrailed* del peix zebra o els gens *hoxa1* i *hoxb1* de ratolí són alguns exemples del model d'evolució per DDC (Force *et al.*, 1999).

Els gens que provenen d'una duplicació presenten la mateixa probabilitat de ser fixats i/o adquirir una nova funció? Hi ha dades que mostren que diferents classes de gens i diferents taxons animals presenten diferències a aquest nivell (Lynch & Conery, 2000). S'ha proposat que els gens metabòlics, particularment els que codifiquen per enzims que realitzen funcions cel·lulars basals ("*housekeeping*") i presenten expressió ubíqua, sols poden ésser fixats per neofuncionalització, mentre que els que presenten patrons restringits, com els gens reguladors del desenvolupament,

podrien a més a més ésser fixats per subfuncionalització (Cooke *et al.*, 1997, Patton *et al.*, 1998). D'altra banda, el fet que la majoria de les mutacions puntuals en gens reguladors (per ex. factors de transcripció) estiguin associades a fenotips deleteris dominants, fa que el temps durant el qual els dos gens es mantindrien de forma redundant s'allargui en relació a gens que codifiquen per enzims (Shimeld, 1999). El destí d'un gen duplicat també depèn de si la duplicació ha afectat a un gen individual o per contra ha englobat una part més gran del genoma (poliploidització), donat que els gens es troben immersos en xarxes, i les seves funcions han de ser considerades en el marc de la interacció de tot els seus components. Després d'una duplicació genòmica, els valors relatius que caracteritzen una xarxa gènica es mantenen constants, incrementant així les oportunitats de supervivència d'ambdues còpies (Shimeld, 1999).

3. Paràmetres relacionats amb l'expansió de famílies gèniques durant l'evolució del genoma dels cordats

L'increment en el nombre de còpies d'un gen és conseqüència d'un o més processos de duplicació del DNA, els quals poden implicar fragments genòmics de mida molt diversa: (1) duplicacions gèniques regionals, de subregions cromosòmiques o cromosomes sencers, i (2) duplicacions de tot el genoma o poliploidies. Els principals mecanismes moleculars que poden donar lloc a duplicacions regionals són l'entrecreuament desigual entre cromàtides homòlogues o entre cromàtides germanes. Aquests intercanvis "erroris" es veuen afavorits per la presència de DNA repetit, que permet l'aparellament de regions no homòlogues. Cal esmentar que les duplicacions de DNA han estat un factor molt important en l'evolució de la mida i l'arquitectura dels genomes, així com per a la distribució i estructura dels gens. L'increment de DNA durant l'evolució dels cordats ha estat acompanyat de mecanismes sofisticats que ajuden a mantenir l'estructura genòmica i contribueixen a un control més precís de l'expressió gènica.

3.1 DNA molt repetit

La mida del genoma haploide d'un organisme es defineix com el valor C, i aquest és constant per cada espècie. L'observació de les grans diferències de mida dels genomes de diferents animals, fou el que portà a Ohno a suposar grans duplicacions a l'inici de l'evolució dels cordats. La manca de correspondència entre la mida del genoma, el nombre de gens i la complexitat dels organismes ha donat lloc a la paradoxa del valor C, que tracta de descriure l'excés aparent de DNA dels organismes en relació a les suposades necessitats funcionals. (Charlesworth *et al.*, 1994). L'increment de la mida dels genomes, especialment en els vertebrats, semblaria clarament relacionat amb grans quantitats de DNA repetit. D'aquesta manera, tots els mecanismes mutacionals derivats de la dinàmica del DNA repetit haurien tingut conseqüències evolutives crucials en la construcció dels genomes actuals (fig. 4).

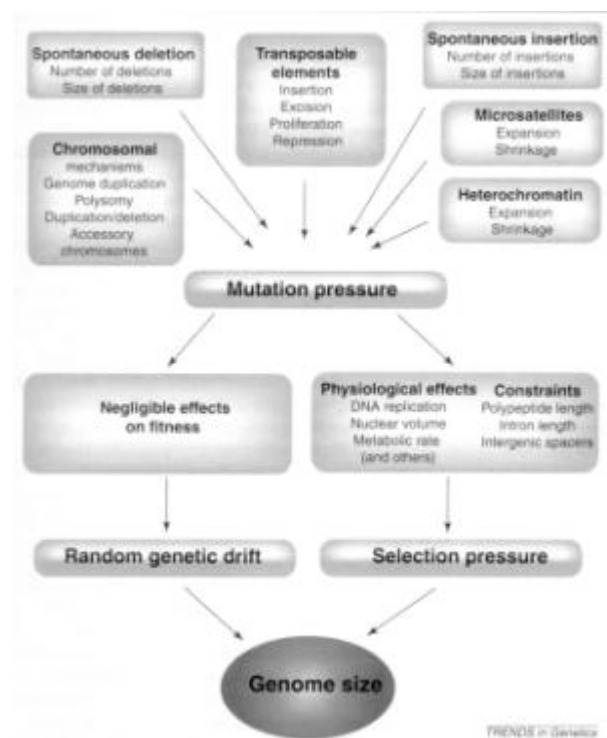


Figura 4. Les principals forces que afecten l'evolució del genoma (Petrov, 2001).

El DNA repetit es pot classificar de forma genèrica en dos grans grups: repeticions en tàndem i repeticions disperses.

3.1.1 Repeticions en tàndem: DNA satèl·lit, minisatèl·lit i microsatèl·lit.

Els DNAs satèl·lits són seqüències curtes, de menys de 10 bp fins a més de 300 bp, repetides en tàndem i freqüentment concentrades en grans blocs prop dels centròmers i telòmers. Els microsatèl·lits i minisatèl·lits són repeticions en tàndem però disperses per tot el genoma. Els microsatèl·lits tenen subunitats de repetició curtes de 2 a 6 bp i els minisatèl·lits tenen subunitats més llargues de 15 a 500 bp. Els *loci* que presenten aquests tipus de repeticions en tàndem mostren un elevat grau de polimorfisme poblacional, degut a la seva variabilitat en el nombre de repeticions. Els micro- i minisatèl·lits per la seva informativitat han estat molt emprats en el camp de la genètica forense, mapatge genètic, i per l'estudi de la diversitat genètica i estructura de les poblacions naturals. Malgrat encara manca molt per desentrellar sobre els processos que generen i mantenen aquesta variabilitat,

s'han proposat dos mecanismes principals per explicar la inestabilitat de les seqüències repetides en tàndem: l'entrecruament desigual inter- o intra-al·lèlic i el *slippage* de la DNA polimerasa en la replicació del DNA (Debrauwere *et al.*, 1997, Jeffreys *et al.*, 1994). La variabilitat dels minisatèl·lits es produeix de forma majoritària en la línia germinal durant la formació dels gàmetes, i en menor grau de forma somàtica durant la vida de l'animal (Buard *et al.*, 2000). Les mutacions en la línia germinal, més abundant en els mascles (Jeffreys *et al.*, 1988), freqüentment impliquen transferències inter-al·lèliques complexes de subunitats de repetició durant la meiosi (Jeffreys & Neumann, 1997). Aquestes mutacions es troben esbiaixades cap un dels extrems de la seqüència i regulades per elements flanquejants (Jeffreys *et al.*, 1994). D'altra banda, la inestabilitat somàtica sembla dur-se a terme per mecanismes diferents, implicant *slippage* i entrecruaments intra-al·lèlics (Jeffreys *et al.*, 1994). Des d'un punt de vista evolutiu, la presència de repeticions en membres d'una família gènica té conseqüències importants ja que incrementa els efectes de l'evolució concertada (Graur & Li, 2000). Això implica que els diferents membres de la família gènica no evolucionen de forma independent, si no que intercanvien informació entre ells produint un efecte d'homogenització intra-familiar. Les taxes de conversió gènica i d'entrecruament desigual, els dos processos principals que promouen l'evolució concertada, es veuen fortament influenciades per factors com el nombre i l'ordre de repeticions, l'estructura interna i restriccions funcionals de la unitat de repetició, i els processos selectius i no-selectius a nivell poblacional (Graur & Li, 2000). Encara més, l'evolució concertada pot tenir un efecte sever en el destí dels dos gens duplicats, i s'han descrit casos on una de les dues còpies duplicades, després d'haver degenerat per processos mutacionals en un pseudogèn, ha recuperat la funcionalitat mitjançant la transferència d'informació de l'altre gen encara actiu (Trabesinger-Ruef *et al.*, 1996). Així, gràcies a l'evolució concertada la segona còpia del gen podria romandre més temps amb la possibilitat d'adquirir una nova funció.

3.1.2 Repeticions disperses: elements genètics mòbils

Són seqüències repetides i disperses en el genoma, d'una mida generalment més llarga que les repeticions en tàndem, i que es

caracteritzen per la seva capacitat de canviar de lloc. Van ser descoberts als anys 50 per Barbara McClintock, malgrat els seus treballs no foren reconeguts amb el premi Nobel fins l'any 1983. Els elements genètics mòbils presenten la propietat de poder-se copiar, sortir d'on són i inserir-se en un nou locus, el que es coneix com a transposició. Aquesta pot ser replicativa o conservativa segons hi hagi o no duplicació acoblada a la seva mobilització. Alguns elements transponibles estan delimitats per seqüències repetides i invertides, i disposen de la seva pròpia maquinària per la transposició (transposasa, transcriptasa reversa,...) mitjançant una molècula intermediària que pot ser RNA o DNA. El lloc d'inserció és en general aleatori, malgrat en alguns casos es reconeix una diana específica d'inserció. Existeixen elements transponibles no-autònoms que han perdut la maquinària necessària per la transposició i fan servir la d'altres elements per mobilitzar-se. Els elements transponibles es classifiquen en diferents famílies segons la seva estructura interna i les seqüències repetides que els delimiten (per una revisió Smit, 1999).

Entre els molts efectes proposats dels elements transponibles sobre l'evolució i expressió dels gens, caldria destacar la seva implicació en reordenacions genòmiques (inversions, translocacions, duplicacions i grans insercions/deleccions) de forma directa, arrossegant fragments de DNA en la pròpia transposició, o indirecta, incrementant la probabilitat de recombinació entre dues zones que prèviament no compartien cap homologia. D'altra banda, insercions d'elements mòbils en un dels membres d'una família gènica pot fer decreixer la taxa de conversió gènica, i per tant afavorir la divergència entre els gens duplicats. Finalment, els elements mòbils poden modificar l'expressió gènica, bé per inserció enmig de regions reguladores o bé per l'addició de noves regions reguladores arrossegades en la transposició.

3.2 Introns

Els introns són seqüències intercalades en la regió codificant dels gens que durant la maduració del RNA són eliminades (Bergert *et al.*, 1977). Alguns autors han proposat que els introns eren ancestrals i s'han anat perdent en diferents llinatges (teoria dels *introns primerencs*). D'altres mantenen que els introns no existien inicialment i que s'han inserit posteriorment (teoria dels

introns tardans) (per una revisió Logsdon, 1998). L'esbiaix en la distribució dels introns respecte a les fases de la pauta de lectura i la correlació de les seves posicions amb els límits dels dominis proteics es consideren evidències a favor de l'antiguitat dels introns. D'altra banda, la distribució filogenètica dels introns en molts casos suggereix que hi ha hagut guanys i pèrdues d'introns afavorint el seu origen tardà. La presència o absència d'introns específics entre diferents organismes ha estat utilitzada com a marcador filogenètic (Venkatesh *et al.*, 1999). De fet, els anomenats "canvis genètics rars" (RGCs) que inclouen les insercions-deleccions d'introns, integracions d'elements mòbils, canvis en l'ordre de gens mitocondrials o dels cloroplasts, les duplicacions gèniques o els canvis en el codi genètic, són marcadors filogenètics complementaris amb un enorme potencial per a la sistemàtica molecular.

El DNA no codificant extragènic ha colonitzat molt més ràpidament que l'intragènic (introns) el genoma animal. La mida i el nombre d'introns té un cost metabòlic ja que la seva transcripció implica una despesa energètica; per tant estan sotmesos a restriccions selectives. S'ha proposat que la mida dels introns juga un paper important en la regulació de l'expressió gènica a dos nivells: primer, modulant la transcripció mitjançant la seva influència en el grau de corbatura i condensació del DNA, i segon, intervenint en la regulació pre- o post-transcripcional (Vinogradov, 2001). Es creu que en la transició de la unicel·lularitat a la multicel·lularitat, el control necessari per mantenir els gens silenciats es va assolir mitjançant la condensació de la cromatina. Per aquest motiu, en organismes multicel·lulars, el nombre i la mida dels introns tendeix a créixer (Logsdon, 1998, Alvarez-Valin *et al.*, 1998). Els vertebrats també presenten la mateixa tendència respecte als invertebrats, possiblement relacionada amb les diferències en les restriccions selectives dels seus cicles biològics, i al necessari desenvolupament en els primers d'una maquinària genètica molt més complexa en el control de l'expressió gènica: *i)* gràcies al *splicing* alternatiu, un únic gen pot codificar per més d'una proteïna; per exemple es creu que en l'home hi hauria *splicing* alternatiu en més del 30% dels gens (Hanke *et al.*, 1999); *ii)* els introns poden contenir elements reguladors; *iii)* o incloure altres gens interns i d'aquesta manera conferir una expressió coordinada (Xia *et al.*, 1995); i finalment, *iv)* la seva presència disminueix el risc de recombinació

il·legítima amb gens paràlegs o pseudogens, malgrat d'altra banda promou la creació de nous gens per reordenacions exòniques (Krickler *et al.*, 1992). S'ha descrit que la selecció potencia l'increment de la mida dels introns, probablement afavorint d'aquesta manera recombinacions intragèniques que augmentarien l'eficàcia selectiva (Comeron & Kreitman, 2000).

3.3 Estat de metilació

Malgrat la metilació hagi estat descrita en els quatre grups principals d'eucariotes (protistes, plantes, fongs i animals), aquesta modificació del DNA no és comuna a tots els organismes. Durant molt de temps, *Drosophila* havia estat considerat l'organisme arquetípic que podia viure sense metilació de DNA (Bird & Taggart, 1980). El fet que *C. elegans* i el llevat també compartissin aquesta característica, va dur a la hipòtesi que la metilació era un mecanisme necessari en genomes complexos per silenciar les transcripcions espúries, i que només genomes molt compactes, com el de *Drosophila*, podrien prescindir d'aquest mecanisme. Investigacions recents han mostrat l'existència d'una maquinària de metilació a *Drosophila* (metiltransferases i proteïnes d'unió a grups metil) que possiblement actuaria de forma puntual durant el desenvolupament sobre dianes molt concretes (Lyko, 2001).

En vertebrats, els dinucleòtids CpG són els substrats prevalents per la metilació. El 70-80% de CpGs a mamífers es troben metilades i la resta de CpGs, no metilades, formen agrupacions o illes aprop del primer exó de molts gens. Clàssicament, s'han distingit dos patrons de metilació en diferents organismes: el típic dels invertebrats, restringit a certes fraccions del genoma, i el patró de distribució uniforme per tot el genoma, característic dels vertebrats (Tweedie *et al.*, 1997). La transició entre els dos patrons semblaria haver-se donat a la base dels cordats, on els ascidis i amfioxos encara presenten un patró típic "d'invertebrat" amb fraccions comparables de DNA metilat i no metilat. S'ha demostrat que existeix una clara correlació entre l'existència i la distribució de DNA repetit i el grau de metilació.

El patrons de metilació s'estableixen i es mantenen mitjançant la maquinària de metilació. Aquesta permet que aquest procés sigui reversible i dinàmic, remodelant-se específicament en cada teixit i moment del desenvolupament. La metilació ha estat associada a diverses funcions: 1) repressió de la transposició d'elements

Introducció-3. Paràmetres relacionats amb l'expansió de famílies gèniques durant l'evolució del genoma dels cordats

endògens, 2) supressió de la recombinació homòloga, 3) inactivació de gens, com la inactivació del cromosoma X en les femelles de mamífers, 4) senyal d'herència parental reguladora de la transcripció per "*imprinting*", i 5) silenciament de l'expressió basal de gens amb poc o gens control transcripcional, gens

"*housekeeping*" (Colot & Rossignol, 1999). Darrerament s'ha suggerit que gràcies a que la metilació permet silenciar transcripcions inapropiades, els vertebrats són l'únic tàxon animal capaç de mantenir un nombre tant elevat de gens (Colot *et al.*, 1999)

4. Evolució del fílum cordats

Els vertebrats juntament amb els cefalocordats (amfiox) i els urocordats (ascidis) formen el fílum dels cordats. Les quatre característiques distintives que presenten *versus* els altres fila són: notocordi en almenys algun dels estadis del desenvolupament, cordó nerviós tubular i buit situat en la part dorsal, fenadures branquials faríngees i cua postanal.

Avui s'accepta de forma general que els vertebrats (en sentit ampli incloent tots els craniats) van evolucionar a partir d'ancestres invertebrats cordats (fig. 5). Endemés, es creu que l'aparició i radiació dels vertebrats està totalment associada a l'increment del nombre de gens generats mitjançant duplicacions a gran escala que s'haurien produït a la base de la seva evolució. No obstant, no s'han trobat "gens innovadors" causants de les diferències del patró de desenvolupament entre invertebrats i vertebrats. La majoria de gens descrits provenen de gens ortòlegs ja presents en els prevertebrats, i l'estudi molecular comparatiu de l'expressió d'aquests gens revela

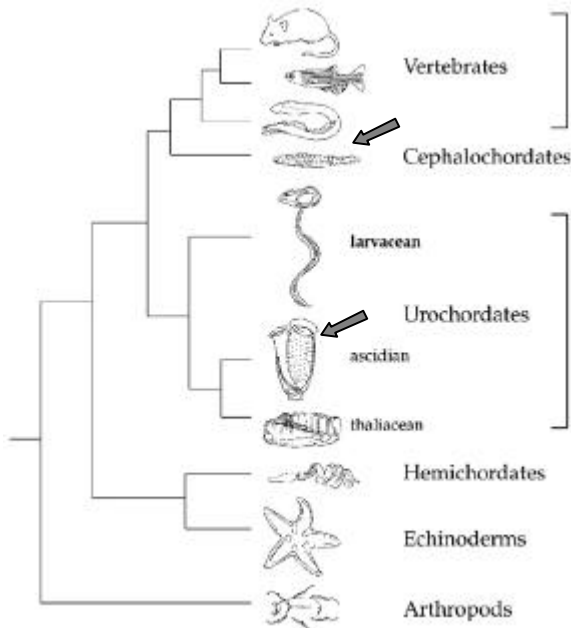


Figura 5. Relacions filogenètiques entre els cordats i altres deuteròstoms, amb un protòstom com a grup extern. La topologia de l'arbre, basada en la dades de rDNA (Wada i Satoh, 1994), mostra els ascidians, dins els urocordats, i els cefalocordats a la base dels vertebrats (fletxes). (Bassham i Postlethwait, 2000)

homologies entre les diferents regions corporals (Holland & Holland, 1999). Així, moltes de les estructures i característiques que prèviament s'associaven exclusivament als vertebrats, ja es troben dissenyades en els procordats i l'origen dels vertebrats s'entén avui com una transició gradual, més que un fet abrupte (Graham, 2000, Shimeld & Holland, 2000, Holland & Chen, 2001). Un dels punts més importants en aquesta transició fou el pas d'una vida sedentària a una vida depredadora mòbil amb l'aparició d'un cap diferenciat. Amb aquest canvi arribaren tres avenços crucials pel subsegüent èxit dels vertebrats: 1) perfeccionament del sistema nerviós associat a l'aparició d'òrgans sensorials per localitzar la presa i el seu entorn, 2) la formació d'un esquelet cranial ben desenvolupat, on més tard apareixeran les mandíbules per capturar les preses, i 3) una bomba muscular i una circulació branquial eficaç per un intercanvi gasós d'alt rendiment.

4.1 Subfílum Cefalocordats.

Són un petit grup d'animals marins sense ulls, sense dents, sense gairebé res, que recorda a un pàl·lid filet d'anxova animat (Gee, 1994) (fig. 6). El seu cos prim està lateralment comprimit, amb una llargada de fins a uns 7 cm. Són coneguts popularment com "llancetes", per tenir un aspecte lanceolat com les puntes de les llances, o "amfiox" (grec: amphí, ambdós extrems, oxys, punxagut).



Figura 6. Fotografia d'un amfiox adult de l'espècie *Branchiostoma lanceolatum*. La longitud de l'animal és d'uns 2 cm.



Figura 7.1. L'amfiox era un dels plats que apareixien en el menú de la festa del 60é aniversari de E. Haeckel (Stokes i Holland, 98)

Van ser descrits per primer cop al 1744 per l'alemany Pallas, creient que eren mol·luscs. Cap al 1830 Costa anticipà la visió moderna de l'amfiox com un vertebrat basal. A la segona meitat del segle XIX, Kovalevskii mostrà que el desenvolupament primerenc de l'amfiox era el típic d'un invertebrat (semblant als eriçons de mar), però donava lloc a una criatura del tipus vertebrat. Finalment Haeckel, inspirat en aquestes idees, adoptà l'amfiox amb especial veneració com a figura principal en la seva teoria de "l'ontogènia recapitula la filogènia", on representava el pas entremig d'invertebrats i vertebrats. La devoció de Haeckel per l'amfiox entrà en conflicte amb els defensors acèrrims de la teoria que l'ancestre invertebrat dels vertebrats era del tipus d'un cuc segmentat. Molts han estat els debats i discussions sobre l'origen dels vertebrats al llarg de la història. L'amfiox ha jugat diferents papers, des de ser simplement un protomol·lusc que havia divergit abans de l'aparició dels cucs segmentats, directes ancestres des vertebrats, fins a ser un vertebrat degenerat que havia perdut el cap, la columna vertebral i moltes de les característiques típiques dels vertebrats (per una revisió amena llegir Gee, 1996). Avui en dia, estudis de biologia molecular han

mostrat que l'amfiox és el parent viu invertebrat més proper als vertebrats. Presenta un pla corporal típic d'un cordat, i se l'ha proposat com un organisme arquetípic no molt diferent a l'ancestre de tots els vertebrats. El registre fòssil ha revelat restes com la Pikaia, pertanyent al Càmbric de Burgess Shale, amb un aspecte molt semblant als actuals cefalocordats. Malgrat tot, els cefalocordats i els vertebrats han estat evolucionant de forma separada durant almenys 450 MA, temps suficient per haver guanyat i perdut caràcters comuns i per que s'hagin produït algunes disjuntives entre les homologies a nivell genètic i morfològic.

Existeixen aproximadament unes 25 espècies de cefalocordats distribuïdes mundialment en zones tropicals i mars temperats (Poss & Boschung, 1996), agrupades en dos gèneres, *Branchiostoma* i *Epigonichthys* (= *Assymmetron*). Molts dels treballs clàssics sobre la biologia de l'amfiox s'han fet en l'espècie *Branchiostoma lanceolatum*, típica de les costes europees atlàntiques i mediterrànies. Les altres dues espècies més estudiades són *Branchiostoma floridae*, "fàcilment" accessible a les platges d'Old Tampa Bay (Florida), i *Branchiostoma*

belcherei, a les costes xineses. La seva taxonomia està basada en caràcters morfològics i hi ha molts poques dades moleculars (Poss & Boschung, 1996). De forma anecdòtica, aquests animals es fan servir pel consum humà (fig. 7.1), frescos o assecats, i s'han registrat pesques de fins a 35 tones en una platja d'una milla d'ample per 6 de llargada a la Xina. A Jamaica s'han observat poblacions de fins a 5000 individus per metre quadrat, i en algunes parts del Brasil deixen "pasturar" els pollastres en les platges perquè s'alimentin d'amfiox.

Malgrat normalment viuen sedentaris ensorrats deixant el cap a l'exterior, també poden desplaçar-se nedant fent servir moviments ondulatoris. Tenen un notocordi prominent des de la cua fins l'extrem més anterior (fig. 7.2). Aquesta és la característica que els dona el nom de cefalocordats. El notocordi té funcions esquelètiques i està envoltada d'una matriu extracel·lular que l'ajuda a mantenir la forma. L'amfiox reflecteix un marcat metamerisme segmentat en blocs de musculatura lateral en forma de V, miòtoms, que es distribueixen al llarg de tot el cos. Pel damunt del notocordi corre un tub nerviós buit, eixamplant-se en una petita vesícula cerebral en la part anterior. Anàlisis moleculars comparatius dels gens que s'expressen en aquesta regió han revelat homologies entre vertebrats i procordats, les quals permeten entreveure les parts del cervell que són innovacions dels vertebrats, i a partir de quines estructures s'han desenvolupat (per una revisió Shimeld & Holland, 2000).

S'alimenten filtrant el fitoplàncton de l'aigua, el qual entra per la boca, travessa un complex aparell faríngi ciliat, surt a l'atri per les fenadures branquials, i finalment és expulsat a l'exterior per l'atriopore. Les partícules d'aliment són atrapades per una substància mucosa segregada per l'endostil, òrgan equivalent a la glàndula tiroides dels vertebrats. L'aliment passa per un solc

epifaringi dorsal cap a l'esòfag i l'intestí mitjà, on començarà a ser digerit. En la unió de l'esòfag i l'intestí apareix un diverticle que es projecta cap a la part anterior per la dreta de la faringe, el qual té funcions d'emmagatzematge de nutrients i secreció d'enzims a l'intestí. Rep el nom de cec hepàtic, i s'ha proposat com a òrgan equivalent al fetge i pàncreas de vertebrats malgrat el seu origen blastodèrmic no sigui el mateix. L'aliment passa després de l'intestí mitjà a la cavitat iliocolònica on acaba de ser digerit ajudat pels moviments rotatoris proporcionats pel tracte ciliat de la paret intestinal. El material que no s'ha digerit passa a l'intestí posterior on és compactat i expulsat per l'anús.

L'amfiox té un sistema vascular ben desenvolupat. Fa l'intercanvi gasós en petits capilars a les fenadures branquials, i la sang és distribuïda posteriorment per les artèries dorsals i retornada per les venes ventrals. No hi ha un cor especialitzat, i la sang és bombejada per un sinus contràctil venós i per l'artèria endostilar. Els cefalocordats són animals diòics i alberguen les gònades a l'atri. L'espècie *Branchiostoma floridae* es troba a la costa de Florida i és la més estudiada a nivell reproductiu. Durant l'estiu és quan assoleixen la maduració i tots els individus de la població alliberen de forma sincrònica els gàmetes a la posta del sol. Aquest fenomen es repeteix aproximadament cada dues setmanes. La fecundació és externa. Segons després que l'oòcit hagi estat fecundat, una membrana externa s'expandeix deixant un gran espai perivitel·lí. El zigot comença a dividir-se de forma radial i holoblàstica per formar la blastula (3.5 h) (fig. 8). La gastrulació (5.5 h) té lloc per invaginació, i el celoma anterior, miòmers i notocordi es formen per enterocèlia. Les gastrules tardanes comencen a moure's impulsades per moviments ciliars, i a les 10 hores en estadi de nèurula eclosionen trencant l'embolcall vitel·lí, i resten de forma planctònica. La nèurula comença a créixer al llarg de l'eix antero-posterior,

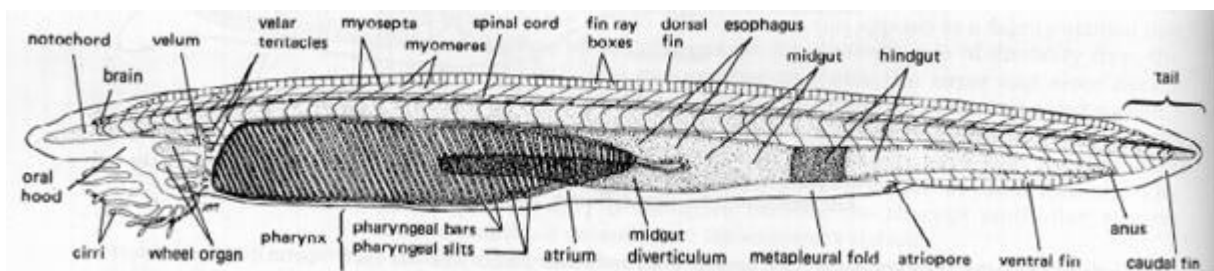


Figura 7.2. Representació esquemàtica d'un amfiox on s'indiquen les diferents estructures de la seva anatomia.

desenvolupant els somites. La larva, amb una forma ja ben allargada, obre la boca a les 38 hores en el costat esquerra del cos, mentre que les fenadures branquials es van obrint progressivament al cantó dret. Aquesta marcada asimetria dreta-esquerra en el desenvolupament, també present a l'adult (els blocs de musculatura, gònades i nervis es troben desplaçats mig miòtom amb el seu company, l'anus està lleugerament desviat cap el cantó esquerra i el cec hepàtic es situa a la dreta de la

faringe), és un dels trets típics de l'amfiox i un dels punts candents en la discussió del seu suposat caràcter arquetípic. Finalment, la metamorfosis comença cap a les 5 setmanes, amb l'aparició de la segona línia branquial i subsegüent migració cap a l'esquerra de l'animal, amb el desplaçament de la boca a l'extrem anterior, l'aparició de les estructures bucals, i amb el desenvolupament del cec hepàtic.

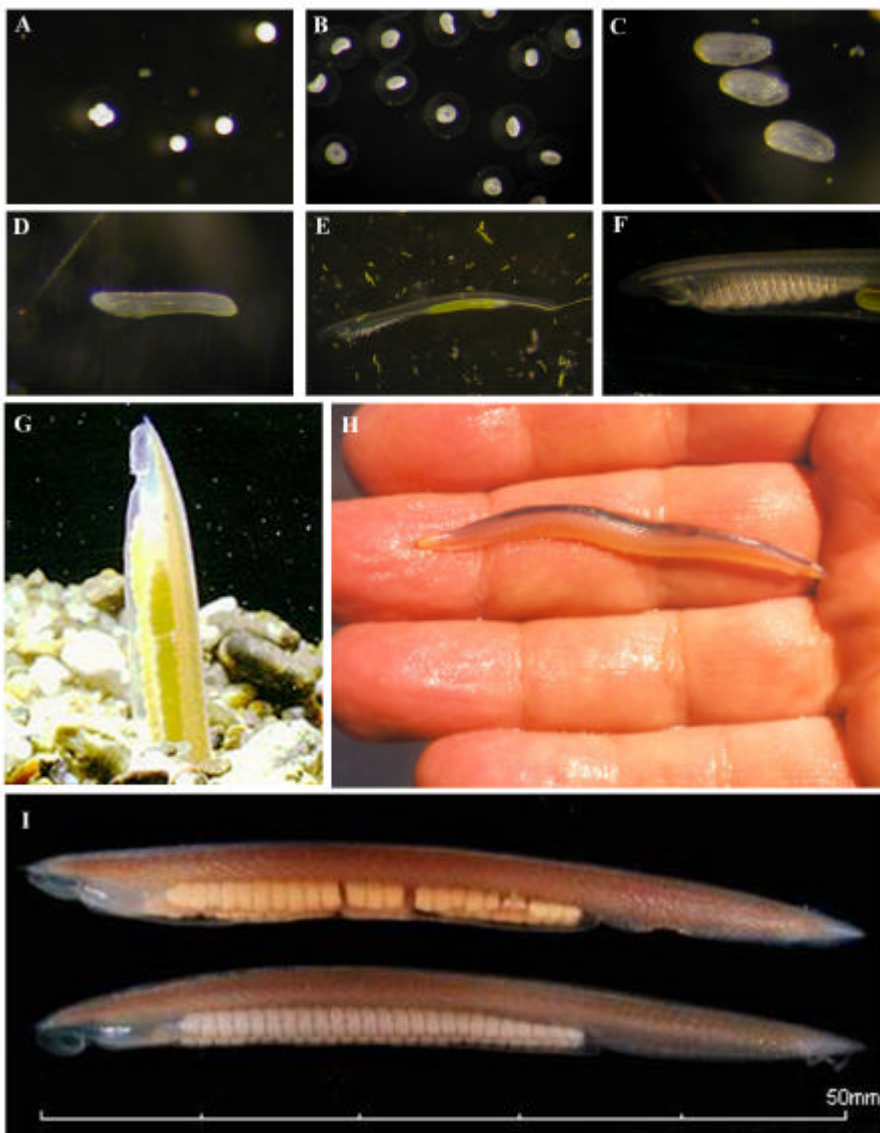


Figura 8. Fotografies de diferents aspectes de la biologia de l'amfiox. Desenvolupament embrionari: Estadis de 1, 2 i 4 cèl·lules (A), estadi de gàstrula(B), estadi de nèurula (C), estadi larval primerenc (D), estadi en fase de metamorfosi (E), part anterior d'un individu finalitzant la metamorfosi (F), individu adult en el seu hàbitat natural (G). Fotografia d'un adult en una mà (H). Amfiox femella (superior) i mascle (inferior) en estadi reproductiu amb les gònades madures (I).

4.2 Subfilum Urocordats

Són els cordats invertebrats més abundants, amb aproximadament unes 2000 espècies marines distribuïdes per tot el món. També són coneguts com a tunicats, degut a que estan protegits per una túnica, estructura rica en cel·lulosa que els protegeix (fig. 9).

Estan dividits en tres classes: Ascidiacis, Larvacis i Thaliacis, i poden ser solitaris, colonials i compostes. Les formes adultes dels tunicats, especialment els ascidis, viuen sèssils i són cordats altament especialitzats. Solen tenir una forma cilíndrica amb un sífó inhalant, corresponent a l'extrem anterior del cos, i un sífó exhalant a la part dorsal (fig. 10). L'aigua entra pel primer sífó, passa a una faringe ciliada, i a través de les fenadures branquials surt a l'atri per ser expulsada finalment pel sífó exhalant. L'aliment és retingut per una xarxa mucosa segregada per l'endostil i reconduït formant un cordó (làmina dorsal) cap a l'estómac on serà digerit. Les restes d'aliments no digerits són expulsats a l'exterior per l'anús, situat al costat del sífó exhalant. Cal destacar la gran similitud que existeix entre el sistema d'alimentació dels tunicats i dels cefalocordats: sífó inhalant/boca, faringe ciliada, fenadures branquials, atri, sífó exhalant/atripor, endostil, xarxa mucosa i solc epifaringi/làmina dorsal.

El sistema nerviós es redueix a un gangli nerviós i un plexe localitzat a la regió dorsal de la faringe. Els tunicats normalment són hermafrodites amb reproducció creuada. De les quatre



Figura 9. Un ascidi de l'espècie *Ciona intestinalis*.

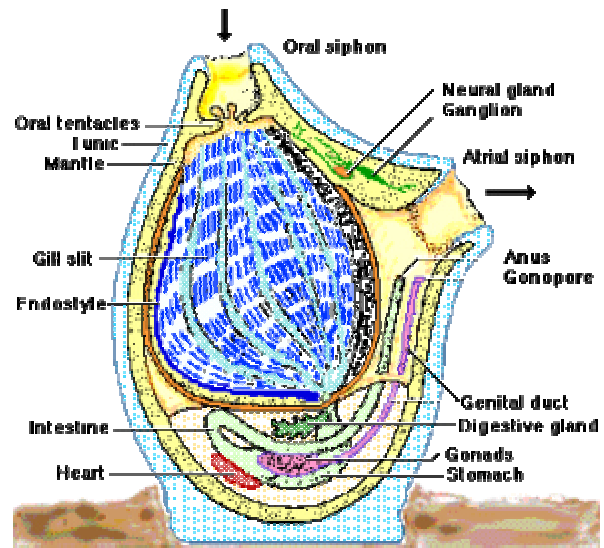


Figura 10. Representació esquemàtica d'un ascidi

característiques distintives dels cordats, els ascidis adults tan sols presenten les fenadures branquials, ja que les altres degeneren per un dràstic procés metamòrfic. La forma larval (fig.11), també coneguda com a cap-gros, presenta un disseny corporal cordat prototípic amb un tub neural buit dorsal, un notocordi axial flanquejat per múscles i una banda endodèrmica ventral, una cua propulsora postanal i fenadures branquials en una gran faringe amb endostil (revisat a Di Gregorio i Levine, 1998, Wada i Satoh, 1999 i Corbo *et al.*, 2001). El desenvolupament embrionari presenta un patró invariable i determinat amb simetria bilateral. La gastrulació i la neurulació és totalment comparable a la dels vertebrats. A l'estadi d'esbós de cua (*tailbud*) l'embrió està format per un cap i una cua. L'estadi de capgrós està format per tan sols 2600 cèl·lules, el llinatge de les quals és conegut (Satoh 1994). En aquest estadi, malgrat tenir un sistema endodèrmic ventral al llarg de tota la cua i una cavitat gàstrica en el cap, mai arriba a assolir un estat funcional fins després de la metamorfosi.



Figura 11. Larva d'ascidi mostrant el disseny corporal típic d'un cordat.

4.3 Models animals invertebrats per inferir les funcions dels gens

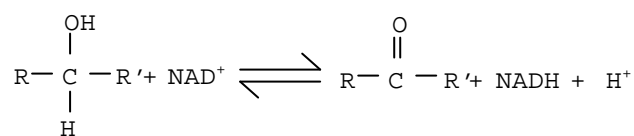
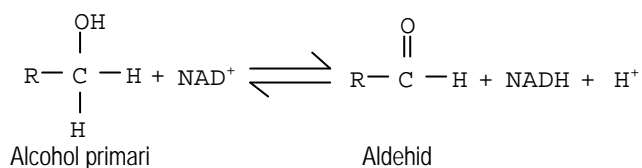
Caenorhabditis elegans i *Drosophila melanogaster* són els dos models animals invertebrats més estudiats. Els seus genomes, petits i compactes, han estat seqüenciats. Nombroses tècniques per l'estudi funcional dels gens han estat posades a punt en aquests dos organismes, dels quals es disposa de grans col·leccions de mutants. Es podria afirmar rotundament que la visió actual de la biologia no podria haver-se assolit sense els treballs que s'han dut a terme en *Drosophila*. El grau de conservació gènica al llarg de l'evolució és prou gran per a que fins i tot gens humans es puguin bescanviar i complementar amb els de *Drosophila* i *C. elegans*. Malgrat tot, aquest dos organismes protostomats tenen una morfologia i un disseny corporal molt diferent als mamífers. Pertanyen al grup dels protostomats i es caracteritzen per un desenvolupament en espiral i determinat, amb un celoma d'origen esquizocèlic, i on el blastopor donarà lloc a la boca. D'altra banda, els cordats, entre els quals està l'home, juntament amb els chaetognats, els hemicordats i els equinoderms formen l'altre gran grup d'animals bilaterals: els deuteròstoms. Aquests es caracteritzen per tenir un desenvolupament radial i indeterminat, amb un celoma

enterocèlic, i on l'anús es forma a partir del blastopor. D'aquest gran grup, el ratolí i el peix zebra són els dos organismes models més emprats. El seu patró corporal és molt més comparable a l'humà, i per tant és molt més fàcil extrapolar els coneixements adquirits amb el seu estudi. Els inconvenients que presenten són el seu elevat grau de complexitat, referida al seu genoma, el nombre de gens, la redundància funcional i l'elevat nombre de tipus cel·lulars diferents. Els seus projectes de seqüenciació genòmica ja s'han iniciat, i en pocs anys disposarem de molta més informació.

Els amfioxos i els ascidis es presenten com les alternatives de models animals que poden complementar aquest escenari. La seva estratègica posició filogenètica, molt més a prop dels vertebrats que *Drosophila* i *C. elegans*, i el fet que varen divergir abans de les duplicacions genòmiques i per tant tenen un genoma amb menys redundàncies funcionals, forma un conjunt de característiques idònies per ocupar aquesta condició prevalent. No obstant això, cal recordar que els amfioxos i ascidis no són els ancestres dels vertebrats, sinó que han evolucionat independentment durant més de 500 MA, i només una anàlisi sintètica dels tres grups (urocordats, cefalocordats i vertebrats) ens pot permetre inferir com era l'estat ancestral, i com ha progressat l'evolució fins a generar les formes actuals.

5. La família gènica de les alcohol deshidrogenases

Les alcohols deshidrogenases (ADHs) formen un gran grup d'enzims capaços de catalitzar l'oxidació d'alcohols (fig. 12) i es troben àmpliament distribuït des en tots els tipus d'éssers vius.



Alcohol secundari

Cetona

Figura 12. Reaccions d'oxidoreducció catalitzades per les ADHs formen part de tres superfamílies:

- 1) La superfamília de les deshidrogenases-reductases de cadena mitjana (MDRs). Els seus membres tenen una mida aproximada d'uns 350 aminoàcids (Jörnvall *et al.*, 1981, Jörnvall *et al.*, 1993, Persson *et al.*, 1994). Comprèn més de 537 enzims, entre els quals destaca la família de les

alcohol deshidrogenases (ADHs), i moltes altres deshidrogenases que actuen sobre substrats tan diversos com el sorbitol, la glucosa o el xilitol (fig. 13A).

- 2) La superfamília de les deshidrogenases-reductases de cadena curta (SDRs). Inclou més de 1056 enzims amb una mida del monòmer d'uns 250 aa, (Jörnvall *et al.*, 1981, Jörnvall *et al.*, 1995). Formen part d'aquesta superfamília l'alcohol deshidrogenasa de *Drosophila* i enzims com les retinols deshidrogenases microsomals o les 17 β - i 11 β -hidroxiesteroide deshidrogenases (fig. 13B).
- 3) La superfamília de les alcohol deshidrogenases-reductases de cadena llarga. Aquest grup és molt més heterogeni i inclou algunes ADHs dependents de ferro i NAD(P)⁺, amb una estructura que presenta alguna semblança amb les MDRs (Scopes, 1983, Williamson & Paquin, 1987, Jörnvall *et al.*, 1999).

En aquesta tesi ens hem centrat en l'estudi de la família gènica de les ADH-MDR. El primer membre caracteritzat va ésser l'alcohol deshidrogenasa de fetge de cavall (Bonnichsen & Wassén, 1948), la qual ha estat estudiada exhaustivament i ha servit de model pels membres descrits posteriorment. Els gens d'aquesta família codifiquen per enzims que metabolitzen una gran varietat de substrats, incloent-hi alcohols aromàtics, primaris i secundaris, retinoides, aldehids, esteroides, i productes de la peroxidació

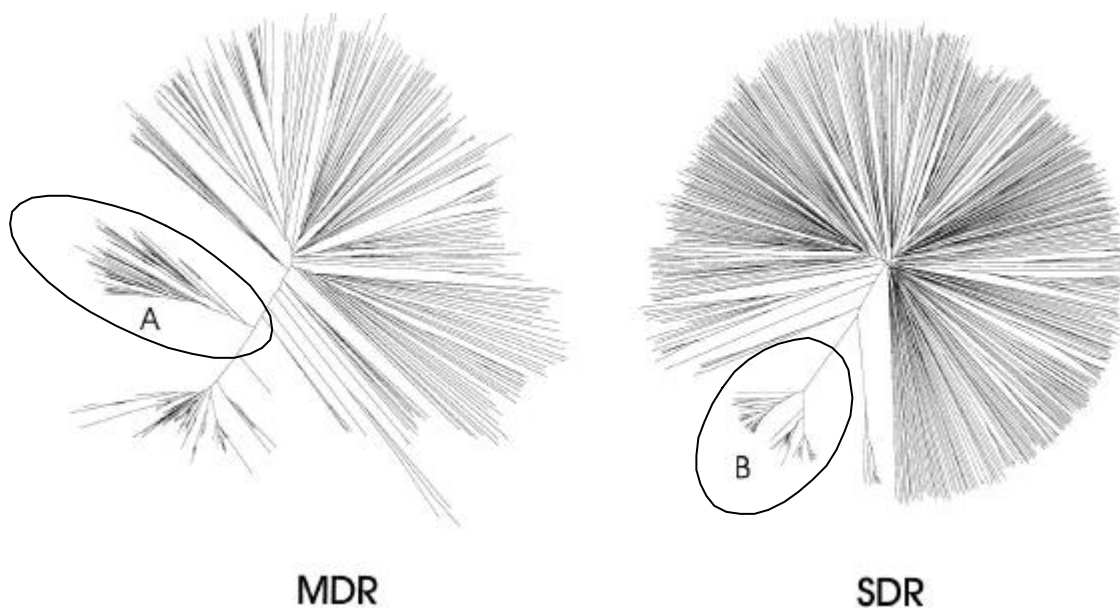


Figura 13. Relacions evolutives de les superfamílies MDR i SDR on apareixen les famílies gèniques de les Adhs-MDR del regne animal (A) i Adhs-SDR de *Drosophila* (B). (Jörnvall *et al.*, 1999)

lipídica. En vertebrats s'han descrit set classes diferents d'ADHs codificades per gens no ortòlegs, definits en base a l'homologia de les seves seqüències, les activitats catalítiques i els patrons d'expressió. De totes les classes, l'ADH de classe 3 (ADH3) és l'única que s'ha trobat en organismes no vertebrats, àmpliament distribuïda en procarïotes i també en tots els eucariotes estudiats (fongs, plantes i animals). La proposta recent de nomenclatura per aquesta família (Duester *et al.*, 1999) és la que emprarem al llarg d'aquest treball (taula 1).

5.1 L'ADH classe 3

L'ADH3 és de fet l'enzim que havia estat prèviament descrit com a formaldehid deshidrogenasa depenent de glutatió, FDH, (EC 1.2.1.1) (Koivusalo *et al.*, 1989) a partir de fetge de vedella (Strittmatter & Ball, 1955) i en l'home (Uotila & Koivusalo, 1974) i posteriorment com a ADH3 en l'home (Parés & Vallee, 1981). Aquest enzim s'ha de considerar com una alcohol deshidrogenasa ja que el seu substrat no és el formaldehid (FA) sinó el S-hidroximetilglutatió, un aducte que es forma

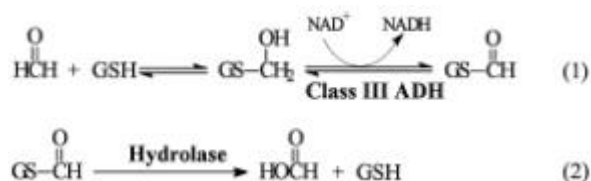


Figura 14. Reacció d'oxidoreducció del formaldehid a S-formilglutatió (1) i posterior hidròlisi a àcid fòrmic (2) (Yang *et al.*, 1997).

espontàniament entre el glutatió i el FA. Aquest és oxidat de forma reversible a S-formilglutatió amb la concomitant reducció d'una molècula de NAD⁺. El S-formilglutatió és posteriorment hidrolitzat generant glutatió i àcid fòrmic (fig. 14), el qual és finalment, o bé oxidat a CO₂ o bé incorporat a alguna via metabòlica com a font de carboni.

Es creu que l'ADH3 és un component important del metabolisme cel·lular per a l'eliminació del FA tant d'origen ambiental (el FA es fa servir en la producció de resines, plàstics, inhibidors de corrosió, preservador de cosmètics, i es genera a partir de substàncies com el metanol i diclorometà, o en combustions incompletes, cigarrets, automòbils i cuines (Dietrich *et al.*, 1996, Casanova *et al.*, 1997), com endogen, generat a partir de serina

Organism	Class I	Class II	Class III	Class IV	Class V	Class VI	Class VII
Human	ADH1A ADH1B ADH1C	ADH2	ADH3	ADH4	ADH5		
Mouse	ADH1	ADH2	ADH3	ADH4			
Rat	ADH1	ADH2	ADH3	ADH4		ADH6	
Deer mouse	ADH1					ADH6	
Horse	ADH1E ADH1S		ADH3				
Baboon	ADH1B ADH1C						
Monkey	ADH1A						
Rabbit	ADH1	ADH2A ADH2B	ADH3				
Gopher	ADH1						
Chicken	ADH1		ADH3				ADH7
Quail	ADH1						
Ostrich	ADH1	ADH2					
Kiwi	ADH1						
Alligator	ADH1						
Cobra	ADH1						
Lizard	ADH1A ADH1B		ADH3				
Frog	ADH1			ADH4			
Codfish	ADH1		ADH3H ADH3L				
Shark			ADH3A ADH3B				
Hagfish			ADH3				

Taula 1. Nomenclatura proposada per la família gènica de les *Adhs* als vertebrats (Duester *et al.*, 1999)

i glicina (Uotila & Koivusalo, 1989). El FA és un potent agent irritant que pot causar llagrimització, rinitis, faringitis i dermatitis de contacte degut a la seva reactivitat (Sanghani *et al.*, 2000). D'altra banda, s'ha descrit que el FA es pot unir a àcids nucleics, formant complexes DNA-Proteïna-FA i aductes RNA-FA (Casanova *et al.*, 1997) i d'aquesta manera produir efectes mutagènics i carcinogènics (Sanghani *et al.*, 2000).

L'ADH3, a l'igual que les altres ADHs del grup de les MDRs, actua com un homodímer de dues subunitats d'uns 375 residus aminoacídics i una massa molecular d'aproximadament uns 40 kDa. Cada subunitat conté 2 àtoms de zinc i presenta un domini Rossman on s'uneix el cofactor (fig. 15).

A diferència d'altres classes d'ADHs, l'ADH3 presenta més activitat per alcohols de cadena llarga, com ara octanol i ω -hidroxiaïcids grassos, que per alcohols de cadena curta com l'etanol, amb una K_m superior a 2 M (Julià *et al.*, 1987). S'ha demostrat que l'Arg-115, típica de la classe 3, juga un paper

central en la interacció amb el substrat i en la seva activació mitjançant l'acció d'anions hidrofòbics (Engeland *et al.*, 1993, Holmquist *et al.*, 1993). Estudis de mutagènesi dirigida han demostrat com canvis entre posicions específiques de classe, poden intercanviar algunes de les propietats característiques de cada classe. Per exemple, els canvis de Asp57 (ADH3) a Leu (ADH1) i Tyr93 (ADH3) a Phe (ADH1) sensibilitzen l'ADH3 al 4-metilpirazol, un clàssic inhibidor de les ADHs de classe 1 (Estonius *et al.*, 1994, Hedberg *et al.*, 1998).

L'ADH3 és majoritàriament un enzim citosòlic, però dades immunocitoquímiques revelen la seva presència en el nucli de cèl·lules glials i d'hepatòcits de rata, on estaria associada a la cromatina condensada i regions intercromàtiques proporcionant protecció al DNA (Iborra *et al.*, 1992). D'altra banda a l'ADH3 se li comença a atribuir una tasca citoprotectora addicional contra processos neurodegeneratius (Mori *et al.*, 2000), ja que és l'única ADH present de forma majoritària en el cervell

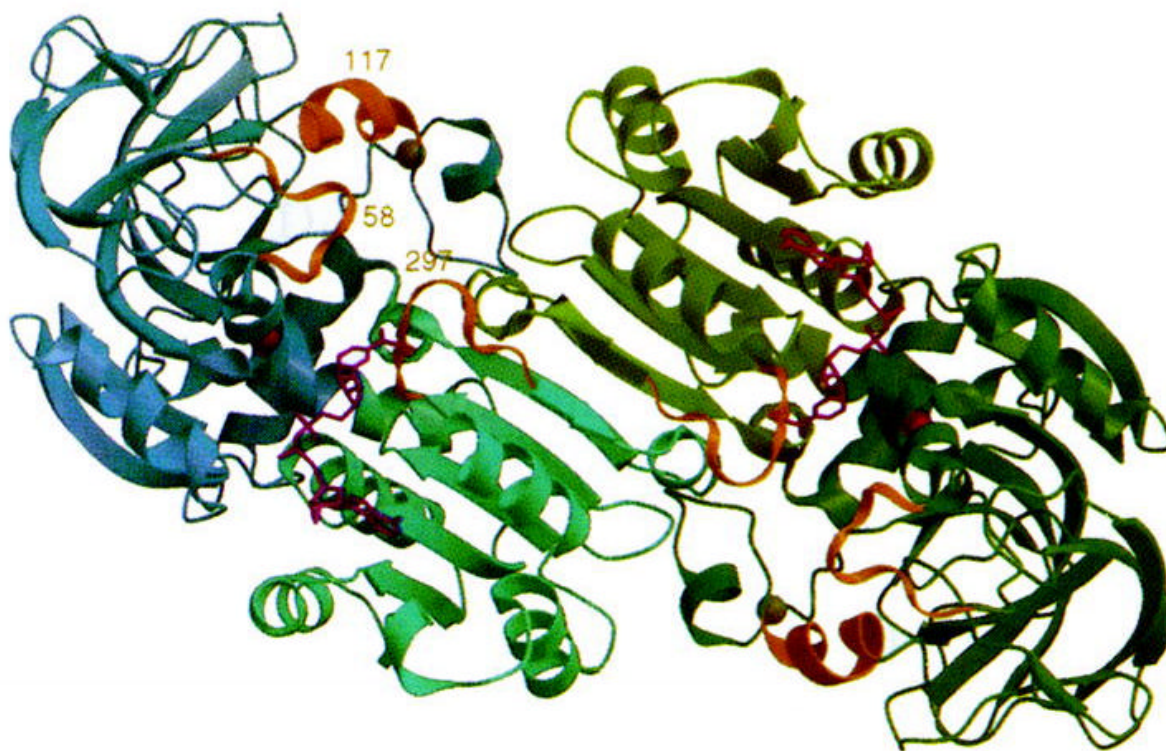


Figura 15. Representació gràfica de l'estructura tridimensional del dímer de l'ADH3 humana (blau i verd). En la part interna de cada subunitat s'observa en un color més fluïd el lloc d'unió al coenzim, i en color més fort els dominis catalítics. Els àtoms de zinc catalítics estan en vermell i els estructurals en gris. El NAD està representat en color porpra i les estructures secundàries clau al voltant dels llocs d'unió al substrat estan en taronja (Yang *et al.*, 1997).

(Beisswenger *et al.*, 1985), i hi ha evidències que demostren que reconeix el 20-OH-leucotriè B₄, un intermediari dels processos inflamatoris (Gotoh *et al.*, 1990).

L'ADH3 és un enzim present en tota l'escala filogenètica, des de bacteris fins a l'home. Malgrat la gran distància evolutiva entre l'ADH3 de procariotes i l'ADH3 de mamífers, les constants cinètiques, K_m i K_{cat} , són molt semblants. Això es deu a l'elevat grau de conservació estructural, on els dominis funcionals s'han mantingut quasi invariables. Tant és així, que monòmers d'ADH3 d'organismes tan diferents com les mixines, l'home o la *Drosophila* poden formar heterodímers (Danielsson *et al.*, 1994b), o que l'anticòs contra l'ADH3 de mamífers reconegui l'ADH3 d'*E. coli* i d'*Arabidopsis* (Gutheil *et al.*, 1992, Martínez *et al.*, 1996). La presència de l'ADH3 en tots els regnes, el seu elevat grau de conservació i la seva expressió ubíqua en tots els teixits analitzats indicarien una funció cel·lular basal (*housekeeping*) per aquest enzim i suggeririen un origen molt antic i un paper crucial pels éssers vius.

A continuació s'enumeren algunes de les característiques de l'ADH3 en alguns dels grans grups d'organismes:

- **Procariotes:** L'ADH3 ha estat caracteritzada en bacteris gram negatiu, mentre sembla absent en gram positiu i arqueobacteris. Soques bacterianes portadores de plasmidis amb còpies extres del gen *Adh3* presenten una elevada resistència al FA (Kuemmerle *et al.*, 1996), mentre que microorganismes metanotròfics i metilotròfics utilitzen l'ADH3 en les vies metabòliques que aprofiten substrats d'un carboni i com una font generadora addicional de poder reductor (Brock *et al.*, 1994, Ras *et al.*, 1995).
- **Llevat:** la sobreexpressió del gen de l'*Adh3* li confereix elevada tolerància al FA, però les soques *Adh3* no perden totalment la capacitat de metabolitzar-lo, el que suggereix l'existència d'altres enzims complementaris (Grey *et al.*, 1996). L'expressió del gen és induïble amb FA i de forma molt més feble amb etanol. Estudis de la seva regió promotora han revelat la presència d'una caixa TATA essencial per l'expressió del gen, a més de diferents elements reguladors negatius que participen en el control de la inducció pels diferents agents químics.
- **Plantes** (*Arabidopsis* i el pèsol): les característiques bioquímiques són molt semblants a les presentades per les

ADH3 d'animals, i l'anàlisi Northern ha mostrat la seva presència en tots els teixits (Martínez *et al.*, 1996, Shafqat *et al.*, 1996).

- **Invertebrats:** En artròpodes (*Drosophila melanogaster*), el patró de distribució tissular no ha estat descrit, però s'ha mostrat que l'*Adh3* s'expressa al llarg de tots els estadis del desenvolupament (Luque *et al.*, 1994). El seu promotor no presenta caixa TATA i existeixen tres inicis de transcripció diferents. En cefalòpodes (pop, sèpia i calamar), s'han trobat nivells variables d'activitat enzimàtica ADH3 en la majoria de teixits analitzats (hepatopàncreas, estómac, brànquies, glàndules salivals, gònades, ull, cor branquial i musculatura) (Fernández *et al.*, 1993).
- **Vertebrats:** En àgnats (*Mixina glutinosa*, *hagfish*) i en peixos (bacallà), s'han detectat varies isoformes d'ADH3, tres i dues respectivament, a partir de fetge, intestí i d'altres teixits com cervell, cor, estómac, cec pilòric i melsa (Danielsson *et al.*, 1994b, Danielsson *et al.*, 1996). L'ADH3 de rèptils (*Uromastix hardwickii*) presenta valors molt similars per les constants cinètiques a les descrites en d'altres vertebrats, però no s'han fet estudis de la seva distribució corporal (Hjelmqvist *et al.*, 1995). En mamífers adults, assajos d'activitat enzimàtica i Northern han revelat que el fetge és l'òrgan on majoritàriament s'expressa, però també és detectable significativament en d'altres parts del

Teixit	Activitat formaldehid deshidrogenasa nmol/min/mg prot
Fetge	73.5
Cólon	28.7
Ronyó	26.6
Estómac	21.6
Intestí prim	18.6
Cervell	17.9
Melsa	13.9
Cor	16.0
Pulmó	10.4
Múscul	7.9
Testicle	6.7
Teixit adipós bru	8.5
Intestí gruixut	7.6
Pell	5.6
Esòfag	5.4
Teixit adipós blanc	2.6

Taula 2. Activitat específica ADH3 en teixits de rata (Adaptat de Uotila i Koivusalo, 97).

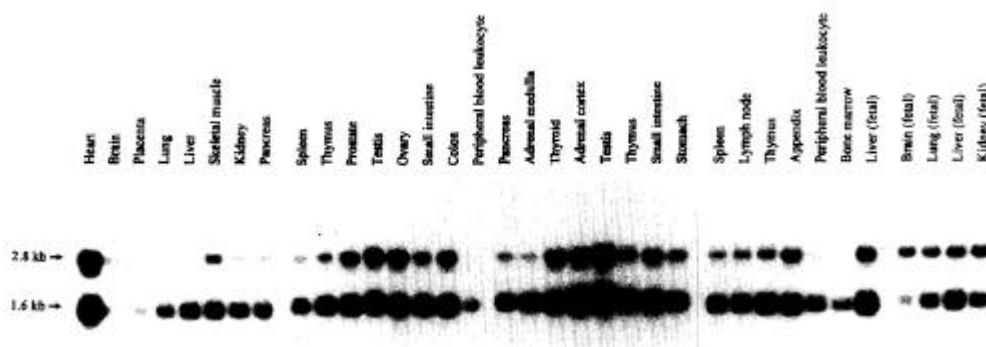


Figura 16. Anàlisi Northern de l'expressió de l'*Adh3* sobre diferents teixits humans (Estonius *et al.*, 1996)

sistema digestiu, muscular i respiratori (taula 2) (Uotila & Koivusalo, 1997, Boleda *et al.*, 1989, Estonius *et al.*, 1993, Höög *et al.*, 1994, Hur *et al.*, 1992, Edenberg, 2000). Les diferències d'activitat poden ser de fins quasi 30 vegades entre diferents teixits (taula 2), malgrat que els estudis de detecció d'mRNA no revelen aquestes grans diferències (fig. 16). En l'home, amb excepció de la placenta i el cervell, on els nivells d'expressió de l'ADH3 són significativament menors, la resta de teixits presenten poques diferències (Estonius *et al.*, 1996). En rosegadors només es detecta un transcrit de 1.6 kb (Hur *et al.*, 1992), a diferència de l'home on hi ha un transcrit addicional de 2.8 kb (fig 16). El significat funcional dels diferents transcrits es desconeix, però es creu que són deguts a diferents llocs de poliadenilació (Sharma *et al.*, 1989). La regió 5' de l'*Adh3* a mamífers mostra característiques típiques dels gens expressats constitutivament, amb presència d'illes CpG i manca de caixes TATA i CAAT (Hur & Edenberg, 1992, Foglio & Duester, 1996).

5.2 Altres classes d'ADH

En vertebrats s'han definit sis classes més d'ADHs, que comparteixen al voltant d'un 60-70% d'identitat d'aminoàcids (Taula 3). A més, cada classe pot presentar diferents isoenzims. No s'ha trobat cap organisme que contingui a la vegada totes les classes d'ADH i per exemple a ratolí s'han descrit les classes 1, 2, 3 i 4 mentre que l'home presenta també la classe 5. La classe 6 ha estat descrita únicament en alguns rosegadors, i la classe 7 només en pollastre (Duester *et al.*, 1999). Recentment s'ha identificat una ADH a amfibis amb un 60% d'identitat a la classe 1, amb la particularitat de preferir el NADP(H) com a cofactor. Aquest nou enzim podria suposar l'aparició d'una nova classe en els vertebrats (ADH8) (Duester, 2000, Peralba *et al.*,

1999a). Finalment, en les plantes també hi ha un grup d'ADHs que presenten activitat enfront l'etanol, la classe P. Són exclusives dels vegetals i també s'haurien originat per duplicació a partir de la classe 3, però de forma totalment independent a les classes d'ADHs d'animals.

- L'ADH de **classe 1** (ADH1), és la "clàssica" alcohol deshidrogenasa descrita a fetge, arribant a assolir l'1% del total de proteïna citosòlica. És activa enfront l'etanol, presentant una K_m baixa, i es creu que juga un paper central en el seu metabolisme. *In vitro* s'ha demostrat que pot oxidar el retinol a retinal (Boleda *et al.*, 1993; Han *et al.*, 1998; Yang *et al.*, 1994), i també s'ha suggerit que podria estar involucrada en el metabolisme de les hormones esteroides (McEvily *et al.*, 1988). La classe I de peixos presenta una estructura proteica més semblant a la classe III, malgrat funcionalment presenta constants cinètiques típiques de la classe I. D'aquesta manera s'ha proposat que podria representar un pas intermig en l'evolució de la família de les ADHs des de la classe 3 cap a les altres classes.
- L'ADH de **classe 2** (ADH2) presenta activitat front l'etanol malgrat té una K_m substancialment més elevada. Es creu que contribueix a l'eliminació de l'etanol quan aquest es troba a concentracions elevades. Quan es comparen les ADH2 d'espècies diferents s'observa que presenten especificitats variables i una baixa conservació a nivell de seqüència. En humans presenta activitat contra el all-trans-retinol, mentre que a rosegadors sembla que tingui més afinitat pel 9-cis-retinol (Svensson *et al.*, 1998; Svensson *et al.*, 1999). Aquests resultats suggereixen que l'ADH2 no ha jugat un paper

evolutivament conservat com a retinol deshidrogenasa, i deu haver adquirit funcions específiques a cada llinatge.

- L'ADH de **classe 4** (ADH4) presenta activitat amb l'etanol amb una K_m intermitja. D'entre les ADHs del grup de les MDRs és la més eficient oxidant els dos tipus de retinol (9-cis- i all-trans-retinol) als seus respectius retinals. L'oxidació del retinol és el pas limitant en la síntesi d'àcid retinoic, i d'aquí la gran importància de l'estudi de l'ADH4 (per una revisió: Duester, 2000).
- L'ADH de **classe 5** (ADH5) només ha estat detectada en rosegadors i primats i pot metabolitzar l'etanol a concentracions moderades (Chen & Yoshida, 1991, Yasunami *et al.*, 1991).
- L'ADH de **classe 6** (ADH6) s'ha trobat únicament en rosegadors i malgrat ser semblant a l'ADH humana classe 5, el fet que tan sols comparteix un 67% d'identitat amb aquesta, ha determinat que es consideri una nova classe. (Zheng *et al.*, 1993, Hoog & Brandt, 1995).

- L'ADH de **classe 7** (ADH7) s'ha identificat exclusivament en embrions de pollastre i presenta preferentment activitat amb l'all-trans-retinol i 3β -5 α -hidroxiesteroïdes, malgrat és inactiva davant dels 3β -5 β -hidroxiesteroïdes (Kedishvili *et al.*, 1997).

Com hem vist al llarg de la descripció de les diferents classes d'ADHs, tres són els principals àmbits funcionals: 1) l'eliminació d'alcohols provinents de la dieta i la seva influència en el risc d'alcoholisme i en els càncers associats a aquest hàbit, 2) metabolisme d'hormones esteroïdes i 3) regulació de la formació d'àcid retinoic com a morfogen fundamental en el desenvolupament animal. Tanmateix, el gran ventall de substrats que pot metabolitzar cadascuna de les classes d'ADHs, juntament amb l'elevada redundància catalítica, les diferents afinitats pels substrats i les diferències quantitatives que presenten els seus patrons d'expressió altament solapants, fa que sigui molt difícil precisar els seus papers fisiològics.

Taula 3. Característiques de les diferents classes d'alcohol deshidrogenases de vertebrats.

Classe	% similitud a l'ADH1A	K _m etanol (mM)	Organismes	Patró d'expressió	Referències
ADH1	100-93	0.05-4	Peixos, amfibis, rèptils, aus, rosegadors i primats	Principalment al fetge i teixits sensibles als efectes dels retinoides (intestí, ronyó, glàndula adrenal, testicles, epidídim, úter i ovaris, certes parts del cervell, mesonefros i sac aeri).	(Danielsson et al., 1992, Bianchone et al., 1986, Ang et al., 1996a; Ang et al., 1996b, Cederlund et al., 1991, Peralba et al., 1999b, Hoffmann et al., 1998, Vonesch et al., 1994, Haselbeck & Duester, 1997, Estonius et al., 1996, Deltour et al., 1997, Haselbeck et al., 1997, Smith et al., 1971, Martínez et al., 2001)
ADH2	60	34	Rosegadors i primats	Patró d'expressió molt limitat al fetge, intestí prim i possiblement també a la pell.	(Estonius et al., 1996, Svensson et al., 1999, Cheung et al., 1999)
ADH3	60	>2000	Àgams, peixos, amfibis, rèptils i mamífers	Patró ubic, especialment en fetge. És l'única classe d'ADH que s'expressa en el cervell.	(Edenberg, 2000, Edenberg et al., 1995, Giri et al., 1989, Hur & Edenberg, 1992, Danielsson et al., 1994b, Cederlund et al., 1991, Estonius et al., 1996, Foglio & Duester, 1996, Danielsson et al., 1996, Ang et al., 1996a)
ADH4	69	28	Amfibis, rosegadors i primats	És detectat a l'epiteli estomacal, pell, ull, glàndula adrenal, teixits reproductius i certes parts del cervell. No s'expressa al fetge.	(Hoffmann et al., 1998, Zgombic-Knight et al., 1995, Moreno & Parés, 1991, Deltour et al., 1997, Haselbeck & Duester, 1997, Zgombic-Knight et al., 1995, Estonius et al., 1996, Hoffmann et al., 1998, Ang et al., 1996a, Martínez et al., 2001)
ADH5	62	20	Rosegadors i primats	S'expressa a estómac i fetge, i amb menor quantitat també a intestí prim i ronyons fetals.	(Estonius et al., 1996, Stromberg & Höög, 2000, Edenberg, 2000, Yasunami et al., 1991, Chen & Yoshida, 1991)
ADH6	60	20	Rosegadors	Només s'expressa a alts nivells en fetge i en menor grau en ronyons i pulmons.	(Zheng et al., 1993, Hoog & Brandt, 1995)
ADH7	67	17	Embrions de pollastre	S'expressa en fetge però no en cor, múscul, testicles o cervell.	(Kedishvili et al., 1997)

OBJECTIUS

Objectius

L'objectiu d'aquest treball ha estat analitzar l'evolució molecular de la família gènica de les *Adhs* al llarg de la història dels cordats a través de la caracterització dels membres d'aquesta família en amfioxos i ascidis. En aquest sentit, els cefalocordats i els urocordats, els dos grups filogenèticament germans dels vertebrats que es van separar abans que es produïssin les duplicacions gèniques a gran escala dels vertebrats, són un bon model d'estudi per inferir les formes ancestrals de gens i proteïnes.

Aquesta anàlisi s'ha fet emprant dues vies d'aproximació complementàries:

1. Una via de caire evolutiu, on a partir d'estudis filogenètics s'ha deduït l'estat ancestral de la família de les *Adhs* i s'ha relacionat la seva expansió amb els esdeveniments de grans duplicacions gèniques durant l'evolució primerenca dels cordats.
2. Una via de caire funcional, amb la caracterització bioquímica i l'estudi dels patrons d'expressió per tal d'entreveure el procés seguit en l'adquisició de noves funcions.

