

UNIVERSITAT DE BARCELONA

FACULTAT DE BIOLOGIA

Departament de Genètica

Tesi Doctoral de

Carolina MINGUILLÓN GIL

amb el títol

**Evo-Devo de la complexitat gènica i dels somites a l'origen dels vertebrats: els gens
extended Hox i *hairy* de l'amfiox.**

DIRECTOR: Dr. Jordi Garcia Fernàndez

Barcelona, 5 de juliol de 2002

Índex	i-ii
Introducció	3-56
Prefaci	4
1. L'EVO-DEVO	5
2. GRANS TRANSICIONS DURANT L'EVOLUCIÓ DELS ANIMALS	6
2.1. Transició 1: L'origen de la multicel·lularitat	7
2.2. Transició 2: Evolució d'organismes bilateris a partir de diblàstics	7
2.3. Transició 3: L'origen dels vertebrats	7
3. LA DUPLICACIÓ GÈNICA	10
3.1. Com pot donar-se una duplicació?	10
3.2. Com mantenir un gen duplicat?	10
3.3. Com adquirir una nova funció (neofuncionalització)	12
3.4. La duplicació d'un genoma sencer (DGS)	12
4. L'AMFIOX	12
4.1. Obtenció d'embrions de <i>Branchiostoma floridae</i>	13
4.2. Anatomia de l'adult	13
4.3. Desenvolupament de l'amfiox	15
4.4. Posició filogenètica de l'amfiox	19
4.5. L'amfiox i les innovacions dels vertebrats	20
5. L'HOMEOBOX	21
5.1. Classificació i evolució dels gens amb homeobox	22
5.2. Els gens Hox	24
5.3. La classe Even-skipped (Evx)	34
5.4. La classe Mox	39
5.5. La classe empty spiracles (Emx)	41
6. EL DOMINI bHLH	44
6.1. Classificació de les proteïnes amb domini bHLH	45
6.2. Proteïnes amb domini bHLH del grup E	46
6.3. La família hairy	47
6.4. Funcions de la família hairy en vertebrats	49
Objectius	57-62
Objectius	59
Objectives	61
Resultats	63-140
Introducció als resultats	65
CAPÍTOL I. Caracterització de la Classe Mox de <i>B. floridae</i>	66
CAPÍTOL II. Caracterització de l'extrem 5' de <i>cluster</i> Hox de <i>B. floridae</i>	67
CAPÍTOL III. Caracterització de la Classe Even-skipped (Evx) de <i>B. floridae</i>	69
ANNEX I. Lligament dels gens Evx al <i>cluster</i> Hox de <i>B. floridae</i>	71

CAPÍTOL IV. Evolució del genoma de l'amfiox. Un nou cas de duplicació: la Classe Empty spiracles (Emx) de <i>B. floridae</i>	74
CAPÍTOL V. Caracterització de la família hairy en l'amfiox <i>B. floridae</i>	75
Apèndix d'articles publicats o en preparació	
Discussió general	141-162
Discussió general	143
General discussion	145
1. Evolution of the extended Hox group of homeobox genes	145
2. Amphioxus somitogenesis and vertebrate novelties	152
3. Cautions with the interpretation of amphioxus gene data	155
3.1. The concept of ancestral species	155
3.2. The evolution of the amphioxus genome	155
3.3. The concept of homology	157
4. Limitations in the use of amphioxus as a model system	158
4.1. Amphioxus maintenance and reproduction in the lab	159
4.2. Complementary approaches and future prospects	160
Conclusions	163-168
Conclusions	165
1. About amphioxus genes	167
2. About gene expression, homologies, vertebrate novelties and amphioxus development	167
3. About the evolution of the extended Hox group of homeobox genes	168
Bibliografia	169-182
Bibliografia	171

Prefaci

El Desenvolupament, o estudi de l'obtenció d'un organisme pluricel·lular a partir d'una única cèl·lula ou o zigot, és una disciplina molt antiga. Així doncs, no és estrany trobar tractats, llibres o articles, en els que es descriu el desenvolupament d'un determinat organisme molt acuradament, que daten de principis del segle XIX. L'estudi del desenvolupament o embriologia ha passat per diferents etapes -o enfoc d'estudi- al llarg d'aquests quasi 200 anys, en funció dels coneixements del moment i de les tècniques disponibles (von Dassow i Munro, 1999). La primera etapa és bàsicament descriptiva i comparativa. Amb la introducció del concepte de *Mecànica del Desenvolupament* per Roux al 1894, comença una anàlisi mecànica del desenvolupament a través d'experiments. Més tard apareix *l'Embriologia Química* com a estudi experimental que pretén revelar els mecanismes subjacents a varis fenòmens que tenen lloc en determinades regions de l'embrió. Aquests fenòmens han estat definits pels estudis embriològics anteriors (per exemple la cerca de substàncies responsables de les propietats de l'organitzador). *L'Embriologia Química* no troba respostes perquè busca en llocs erronis: els embriologistes d'aquella època contempen la Genètica com a irrel·levant, perquè no entenen com els mateixos gens que estan presents des de l'ou fins a l'adult, poden tenir res a veure amb l'epigènesi. Amb la demostració de la regulació de l'expressió gènica en procariotes per Jacob i Monod al 1961, apareix la *Genètica del Desenvolupament*, com a versió moderna de *l'Embriologia Química*. Aquesta pretén explicar com els gens heredats poden influir directament en la fisiologia d'una manera epigenètica (entenent epigènesi com a processos dinàmicament regulats que tenen lloc en un ésser viu).

L'Evolució és també una disciplina força antiga, que estudia el procés pel qual les formes dels éssers vius s'han desenvolupat per descendència i amb modificació a partir de formes que han existit anteriorment. Una de les primeres formalitzacions de la matèria és atribuïble a Lamarck (1809), que proposa un origen comú dels éssers vius, i seguint a Linné, estableix subdivisions d'organismes en taxons jeràrquics a diferents nivells. També cal destacar altres evolucionistes com Geoffroy St. Hilaire que, comparant diferents fílums, intenta establir un únic pla d'organització. Tres dècades més tard que St. Hilaire, Darwin (1859) extén la idea d'evolució, proposant la selecció natural com a mecanisme d'evolució (David, 2001).

A finals del Segle XIX i principis del XX, aquestes dues ciències eren inseparables. Les comparacions entre el desenvolupament embrionari de dues espècies, s'usaven com a evidència de l'evolució [l'ontogènia recapitula la filogènia (Haeckel)] i, alhora, la història evolutiva es veia com a explicació suficient per a gairebé qualsevol estructura o procés observat durant el desenvolupament. No obstant, l'evolució s'oblida del desenvolupament quan cap als anys 20 comença a basar-se en la Genètica. Així, l'estudi del desenvolupament passa a veure's com insuficientment precís o quantitatiu per a aportar res a aquesta cada vegada més rigorosa ciència. Encara que alguns científics veuen la necessitat de reconciliació entre la ciència del desenvolupament i la ciència de l'heredabilitat ja al llarg del segle XX, aquesta no es duu a terme fins fa uns quinze anys. La reconciliació entre les dues permet el naixement a la seva interfase d'una nova disciplina: l'Evolució del Desenvolupament (de l'anglès *Evolutionary Developmental Biology*), o "Evo-Devo" (Holland, 1999).

1. L'EVO-DEVO

Malgrat per molts autors aquesta matèria neix formalment al 1999 durant un congrès de la SICB (de l'anglès *Society for Integrative and Comparative Biology*) que va tenir lloc a Atlanta, Georgia (Dalton, 2000), molts altres recorden el simposi "*Evolution of Development*" del congrés de la Societat Britànica de Biologia del Desenvolupament d'Edimburg al 1994, com el primer fòrum de discussió de la nova visió de l'Evolució com a resultat dels canvis als mecanismes de desenvolupament. Els biòlegs del camp de la genètica evolutiva intenten entendre com evolucionen els organismes i com canvia la seva forma. A la base d'aquests canvis morfològics es troben els mecanismes de desenvolupament que controlen la forma final dels organismes. Els biòlegs del desenvolupament intenten entendre com alteracions en l'expressió dels gens i en la seva funció donen lloc a canvis en la forma i el patró corporals. Així, l'Evo-Devo neix fa una mica més d'una dècada quan alguns científics es basen en l'expressió de gens del desenvolupament per a intentar explicar com han evolucionat els diferents grups d'organismes (Goodman i Coughlin, 2000). Un dels temes centrals en l'Evo-Devo és entendre com l'arquitectura del desenvolupament confereix capacitat d'evolució (de l'anglès *evolvability*) (von Dassow and Munro, 1999). Rudolph Raff, primer president de la divisió d'Evo-Devo de la SICB, espera que aquesta expliqui com varia el curs del desenvolupament entre organismes, com evoluciona el mateix, i com les tries evolutives fetes durant el desenvolupament restringeixen l'evolució. Les constriccions del desenvolupament són una de les raons per les que els organismes no poden evolucionar cap a una varietat infinita de formes. Una altra tasca de l'Evo-Devo serà intentar entendre el perquè de l'existència d'aquestes restriccions, i també definir-les (Raff, 2000).

És precisament en aquesta nova disciplina -en l'Evo-Devo- en la que es fonamenta aquesta Tesi Doctoral; en intentar estudiar com han evolucionat els mateixos programes de desenvolupament, com poden modificar-se per canvis genètics i com aquestes modificacions produeixen la passada i present diversitat de morfologies i plans corporals.

1.1. La paradoxa de l'Evo-Devo

Ja al 1977, François Jacob hipotetitzava que allò que fa diferent a dos animals no és una diferència en el contingut de constituents químics (avui en dia entesos com a gens), sinó en l'organització i distribució d'aquests (avui en dia entès com a interacció o combinació de productes gènics). De manera oposada, l'àmplia majoria de científics de l'època de Jacob creuen que la diversitat de formes animals era generada per una comparable diversitat de mecanismes genètics. Molts treballs d'aquests darrers anys han demostrat que la visió correcta era la de Jacob: no només la majoria de molècules reguladores s'han mantingut remarcablement constants en els metazous, sinó també els mecanismes moleculars en els que intervenen. Així, l'"enginyer" evolutiu apareix com a molt cautelós pel que fa referència a la tria de materials bàsics per a la construcció, mentre que com a molt creatiu pel que fa referència a l'ús d'aquests productes gènics en la diversificació de les formes animals al llarg de l'evolució (Duboule i Wilkins, 1998).

És llavors quan sorgeix la paradoxa de l'Evo-Devo: com explicar l'evolució morfològica, si per a construir animals molt diferents fem servir els mateixos gens i xarxes gèniques. L'evolució apareix com a eminentment conservadora fent ús de combinacions ja provades i reprovades una

a una en altres processos de desenvolupament o en organismes ancestrals. Per a intentar resoldre aquesta paradoxa s'han formulat algunes hipòtesis com:

- que alguna d'aquestes noves funcions hagi aparegut gràcies a les vàries formes d'intercanvi d'exons (*exon shuffling*) de determinades parts de gens pre-existents (Frigerio i col, 1986).
- *syntagmata* (xarxes gèniques) usats diferencialment en l'espai i el temps (García-Bellido, 1994). Es necessiten determinats fenòmens mutacionals que permetin que gens que ja són presents a l'organisme siguin reclutats (co-optats) per a la realització de noves funcions durant el desenvolupament. De fet, el que aquests 20 darrers anys d'estudi han demostrat, almenys pel que fa referència a la funció dels gens reguladors, és que la regla general és que un únic producte gènic s'utilitzi múltiples vegades durant el desenvolupament d'un organisme. Així, pot ser que l'increment en la complexitat del desenvolupament d'un organisme vagi lligat a l'increment de funcions biològiques per gen regulador (Duboule i Wilkins, 1998).
- La duplicació gènica seguida de divergència funcional i adquisició de noves funcions també apareix com a mecanisme per a la producció de diversitat (veieu secció 3 de la Introducció).

2. GRANS TRANSICIONS DURANT L'EVOLUCIÓ DELS ANIMALS

Al llarg de l'evolució dels animals hi ha hagut algunes fases en les que s'han patit canvis radicals dels mecanismes de desenvolupament. Aquestes fases han estat anomenades grans transicions. Cal recolzar-se en una bona filogènia del Regne Animal per a la correcta identificació de quan en l'evolució aquests mecanismes són radicalment alterats.

Holland (1998), defensa l'existència de sis transicions, mentre que Lundin (1999) en defensa l'existència de tan sols tres. Pel primer, aquestes sis fases serien:

- 1) l'origen de la multicel·lularitat
- 2) l'origen de la simetria, de dues capes embrionàries i de les neurones
- 3) l'origen de la simetria bilateral, de les tres capes embrionàries, d'un cordó nerviós axial i d'un intestí al llarg de l'eix corporal
- 4) la inversió de l'eix dorso-ventral
- 5) l'origen de la cresta neural i nous tipus cel·lulars
- 6) i per últim, l'origen d'un mesoderm amb propietats migratòries, apèndix parells i mandíbules.

Les tres transicions de Lundin serien:

- 1) l'origen dels metazous (coincideix amb la transició 1 de Holland)
- 2) l'origen dels "urbilateria" (coincideix amb la transició 3 de Holland)
- 3) l'origen dels vertebrats (coincideix amb les dues darreres transicions de Holland).

Aquestes grans transicions -vistes com a successos de macro-evolució- no poden ser degudes a fenòmens de micro-evolució convencionals, sinó que calen altres mecanismes per a explicar-les, com la duplicació gènica o l'intercanvi de dominis (de l'anglès *domain shuffling*) (Ohno, 1970).

Com que l'objectiu d'aquesta Tesi no és l'estudi de totes aquestes transicions, sinó tan sols la d'una (l'origen dels vertebrats; és a dir, la transició 3 de Lundin, ó les dues darreres de Holland), s'explicaran breument les dues primeres transicions de Lundin, posant especial èmfasi a la tercera (qüestió que serà àmpliament tractada al llarg d'aquest treball).

2.1. Transició 1: L'origen de la multicel·lularitat

S'ha descrit l'existència de múltiples duplicacions gèniques anteriors a la separació del llinatge de les esponges de la branca principal dels metazous en l'arbre evolutiu. Un conjunt de funcions genètiques compartides entre esponges i altres metazous són les referides a la transducció de senyals. Així, la majoria de duplicacions en subtipus de proteïnes amb activitat tirosina-quinasa (Suga i col., 1997), de fosfodiesterases de nucleòtids cíclics (Koyanagi i col., 1998), de proteïnes amb activitat tirosina-fosfatasa (Ono i col., 1999) i, per últim, de proteïnes G (Suga i col., 1999) són anteriors a la divergència d'esponges, cnidaris i altres metazous.

Si tots aquests nous gens es van originar a partir una duplicació del genoma sencer (DGS) o per una onada més limitada de duplicacions gèniques a l'ancestre dels metazous, és una qüestió que encara roman oberta.

2.2. Transició 2: Evolució d'organismes bilateris a partir de diblàstics

Tres mecanismes podrien haver estat implicats en aquesta transició:

- (i) canvis en els sistemes de regulació gènica
- (ii) extens intercanvi de mòduls o dominis proteics
- (iii) expansió per duplicació gènica de determinades famílies de gens.

Pel que fa referència als dos primers, la visió actual seria una barreja d'ambdós, doncs la majoria de proteïnes que trobem en esponges i cnidaris ja són modulars. Per exemple: la interacció d'algunes proteïnes Hox amb altres proteïnes amb homeodomini -les del grup EXD (*Drosophila*) o els ortòlegs de vertebrats (PBX)- es duu a terme a través d'un hexapèptid conservat de les proteïnes Hox. Aquest hexapèptid està present en triblàstics, però manca en la majoria de proteïnes Hox de cnidaris (Kuhn i col., 1999).

La duplicació de determinades famílies gèniques és un fenomen que està àmpliament acceptat almenys per a tres d'elles: els col·làgens, els gens Hox, i els receptors nuclears (Escribà i col., 1997). Trobem un nombre de gens més elevat de cadascuna d'aquestes famílies gèniques en bilateris que en organismes diblàstics.

2.3. Transició 3: L'origen dels vertebrats

Aquesta transició comporta l'aparició d'un conjunt d'estructures "inventades" pels vertebrats -caràcters derivats- com les cèl·lules de la cresta neural i els seus derivats, l'elaboració d'estructures derivades dels placodes, l'elaboració d'un cervell molt complex amb segmentació dels rombòmers, l'esquelet axial i del cap, els teixits cartilaginós i ossi, i un increment en el nombre de gens al genoma (revisat en Shimeld i Holland, 2000). El que encara està en discussió és com s'ha produït la duplicació gènica en aquesta transició. Alguns autors hipotetitzen l'existència de dues rondes de duplicació del genoma sencer (DGS) (hipòtesi 2R) que provoquen l'evolució dels vertebrats mandibulats (Holland i col., 1994; Sidow, 1996; Amores i col., 1998), mentre que d'altres la critiquen (Skrabaneck i Wolfe, 1998; Hughes i col., 2001).

2.3.1. Arguments a favor i en contra de l'existència de dues rondes de DGS (hipòtesi 2R) durant l'evolució del llinatge dels vertebrats

Els arguments a favor estan basats en (i) l'estimació i comparació del nombre de gens d'una mateixa família presents en determinades espècies de vertebrats respecte d'invertebrats i, (ii) en l'existència i mapatge de gens paràlegs presents en diferents espècies de vertebrats (regions sintèniques).

Ohno va ser un dels pioners en la introducció de la duplicació gènica com a font per a la macro-evolució. Ell mateix va proposar un model en el qual es patien com a mínim dues rondes de DGS durant l'evolució dels primers cordats. Es basa en l'estudi del contingut de DNA haploïd en diferents espècies. En l'amfiox (cefalocordat), aquest era de 0,6 pg per cèl·lula, que és similar al del genoma més petit de vertebrat i tres vegades superior al del genoma d'un ascidi (urocordat) (Atkin i Ohno, 1967). Això el portà a suggerir que hi havia hagut una expansió, possiblement per poliploïdia, dins el fílum dels cordats després de la separació dels urocordats i anterior a l'origen dels cefalocordats i vertebrats (Ohno, 1970).

Altres autors defensen també la hipòtesi 2R; la diferència es troba en quina posició de l'escala evolutiva tenen lloc aquests fenòmens (Fig. 1).

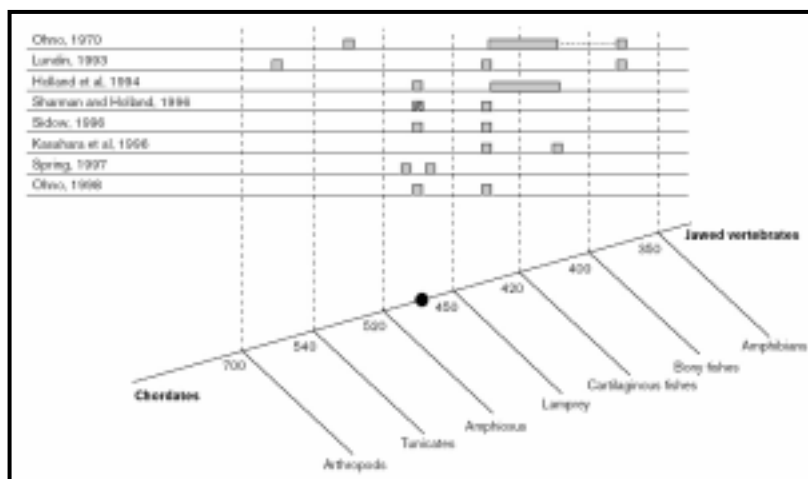


Figura 1. Propostes per al moment de duplicació durant l'evolució dels vertebrats. Els rectangles grisos indiquen les proposades DGSs, i el rectangle ratllat una proposada onada de múltiples duplicacions en tàndem. El cercle marca a l'origen dels vertebrats (extret de Skrabanek i Wolfe, 1998).

Lundin (1993) suggereix l'existència de quatre grups de regions cromosòmiques paràlogues (que contenen gens lligats o sintènics) al genoma humà i al de ratolí. Per a ell, aquestes regions representen restes conservades de dues i fins i tot de tres rondes de DGS durant l'evolució dels vertebrats (Fig. 1). Treballs més actuals similars al de Lundin en els que es troben regions sintèniques repartides en quatre cromosomes de diferents espècies de vertebrats, també han recolzat la hipòtesi 2R (Pébusque i col., 1998).

Holland i col. (1994) proposen l'existència de dues rondes de DGS durant l'evolució del llinatge dels vertebrats; la primera després de la separació dels cefalocordats del llinatge de vertebrats i la segona propera a l'origen dels vertebrats mandibulats (Fig. 1). Es basen en diferències en el nombre de gens de determinades famílies gèniques en l'amfiox respecte els vertebrats. Més tard, l'estudi de la complexitat de determinades famílies gèniques (de nou per comparació entre l'amfiox i els vertebrats), la construcció d'arbres filogenètics amb aquests gens, el mapatge cromosòmic d'algun d'aquests gens i la comparació de la mida del genoma de l'amfiox i de determinats vertebrats els permeten redefinir aquests dos fenòmens de duplicació gènica. La primera fase comportaria una duplicació extesa (però no per DGS), mentre que la segona sí que comportaria una DGS (Sharman i Holland, 1996).

La majoria dels autors contraris a la hipòtesi 2R es basen en les relacions filogenètiques dels gens duplicats en vertebrats. Si dues rondes de DGS van ocórrer durant l'evolució dels

vertebrats, al constringer un arbre filogenètic amb les 4 seqüències de vertebrat més la d'un grup extern, aquest ha de tenir una topologia [grup extern [(1,2)(3,4)] és a dir, una topologia 2+2. Bailey i col. (1997) analitzen les relacions filogenètiques dels gens dels col·làgens que es troben lligats als *clusters* Hox, així com les dels mateixos complexos Hox. Els arbres resultants no mostren una topologia 2+2 i per tant rebutgen la hipòtesi 2R. Hughes i col. (2001) analitzen les seqüències de 42 famílies gèniques que tenen membres en dos o més dels cromosomes humans on mapen els *clusters* Hox (HSA 2, 7, 12 i 17). Per a la majoria d'aquestes famílies (32), troben evidències contra la hipòtesi de que la seva duplicació es donés alhora que la dels *clusters* Hox (que seria el que esperaríem si tot el genoma s'hagués duplicat alhora). Skrabanek i Wolfe (1998) critiquen el fet de que a vegades s'hagin definit regions sintèniques que no contenen més de 2 o 3 gens similars.

2.3.2. La seqüència del genoma humà

La seqüenciació i posterior anàlisi del 94% del genoma humà i alliberament de les dades a principis de l'any passat (International Human Genome Sequencing Consortium, 2001) han permès corroborar l'existència de duplicacions gèniques al llarg de l'evolució dels vertebrats. La comparació d'aquest amb genomes d'invertebrats totalment seqüenciats com el de *Drosophila melanogaster* i *Caenorhabditis elegans*, revela que aproximadament el 60% de famílies gèniques són més complexes (pel que fa referència al nombre de gens) en humans que en qualsevol d'aquests dos invertebrats. Malgrat això, diferents aproximacions els porten a concloure que no hi ha evidències convincents per a l'acceptació de la hipòtesi 2R:

- Troben molts parells de segments que contenen uns quants gens paràlegs, però aquests estan en una menor densitat que les observades en llevat o *Arabidopsis thaliana*, organismes en els que l'ocurrència d'un fenomen de poliploidització durant la seva evolució està àmpliament acceptada.
- En comparar el genoma humà amb el de *D. melanogaster* i *C. elegans* no troben una ocurrència més elevada de relacions 4:1:1 que de relacions 3:1:1, 2:1:1, o 1:1:1 entre gens ortòlegs.
- Només pel 24% de 57 famílies observen la topologia 2+2 esperada en arbres filogenètics.
- En buscar grups de 4 cromosomes en els que es trobin gens paràlegs amb representació a tots ells, troben que l'exemple més evident és el de les regions que contenen els 4 *clusters* Hox. Malgrat troben altres grups de 4 cromosomes, aquests estan connectats però, per un baix nombre de gens paràlegs.

Podria ser però, que la interpretació d'aquests resultats no fos tan senzilla. Per exemple, fenòmens de pèrdua massiva de gens després de les DGS seguits de fenòmens de duplicació no relacionada d'alguns gens extra, dificultaria la detecció de DGS. Les úniques famílies de gens de vertebrats que tindrien una relació 4:1 amb el seu ortòleg d'invertebrats serien aquelles en que cap dels gens duplicats s'ha perdut posteriorment o no han patit una nova duplicació (no poliploidia) (Wolfe, 2001). L'elevada ocurrència d'elements mòbils al genoma humà podria ser també una altra causa per a la desorganització del genoma que s'ha duplicat i, per tant, dificultar també la detecció de cromosomes o regions paràlogues. Un altre motiu que pot complicar la detecció de les DGS podria ser degut a l'entrecruament desigual entre cromosomes homeòlegs en fases inicials posteriors a la DGS. Una millor interpretació de les dades podria sorgir de la comparació del genoma humà amb el genoma d'algun organisme molt proper als

vertebrats, però que hagués divergit dels vertebrats al moment oportú (abans de les hipotetitzades duplicacions), com l'amfiox.

3. LA DUPLICACIÓ GÈNICA

Com ja s'ha anat esmentant, els fenòmens de duplicació gènica han estat claus en alguns moments de l'evolució (revisat en Lundin, 1999). S'entén fàcilment que la duplicació d'un(s) gen(s) o d'un genoma sencer doni a l'organisme un amplíssim poder de canvi. L'existència de nou material genètic, en principi prescindible, que pot variar sense restricció, és on radica aquest poder. Ara cal preguntar-se com de fàcil és patir una duplicació i, després, com de fàcil és mantenir les còpies duplicades.

3.1. Com pot donar-se una duplicació?

Tres mecanismes principals provoquen la duplicació d'un(s) gen(s) o d'un genoma complet:

- (i) Entrecreuament desigual: porta com a conseqüència la duplicació de la zona implicada (en el cas que ens interessa, d'un gen) en tàndem. Podem pensar en la creació de famílies gèniques en regions cromosòmiques específiques.
- (ii) Transposició: explicaria la duplicació discreta d'un gen a un altre lloc del genoma i no en tàndem, a diferència del mecanisme anterior.
- (iii) Poliploidia (auto o alopoliploidia): mecanisme que explica la DGS i pel qual un organisme diploid esdevé tetraploid. És un fenomen d'ocurrència esporàdica en l'evolució dels animals. Pot arribar a ser comú en determinats llinatges com el dels peixos ossis (Schultz, 1980), i rar en d'altres com el d'insectes (Lokki i Saura, 1980).

3.2. Com mantenir un gen duplicat?

No està massa clar com dues còpies d'un gen duplicat poden passar d'un estat en que són totalment redundants -en el que una còpia és per tant prescindible- o fins i tot deletèries, a una situació estable en la que les dues còpies es mantenen per selecció natural. Hi ha quatre vies alternatives després d'un fenomen de duplicació (Lynch i Conery, 2000):

- no-funcionalització: una de les còpies es silencia per mutacions degeneratives (de manca de funció).
- neofuncionalització: una de les còpies adquireix una nova funció (que serà mantinguda per selecció natural) i l'altra manté la funció original.
- subfuncionalització: ambdues còpies acumulen mutacions degeneratives que fan que només la suma de les funcions d'ambdues sigui igual a la funció original (repartiment de X funcions entre les dues còpies). Així, és necessari el manteniment d'ambdues còpies al genoma.
- també és possible una combinació entre el segon i el tercer mecanisme.

El model clàssic només contemplava que les dues còpies es mantinguessin si una d'elles adquiria una nova funció (neofuncionalització). Com que la taxa de mutació cap a al·lèls nuls és més elevada que la incidència de mutacions que confereixen una nova i beneficiosa funció, en la majoria de models matemàtics desenvolupats és molt més probable que els gens esdevinguin silenciosos, que no que adquireixin noves funcions (Kimura i King, 1979). Com que aquests models matemàtics afavoreixen el concepte de no-funcionalització, les dades empíriques no queden resoltes: la majoria de genomes eucariotes contenen un elevat nombre de gens duplicats funcionals, que s'han originat fa desenes o centenars de milions d'anys (Postlethwait i col.,

1998). Altres estudis però, demostren que si el tamany poblacional és suficientment gran, el destí de molts dels duplicats serà el d'adquirir una nova funció (Walsh, 1995).

Per tant, calen altres mecanismes (diferents a la neofuncionalització) per a explicar la preservació de la redundància. Alguns han estat proposats, com un benefici acumulatiu de les còpies del gen (efecte de dosi), o l'efecte de fenotips dominants negatius en el cas de proteïnes amb multidominis (en les que una mutació puntual tendeix a ser dominant i per tant deletèria (Gibson i Spring, 1998). Un dels més acceptats ha estat el model DDC (de l'anglès *Duplication-Degeneration-Complementation*) que pretén explicar el fenomen de la subfuncionalització abans esmentat.

En un estudi molt elegant, Force i col. (1999) proposen el model DDC basant-se en gens duplicats de peix zebra (gens engrailed: *eng1*, *2*, *3* i *1b*), de blat de moro (*ZAG1* i *ZMM2*) i de ratolí (*Hoxa1* i *Hoxb1*). Com a exemple, el que fa referència als gens engrailed: Els tetràpodes tenen dos gens engrailed (*En1* i *En2*), mentre que el peix zebra en té quatre (*eng1*, *eng2*, *eng3* i *eng1b*). Anàlisis filogenètics demostren que els fenòmens de duplicació que produeixen els parells *eng1/eng1b* i *eng2/eng3* de peix zebra tenen lloc al seu llinatge després de la separació del de tetràpodes, perquè els gens *eng1* i *eng1b* queden com a semi-ortòlegs del gen *En1*, mentre que *eng2* i *eng3* queden com a semi-ortòlegs d'*En2* (Fig. 2A). En mapar aquests gens troben que el peix zebra té dues còpies de segments cromosòmics grans al voltant dels gens engrailed, i que aquestes regions sintèniques estan conservades en genomes de mamífers (Fig. 2B). En comparar l'expressió d'aquests gens respecte els de ratolí se n'adonen d'aquest fenomen de subfuncionalització: mentre que *eng1* s'expressa a l'esbós de l'aleta pectoral, *eng1b* ho fa en un conjunt de neurones específiques del cervell posterior/medulla espinal (Fig. 2C). En ratolí, el gen *En1* s'expressa en ambdós dominis.

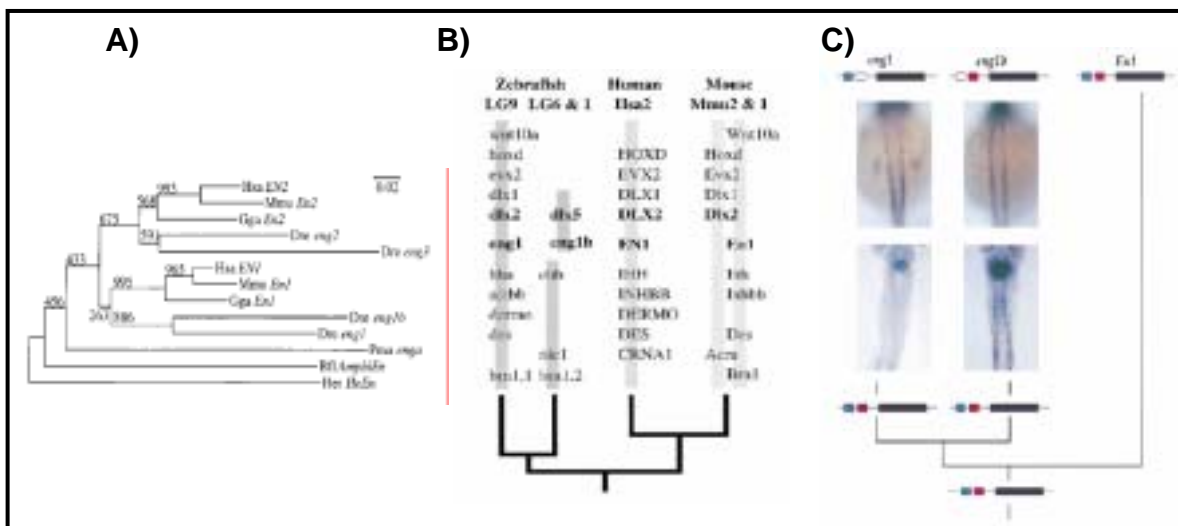


Figura 2. A) Arbre filogenètic dels gens engrailed de peix zebra (Dre), humans (Hsa), de ratolí (Mmu) i de pollet (Gga). B) Regions sintèniques dels cromosomes de peix zebra (LG9, LG6 i LG9), de l'humà (Hsa 2) i de ratolí (Mmu 2 i 1) que contenen els gens engrailed. C) Expressió dels gens *eng1* i *eng1b* de peix zebra, on es mostra un esquema amb l'hipòtesi de l'escenari evolutiu que ha pogut tenir lloc: els rectangles colorejats representen expressió (blau= element regulador de l'aleta pectoral, i vermell= element regulador en neurones), mentre que els buits representen manca d'expressió (extret de Force i col., 1999).

A partir d'aquest treball altres han estat desenvolupats amb la finalitat de refinar el model DDC i determinar paràmetres que facilitin/dificultin el fenomen de subfuncionalització. Algun

d'aquests paràmetres és el tamany poblacional i la quantitat de regions reguladores independents del gen duplicat en qüestió (Lynch i Force, 2000). Aquest model, que ha estat originàriament basat en la degeneració de regions cis reguladores dels gens duplicats, pot aplicar-se també a les regions codificants (Dermatzakis i Clark, 2001). Per exemple quan un dels duplicats perd la funció en un dels dos (o més) dominis de la proteïna, mentre que l'altre la perd en un domini diferent.

3.3. Com adquirir una nova funció (neofuncionalització)

Si sota el model clàssic una de les còpies ha de patir un fenomen de neofuncionalització per a que la selecció sigui suficient per a fixar-la i mantenir-la a la població, cal plantejar-se una nova pregunta: quins factors determinen que un gen pugui adquirir una nova funció? L'atzar, en forma de mutació tant en zones reguladores com transcrits, hi juga un paper central. Alguns autors defensen que l'efecte específic d'aquesta mutació en el gen depèn alhora de la mateixa naturalesa del gen. Per tant, l'organització i funció d'un determinat gen influirà en el seu destí (revisat en Shimeld, 1999).

Els gens reguladors i específics de teixit tenen més possibilitats d'adquirir una nova funció després de la duplicació que els gens *housekeeping* (revisat en Cooke i col., 1997). En general, no és possible que aquests últims adquireixin una nova funció per canvis en la seva regulació, ja que ja estan expressats de manera ubíqua. Així, han de patir canvis a la mateixa seqüència proteica per tal d'adquirir una nova funció. De manera oposada, els gens reguladors poden adquirir nous dominis d'expressió i així, molt probablement, noves funcions per simples canvis als promotors i *enhancers* sense necessitat de canviar la seqüència proteica (p. ex., Li i Noll, 1994).

3.4. Duplicació d'un genoma sencer (DGS)

Com a resultat de la poliploidia es duplica tot el complement genètic d'un organisme, fet que podria ser compatible amb la retenció de molts dels gens duplicats. S'ha suggerit que en el cas dels vertebrats (suposant que han existit una o dues DGS), aquesta retenció podria haver estat del 50% (Nadeau i Sankoff, 1997). Tenint en compte les dificultats amb les que es troben els duplicats per a fixar-se (citades anteriorment), aquest és un valor molt alt. Altres vegades aquesta retenció no és tan elevada: per exemple, la DGS patida al llinatge de llevats ha estat seguida per la pèrdua de més del 90% dels nous gens duplicats (Wolfe i Shields, 1997).

A diferència de les duplicacions d'un sol gen o d'un nombre discret de gens, les poliploidies doblen xarxes genètiques al complet. Per tant, el nombre relatiu de gens que codifiquen per a un determinat component d'una xarxa i el nivell relatiu d'expressió d'aquests és el mateix, quan es compara amb els altres components de la mateixa xarxa (Lynch i Conery, 2000). Aquest equilibri dels components probablement explica com és que molts vertebrats són capaços de sobreviure com a tetraploids, mentre que la duplicació de quantitats més petites de DNA -com les trisomies- tingui conseqüències tan i tan severes pel desenvolupament.

4. L'AMFIOX

Pertany al *Phylum* dels *Chordata*, que comprèn els *Subphyla* dels *Urochordata* (*Tunicata*), *Cephalochordata* (amfiox) i el dels *Vertebrata* (Brusca i Brusca, 1990). Com a cordats tenen tota una colla de característiques que els defineixen, tals com l'existència d'un cordó nerviós buit dorsal a la notocorda, de blocs musculars laterals i d'una cua post-anal, però els manquen

estructures típiques de vertebrat com l'existència de cèl·lules de la cresta neural i d'esquelet, entre d'altres.

Els cefalocordats es coneixen col·loquialment com a *amphioxus* (amfiox en català) o *lancelets* (llancetes en català, degut a la forma allargada del seu cos) (Fig. 3). Se'ls ha definit com a "petits filets d'anxova pàl·lids" (Gee, 1994). Es coneixen unes 29 espècies dioiques cosmopolites, repartides en tan sols dos gèneres (*Branchiostoma* i *Epigonichthys*), que es troben distribuïdes al llarg de les costes marítimes temperades. En aquest treball s'ha utilitzat una única espècie -*Branchiostoma floridae*- que és de la que molt probablement es tenen més dades. Els individus són recol·lectats a Tampa Bay, Florida, USA (Holland i Holland, 1993; i següent apartat).

Figura 3. Fotografia d'amfiox adults, en la que s'aprecia el seu caràcter dioic: els mascles presenten les gònades de color cremós (imatge inferior) i les femelles groguenc (superior). Extret de la pàgina (<http://fukumoto.nsc.nagoyacu.ac.jp/wwwdata/namekuje.htm>)



4.1. Obtenció d'embrions de *Branchiostoma floridae*

Els individus de *B. floridae* creixen des de 4 mm (quan s'acaben d'assentar a la sorra) fins a adquirir una mida màxima de 58 mm, arribant a la maduresa sexual a una longitud mitja de 22 mm. Alliberen ous cada 15 dies aproximadament durant l'època de reproducció que és des de finals de maig fins a principis de setembre. Les femelles tenen una mitjana de 26 parells d'ovaris quan estan plenament madures i és llavors quan el teixit gonadal pot arribar a ser el 20% de la massa corporal total (Stokes i Holland, 1996).

Els adults es van a recol·lectar al mar cap al migdia. Viuen ensorrats a una profunditat que, depenent de la marea, és d'entre 0,5 i 2 m. Amb una pala s'agafa sorra que es passa per un sedàs de mida apropiada (330-1000 µm) per descartar la sorra i on els amfiox, entre d'altres animals, queden retinguts. L'hora en que acostumen a posar ous al mar es després de la posta de Sol. Per això, és en aquest mateix moment en que són separats al laboratori en vasos de plàstic per sexe (uns tres individus per vas) i se'ls aplica un xoc elèctric (de 50V en pulsos de 10 msegons [9 pulsos/seg] durant uns 4 segons aproximadament). Si aquella nit haguessin posat ous al mar, també ho faran al laboratori després del xoc, i si no ho anaven a fer al mar, tampoc ho faran al laboratori. Encara que hi ha hagut molts estudis sobre les condicions òptimes per a la posta, fins al moment no s'ha trobat cap correlació entre una variable com ara la temperatura, salinitat o marea, i l'existència d'una posta (Stokes i Holland, 1996). Si hi ha posta, s'obtenen embrions per fecundació *in vitro* i es fixen els principals estadis de desenvolupament.

4.2. Anatomia de l'adult

El seu *Bauplan* mostra moltes qualitats intermèdies entre aquelles d'invertebrats i de vertebrats. És per això que s'ha dit sobre ells que podrien representar "descendents vius" de l'ancestre dels vertebrats (Brusca i Brusca, 1990). La figura 4 mostra un esquema de la morfologia de l'amfiox adult, on s'indiquen la majoria de les parts a les que s'anirà fent referència en aquest apartat de la introducció.

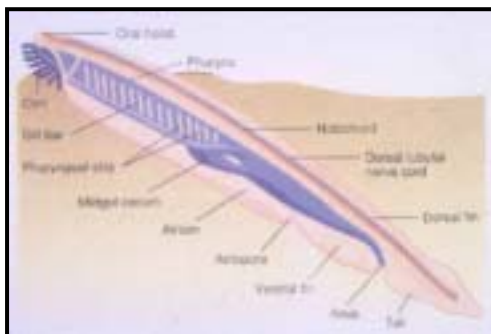


Figura 4. Esquema de la morfologia d'un amfiox adult. Extret de Harris, 1996.

4.2.1. Paret cel·lular, celoma, suport i locomoció

El seu cos està recobert per una epidermis en forma d'epiteli columnar simple, a sota de la qual hi ha un teixit connectiu de dermis molt prim. Els músculs de la paret corporal -que tenen forma de V- són clarament similars als de vertebrats i s'anomenen miòtoms. Aquests ocupen la majoria de l'interior del cos, fent que el celoma quedi molt reduït. Una notocorda -que s'extén fins a la part més anterior de l'animal- persisteix en l'adult i li dóna el major suport estructural al cos.

La locomoció es duu a terme gràcies a la notocorda, als miòtoms segmentats al llarg del cos i a la presència d'una aleta caudal vertical. L'acció propulsora produïda per moviments corporals gràcies als miòtoms, s'amplifica amb ajuda de l'aleta caudal. La notocorda, com que és elàstica i flexible, permet que el cos no és faci més curt al contraure's els músculs, causant enlloc, un *zig-zag* lateral.

4.2.2. Alimentació i digestió

S'alimenten de partícules en suspensió que provenen de la filtració d'aigua. L'aigua entra per la boca i la faringe i surt a través de les fenedures branquials (unes 200 en cefalocordats) a l'atri que les envolta. La funció principal de les fenedures branquials en l'amfiox és l'alimentació; tenen poc paper en l'intercanvi de gasos. L'aigua surt del cos a través del atriopor que es troba en una posició ventral. Les corrents per a captar el menjar són produïdes pels cilis faringis.

Un component essencial per a l'alimentació és el moc, del que tres òrgans en són principalment secretors: (i) La *Hatschek's pit* secreta moc que atrapa les partícules de menjar i les porta cap a dins de la boca. (ii) L'endostil, que es troba a la superfície ventral de la faringe, produeix moc que atrapa el menjar de l'aigua a mida que va passant per les fenedures faríngies dirigint-se cap a l'atri. (iii) La *club-shaped gland*, igual que l'endostil, també es situa dins la faringe. Està estretament associada amb aquest tant espacialment com funcionalment (Fig. 5).

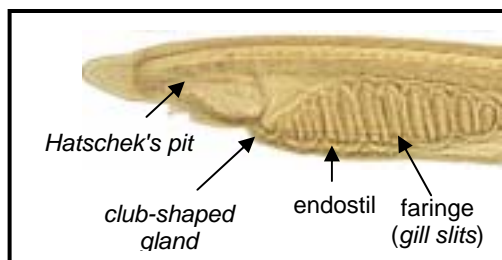


Figura 5. Estructures de la regió anterior relacionades amb l'alimentació de l'amfiox.

Prop de la unió de la faringe amb l'esòfag hi ha una projecció anterior anomenada cec hepàtic. La seva funció d'emmagatzemament de lípids i glicogen i de síntesi proteica ha dut a molts investigadors a pensar en ell com al precursor del fetge de vertebrats. El menjar captat

pel moc es transporta dorsalment al llarg de les parets de la faringe cap a l'esòfag. L'intestí s'allarga posteriorment i s'obre a través de l'anus, situat just enfront de l'aleta caudal. La digestió es comença a fer extracel·lularment a la llum de l'intestí i es completa intracel·lularment a la parets del mateix i al cec (on també s'emmagatzemen els nutrients).

4.2.3. Circulació, intercanvi de gasos i excreció

El sistema circulatori està molt desenvolupat. Està format per un conjunt de vasos tancats pels quals la sang flueix de manera molt semblant a com ho fa en vertebrats primitius. No té cor. La sang flueix posteriorment al llarg de la regió faríngea per un parell d'aortes dorsals que s'uniran en una aorta media dorsal just posterior a la faringe que s'extén cap a la regió de l'aleta caudal. La sang arriba als miòtoms i a la notocorda a través d'una sèrie d'artèries segmentals curtes, i a l'intestí a través de les artèries intestinals. Un conjunt de capil·lars a la paret de l'intestí recol·lecten la sang plena de nutrients. Aquestes donen lloc a les venes intestinals, que alhora s'uniran donant lloc a una gran vena subintestinal que rep el nom de vena porta. Els nutrients van de la vena porta cap a un altre conjunt de capil·lars situats al cec hepàtic. Dels capil·lars del cec se'n deriva la vena del cec, que serà unida per un parell de venes cardinals que tornen dels teixits. Aquests vasos s'uneixen per formar-ne un de sol, l'aorta ventral, que corre per sota la faringe. Des d'aquí la sang és portada a través de les barres faríngies via artèries branquials aferents i eferents cap al parell d'aortes dorsals completant-se, així, el cicle.

La sang no conté pigments ni cèl·lules i està més implicada en la distribució de nutrients que en l'intercanvi de gasos i el transport. Encara que hi pot haver una mica de difusió d'oxigen i diòxid de carboni a través de les fenestres faríngies, la majoria de l'intercanvi gasós té lloc a través de les parets dels solcs metapleurals, que són unes evaginacions primes de la paret corporal situades just anteriorment a l'atriopòr.

El sistema excretor consta d'uns 90 parells de protonefridis situats dorsalment a la regió faríngea, que estan en associació amb les dues aortes dorsals (Whittaker, 1997).

4.2.4. Sistema nerviós i òrgans sensorials

El seu sistema nerviós central és relativament simple. Està format per un cordó nerviós dorsal que s'allarga al llarg de quasi tot el cos, amb una petita expansió anterior que rep el nom de vesícula cerebral. Del cordó nerviós surten nervis de manera segmental, amb el típic patró en arrels dorsals i ventrals que trobem en vertebrats.

L'epidermis presenta molts acabaments de nervis sensorials, la majoria dels quals són probablement tàctils i importants pel l'ensorrament. També tenen un sol ull senzill ciclopi prop del final anterior del cordó nerviós dorsal.

4.3. El desenvolupament de l'amfiox

El desenvolupament de l'amfiox ha estat detalladament descrit en tres tractats principals, que per ordre cronològic són: Hatschek (1893), Cerfontaine (1906) i Conklin (1932). En ser tan antics, encara que molt acurats, a vegades resulten força incomplets. Per exemple, no es disposa de cap mapa de destí de la blàstula d'amfiox, com es té per a la d'ascídies, encara que alguns autors asseguren que aquest és semblant entre aquests dos cordats pre-vertebrats (Conklin, 1932). A la taula I es resumeixen els principals fenòmens que tenen lloc durant el desenvolupament de *B. floridae* a 25°C.

Taula I. Temps i estadis de desenvolupament a 25°C per *B. floridae*

TEMPS	ESTADI
0 h	inseminació
55 min	2 cèl·lules
75 min	4 cèl·lules
95 min	8 cèl·lules
3,5 h	blàstula
5,5 h	gàstrula
10 h	eclosió de la nèurula
38 h	obertura de la boca larvària
18 d	Inici de la metamorfosi
23 d	metamorfosi completa

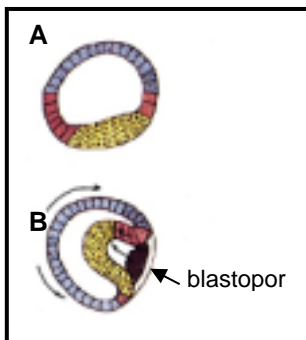
Adaptat de Holland i Holland, 1993.

4.3.1. Crivellament i gastrulació

L'ou de l'amfiox és petit (100-120 µm) i isolecític. El seu crivellament és holoblàstic i radial. Pateix divisions aproximadament sincròniques fins a l'estadi de 128 cèl·lules (7a divisió). La 8a divisió marca el començament de l'estadi de blàstula (3,5 hores post-fertilització, 25°C). Aquesta blàstula és esfèrica i buida.

Al començament de la gastrulació, la blàstula esdevé una mica aplanada per un cantó (Fig. 6A). Al final de la gastrulació s'ha convertit en un embrió de dues capes cel·lulars de gruix, amb un petit blastopor que marca la part posterior (5,5 hores post-fertilització, 25°C) (Fig. 6B). La capa interna consta de (veure també Fig. 7):

- cordamesoderm dorsalment,
- mesoderm diferent al de la notocorda dorso-lateralment i posteriorment,
- i endoderm antero-ventralment.



Cada cèl·lula de la capa externa ha desenvolupat un cili que permet el moviment de l'embrió dins la membrana de l'ou (Hatschek, 1893). Poc temps més tard, aquesta membrana es trenca (10 hores post-fertilització, 25°C) i l'embrió segueix rotant en l'aigua alhora que neda. Aquesta capa externa formarà l'ectoderm i el neuroectoderm.

Figura 6. Esquema de l'estadi de blàstula (A) i de gàstrula (B) de l'embrió d'amfiox. El codi de colors representa a quin tipus de teixit donarà lloc (blau= ectoderm, vermell=mesoderm i groc= endoderm). Adaptat de Langeland i col., 1998.

Fins a l'estadi de gàstrula, el seu desenvolupament és genèricament típic de deuteròstom no vertebrat. Almenys superficialment, és indistingible del desenvolupament típic d'un equinoderm. De manera oposada, els següents estadis són marcadament similars als d'un vertebrat, amb l'aparició d'un cordó nerviós dorsal, d'una notocorda i de somites musculars durant l'estadi de nèurula. (Holland i Holland, 1999).

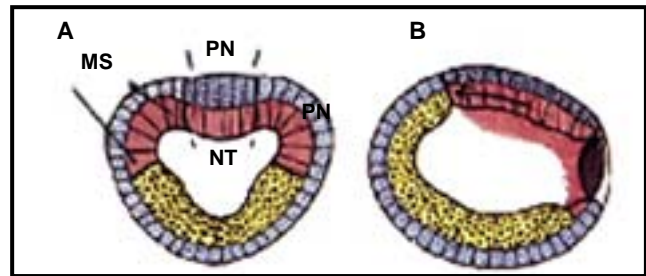
4.3.2. La neurulació (Fig. 7)

És l'enfonsament del neuroectoderm per a formar el cordó nerviós dorsal. Aquest fet comporta un lleuger aplastament dorsal de la gàstrula, que esdevé un ampli solc en forma de V que acaba posteriorment a nivell del blastopor. Aquesta V es va fent més obtusa fins que al final

només una petita obertura, el neuropor, queda prop del final més anterior de l'embrió. El terra del sòlc original, que serà el futur tub neural, queda llavors recobert per una fila d'ectoderm. Un cop acabada la neurulació, l'embrió rep el nom de nèurula.

Mentrestant, la capa cel·lular interna de la nèurula està patint molts canvis. Tot al llarg de la línia mitja dorsal de l'arquènteron, l'epiteli s'evagina dorsalment per a formar la notocorda. Just a dreta i esquerra de la notocorda, l'epiteli de l'arquènteron s'evagina dorso-lateralment per a formar els somites mandibulars i els somites més posteriors. Les parts més ventrals de l'epiteli de l'arquènteron donaran lloc a l'intestí i al primer parell de somites o somites pre-mandibulars.

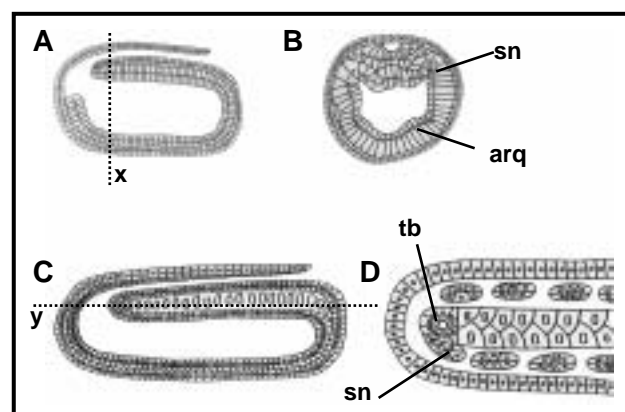
Figura 7. Esquema [tall transversal (A) i longitudinal (B)] de la neurulació d'amfiox on s'aprecia l'aplanament de la placa neural (PN) i de la formació de la notocorda (NT) i del mesoderm somític (MS) a partir de l'arquènteron. El codi de colors és el mateix que en la figura 5. Adaptat de Langeland i col., 1998.



4.3.3. La somitogènesi

Els somites es formen en sentit antero-posterior a mida que l'embrió creix en el mateix sentit. Els somites més anteriors i, per tant, els primers que s'han format, produeixen el celoma per enterocèlia a partir de la gemmació de la paret dorso-lateral de l'arquènteron (Fig. 8A i B), mentre que els més posteriors ho fan per esquizocèlia a partir del *tail bud* (Fig. 8C i D). El nombre de somites de cada tipus varia segons l'autor. Una característica dels somites de l'amfiox és la seva asimetria: els somites de la banda esquerra estan disposats una longitud aproximada de 1/2 somita més rostrals que els de la dreta (Fig. 8D). A partir de quin somita es fa palesa aquesta asimetria ha estat també en debat; Conklin (1932) i Hatshek (1892) coincideixen en defensar que comença a partir del 7è o 8è parell de somites, mentre que Cerfontaine (1906) afirma que aquesta asimetria és present ja des del primer parell.

Figura 8. Representació esquemàtica de les dues fases de somitogènesi de l'amfiox. En la primera (A, B), els somites naixents (sn) s'originen a partir de l'evaginació de les parets dorso-laterals de l'arquènteron (arq). En la segona (C, D) s'originen a partir del *tail bud* (tb). En les visions laterals (A, C) anterior és cap a la dreta i dorsal cap a dalt. B representa un tall transversal a través del nivell x en A amb dorsal cap a dalt. D representa la visió dorsal d'un tall frontal a través del nivell y en C, amb anterior cap a la dreta.



Al seu origen, la notocorda no s'extén fins al final frontal de la capa interna de la nèurula. Una vegada s'han format els 9 o 10 primers parells de somites, la porció d'arquènteron que queda davant de la notocorda es dilata cap a dalt a mida que aquesta va creixent cap endavant, formant els diverticles intestinals anteriors o somites pre-mandibulars.

Els somites dorso-laterals es comencen a estendre a cada cantó del arquènteron fins a trobar els seus antimers a la línia medial ventral per sota de l'intestí. Llavors es fusionen trencant les parets transversals entre ells. Això fa que els signes de segmentació persisteixin només dorsalment. Alhora, els somites pre-mandibulars s'han separat totalment del arquènteron i esdevenen asimètrics. Més tard -en la vida larvària- l'esquerre desenvoluparà una obertura cap a l'exterior just anterior a la boca: la *Hatschek's pit*. El de la part dreta forma el celoma anterior de la larva, que esdevindrà el celoma rostral de l'adult.

Al final de la neurulació i principi de l'estadi de larva, l'animal ha esdevingut un representant quasi perfecte de la forma comú i generalitzada del cos embrionari primitiu dels embrions de vertebrats, que d'acord amb Nelsen (1953) té 3 característiques:

- És una estructura elongada, cilíndrica i una mica comprimida lateralment.
- Té quatre tubs epitelials formadors d'òrgans (epidèrmic, neural, mesodèrmic i endodèrmic) que estan orientats al voltant d'un eix primitiu: la notocorda
- Té diferents regions corporals (cap, àrea faríngea, tronc i cua).

4.3.4. Desenvolupament de la faringe i estructures associades

A la part anterior de la larva tenen lloc molts fenòmens simultanis (Whittaker, 1997):

- Formació de la boca larvària (38 hores post-fertilització, 25°C): es forma un engruiximent a la part ventral de la paret esquerra de l'arquènteron que es fusiona amb l'ectoderm que el contacta. La boca apareix com un forat circular a la banda esquerra de la larva. L'aparició de la boca defineix l'inici de l'etapa larvària.
- De manera similar, l'engruiximent de la paret dreta dóna lloc a la *club-shaped gland*. Encara que aquesta es desenvolupa al cantó dret de la faringe, s'obre a l'exterior pel cantó esquerre.
- Formació de les fenedures: apareixen també per un engruiximent, en aquest cas a la línia mitja ventral al tercer parell de somites. El centre de l'engruiximent es fusiona amb l'ectoderm que el contacta i es desenvolupa un forat que constitueix la fenedura en sí. Hi ha tres grups de fenedures en funció de quan s'originen (Willey, 1894): (i) Les primàries es formen a la larva i es situen al cantó esquerre. (ii) Les secundàries es formen durant la metamorfosi i morfològicament són al cantó dret. (iii) Les terciàries es formen a l'adult i es situen als dos cantons.
- Formació de l'endostil: apareix com a un engruiximent de la faringe just anterior a la *club-shaped gland*.

4.3.5. La metamorfosi

L'aparició de les fenedures secundàries marca l'inici de la metamorfosi (18 dies post-fertilització, 25°C). És el procés pel qual la larva asimètrica dóna lloc a un adult que és bàsicament simètric. Per exemple, té lloc la migració de la boca -que estava al cantó esquerre de la larva- cap a la regió ventro-medial. En comparació amb altres invertebrats i cordats en els que la metamorfosi representa un fenomen que comporta molta "crisi", en l'amfiox el patró adult emergeix gradualment. Es completa la metamorfosi als 23 dies post-fertilització a 25°C. És ara quan els juvenils tornaran a ensorrar-se, tancant així el cicle vital.

L'asimetria ha estat usada com a argument clau per a mostrar les afinitats de l'amfiox amb els invertebrats. Altres autors però, veuen aquestes asimetries com a conseqüència d'adaptacions

especials de la larva que poden ser explicades com a part de mecanismes adaptatius a l'alimentació (procés caenogenètic) (Presley i col., 1996). Defensen que al llarg de la majoria de fases inicials de desenvolupament l'embrió és simètric i fins i tot durant les fases larvàries asimètriques, molts dels caràcters dorsals i axials retenen la simetria essencial i són molt semblants als dels vertebrats. Al llarg dels diferents estadis de desenvolupament, i de fet continuant fins l'edat adulta, hi ha un invariable, sostingut i constant desenvolupament d'addicionals segments simètrics a la cua.

4.3.6. Consideracions filogenètiques respecte el desenvolupament de l'amfiox

L'estadi de *Bauplan* correspon a aquell nivell d'organització que és més estable entre organismes relacionats. Per tant, podria dir-se que el verdader *Bauplan* dels cordats estaria representat, més que per la larva, per la forma adulta de l'amfiox (Presley i col., 1996).

El fet de que els amfiox retinguin alguns caràcters que són típics d'invertebrats com l'existència de gònades segmentals, connexions musculars al cordó espinal enlloc d'axons de motoneurons, especialitzacions tals com corpuscles a la sang o endoteli, etc., pot simplement oferir pistes per a la traça de la seva pròpia ancestralitat dins dels invertebrats. L'amfiox també mostra especialitzacions úniques com l'atri, la natura muscular de la notocorda, entre d'altres, que probablement representen caràcters de cordats primitius.

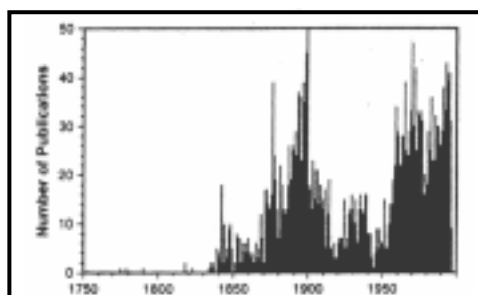
A diferència de tots els anteriors, trobem caràcters que són clarament típics de vertebrats, essent tots ells molt suggestius per a situar l'amfiox com a organisme molt proper a aquests. El més important és l'homologia del procés de desenvolupament comú que posiciona antero-posteriorment els segments bilaterals junt amb la notocorda i el tub neural a la línia mitja. Aquest fet no s'observa ni en equinoderms ni en tunicats. En aquest aspecte, els amfiox s'ajusten totalment als requeriments d'un cordat ancestral. El *Bauplan* de l'amfiox representaria la forma més plausible que havia de tenir l'ancestre comú de cefalocordats i vertebrats (Presley i col., 1996).

Aquesta conclusió està també recolzada pel registre fòssil. Un ancestre tal hauria de semblar-se a algun dels fòssils càmbrics analitzats. El fòssil *Pikaea gracilens* de la fauna de Burgess Shale (Càmbric mitjà) està interpretat com a cordat, més concretament, com a cefalocordat i és sorprenentment semblant als amfiox (Conway Morris, 1982). El fòssil xinès de la fauna Chengjiang *Cathaymyrus diadexus* predata a l'anterior en 10 milions d'anys (Càmbric inferior) i, igual que *Pikaea*, ha estat interpretat com a cefalocordat (Chen i col., 1995).

4.4. Posició filogenètica de l'amfiox

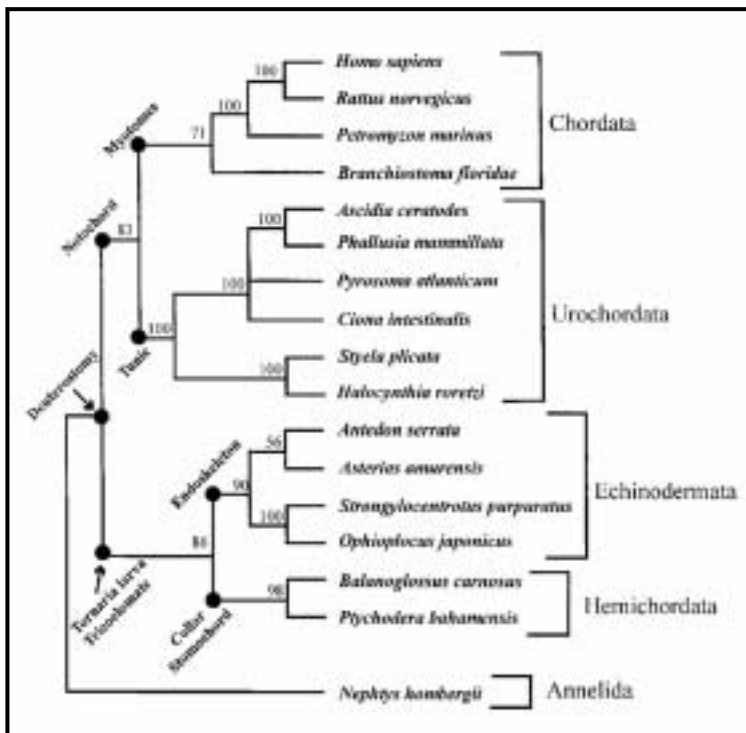
L'interès per l'amfiox ha estat molt discontinu al llarg dels anys, pel que fa referència a intentar respondre la qüestió de com els vertebrats han evolucionat a partir dels invertebrats (Fig. 8).

Figura 8. Gràfica del nombre de publicacions científiques respecte l'amfiox al llarg dels segles XVIII a XX. A partir de Gans, 1996.



Durant les últimes dècades del Segle XIX, es van publicar centenars d'articles sobre l'amfiox en els que es descrivia l'anatomia de l'adult i el seu desenvolupament. Això va ser segurament a causa de la passió que el conegut biòleg alemany Ernst Haeckel sentia per ells, situant-los en una privilegiada i clau posició filogenètica com a invertebrat ancestre immediat dels vertebrats. Aquesta visió però, no era compartida per tota la comunitat científica del moment. L'amfiox també havia estat considerat com a vertebrat simplificat per degeneració (Anton Dohrn) o fins i tot com a proto-molusc (Karl Semper). L'estudi del seu desenvolupament amb característiques pròpies d'invertebrat al principi, però donant lloc a la producció de caràcters molt propers als de vertebrats (Kovalevskii), era un dels motius pels quals Haeckel venerava aquests animals (per ell aquest fet revelava una connexió evolutiva entre els invertebrats -l'amfiox- i els vertebrats).

A principis del Segle XX, les publicacions que feien referència a l'amfiox van caure dràsticament i, no és fins als anys 60, amb els avenços en microscopia electrònica, que va resorgir l'interès per aquestes criatures. En un llibre de text del 1962 Grove i Newell diuen sobre



ell: "Si l'amfiox no hagués estat descobert, hauria d'haver estat inventat". Un segon resorgiment té lloc fa uns deu anys quan es comencen a aplicar mètodes de Biologia Molecular. La visió actual i àmpliament acceptada és la que els cefalocordats es situen com a grup germà dels vertebrats en l'arbre filogenètic (Fig. 9) (Wada i Satoh, 1994; Cameron i col., 2000).

Figura 9. Arbre filogenètic dels deuteròstoms. Extret de Cameron i col., 2000.

4.5. L'amfiox i les innovacions dels vertebrats

Així doncs, si es vol estudiar com els vertebrats han estat capaços d'"inventar" tota una colla de caràcters que els són únics -caràcters derivats- des del punt de vista de l'Evo-Devo, cal escollir un organisme invertebrat que, tot i ser el més proper i semblant a ells li mancaran, per tant, totes aquestes innovacions. Proves que provenen de diferents fonts (anatomia, desenvolupament, filogènia molecular i registre fòssil) indiquen que aquest organisme és l'amfiox. Com que són els parents vius més propers als vertebrats, les comparacions dels gens o genoma d'aquests amb els de vertebrats ens han de donar pistes dels fenòmens genètics que acompanyen l'origen de les innovacions morfològiques dels vertebrats.

Com ja s'ha esmentat, s'ha hipotetitzat que dos fenòmens de DGS van ocórrer a la base de l'evolució dels vertebrats just després de la divergència de l'amfiox, ja que molts dels gens humans (com a representats de vertebrats) són presents en conjunts de 4 paràlegs (per a una revisió veure Spring, 1997). Aquesta hipòtesi però, ha estat debatuda per altres (Skrabaneck i Wolfe, 1998; International Human Genome Sequencing Consortium, 2001), que afavoreixen l'existència de duplicacions no relacionades de regions específiques, més que de successives DGS (per a una revisió dels arguments en contra i a favor veure Wolfe, 2001).

Sigui quina sigui la forma de duplicació que hi està implicada (de tot el genoma, o de determinades parts), el que sí que està clar i acceptat és que per a la majoria de famílies gèniques trobem més representants en els genomes de vertebrats, que en el genoma d'amfiox. Així doncs, s'ha anomenat al genoma de l'amfiox com a arquetípic (semblant però més senzill, almenys pel que fa referència al nombre de gens) en comparació al dels vertebrats (García-Fernández i Holland, 1994).

Per tant, esdevé plausible la següent aproximació: per comparació de l'expressió del (en principi) únic gen X d'amfiox amb els seus semi-ortòlegs X1-Xn de vertebrats, podem extreure les següents conclusions:

- I. Aquells llocs d'expressió comuns als dos representen estructures homòlogues [encara que aquest no ha de servir com a únic criteri per assignar homologies (veieu la discussió)], que per tant ja eren presents al seu ancestre comú, en les que també s'hi expressaria el gen ancestral.
- II. Aquells llocs d'expressió que són únics a l'amfiox, estarien mostrant caràcters específics del seu llinatge.
- III. Aquells llocs d'expressió únics al vertebrat en qüestió, serien resultat de la co-opció d'aquests gens en funcions inventades pels vertebrats, que podria haver tingut com a conseqüència l'aparició d'aquests nous caràcters.

5. L'HOMEBOX

Com que la morfologia final dels organismes la dicten els gens que regeixen el seu desenvolupament, per tal de fer una aproximació evolutiva (com s'han donat canvis morfològics al llarg de l'escala evolutiva) cal triar els gens del desenvolupament apropiats a les innovacions (o canvis) que es volen estudiar. Amb aquesta finalitat, una de les famílies de gens del desenvolupament que han estat més estudiades ha estat la dels gens amb homeobox. Això no vol dir que aquests gens siguin els més importants pel canvi morfològic, sinó que com són dels que es tenen més dades en més espècies diferents, un l'estudi comparatiu és fàcilment aplicable.

Ja al 1894, William Bateson defineix l'homeosi com a fenomen que comporta una pèrdua de diversitat deguda a la substitució d'una estructura corporal per un altre de pre-existent (McGinnis, 1994). El primer mutant homeòtic de *Drosophila melanogaster* és descrit 20 anys més tard per Bridges; es tracta del mutant *bithorax*, conegut actualment com a *Ultrabithorax*, en el que el parell d'halteris del tercer segment toràcic ha estat substituït per un parell d'ales característiques del segon segment toràcic, donant lloc a una mosca amb dos parells d'ales (Gehring, 1992). No és fins fa uns 25 anys que Ed Lewis caracteritza genèticament alguna d'aquestes mutacions homeòtiques i se'n adona que els gens causants (gens homeòtics) es troben

arreglats al genoma de la mosca formant un complexe gènic que anomenà complexe *Bithorax*. Proposa que aquests gens constitueixen una família gènica i que s'han duplicat i han divergit del mateix gen ancestral, i que ha estat durant aquest procés quan han adquirit funcions divergents que són les responsables de les diferències morfològiques que permeten distingir el pla corporal de *Drosophila* del d'altres artròpodes més primitius. Hipotetitzava l'existència de 8 gens homeòtics, cadascun responsable de la identitat d'un dels 8 segments abdominals (Lewis, 1978). Quan les tècniques de Biologia Molecular estan a l'abast dels investigadors, el clonatge i la caracterització d'aquests gens homeòtics permeten l'acceptació general del model de Lewis amb algunes consideracions.

La tècnica de *chromosome walking* va permetre l'aïllament dels dos complexos de gens homeòtics de *D. melanogaster*; el *Bithorax* (BX-C) i l'*Antennapedia* (ANTP-C), ambdós localitzats al cromosoma III [el conjunt dels 2 complexos rep el nom de complexe homeòtic (HOM-C)]. Durant el procés de mapatge de diferents parts del cDNA del gen *Antennapedia* (*Antp*) dins l'ANTP-C es detectà una regió d'homologia creuada entre aquest i el seu gen veí *fushi tarazu* (*ftz*), i també amb el gen *Ultrabithorax* (*Ubx*) del BX-C. La posterior seqüenciació de la regió d'hibridació creuada entre aquests gens revelà l'existència d'un segment de DNA de 180 parells de bases molt conservat entre aquests 3 gens: l'*homeobox* (o caixa homeòtica, fent referència a les mutacions homeòtiques de les que aquests gens en són els responsables) que codifica per a un motiu proteic de 60 aminoàcids, l'homeodomini (McGinnis i col., 1984; Scott i Weiner, 1984). L'ús de l'*homeobox* com a sonda, revelà que aquest domini no és exclusiu dels gens homeòtics de *Drosophila*, sinó que també es troba en altres gens de la mosca i fins i tot en gens d'altres organismes.

L'homeodomini està estructurat en forma de tres (o quatre) hèlixα amb un gir que separa la segona de la tercera. Així, les proteïnes que el contenen actuen com a factors de transcripció del tipus *helix-turn-helix*. Aquest domini permet la unió de les proteïnes que el contenen al DNA, on reconeix la seqüència consens ATTA (revisat en Gehring, 1985). D'aquest reconeixement en són principals responsables els residus 47 i 51 de l'hèlix III, residus molt conservats entre les diferents proteïnes amb homeodomini.

5.1. Classificació i evolució dels gens amb homeobox

Els primers gens que codifiquen per proteïnes amb homeodomini identificats van ser els gens homeòtics de *Drosophila*. Des de llavors, s'han publicat centenars de gens amb el mateix motiu no només en *Drosophila*, sinó també en altres animals i fins i tot plantes i fongs; a més, no tan sols gens homeòtics, sinó també d'altres classes de gens amb homeobox. Calia, per tant, una classificació de tots aquests; Thomas Bürglin ha anat mantenint una base de dades amb tots els nous homeodomini que s'anaven descobrint amb aquesta intenció (Bürglin, 1994; Bürglin, 1995; i <http://www.biosci.ki.se/groups/tbu/homeo.html>). Avui en dia, es coneixen 26 classes de gens amb homeobox; al llarg d'aquesta Tesi s'ha treballat exclusivament amb alguns gens de la superclasse *Antennapedia* (*Antp*) que comprèn diferents famílies i classes de gens amb homeobox, que s'agrupen dins la mateixa superclasse per diferents criteris com ara:

- I. similitud de seqüència
- II. disposició en clusters al genoma
- III. presència d'altres dominis conservats com l'hexapèptid a 5' de l'*homeobox*
- IV. conservació de posicions intròniques

La disponibilitat de la seqüència del genoma humà, del de *Drosophila*, de part del de ratolí i dades genòmiques fraccionades d'altres organismes, ha permès l'estudi de la disposició en complexes (*clusters*) de tots els membres de la superclasse Antp (Pollard i Holland, 2000; Coulier i col., 2000). Així, es parla de l'existència de quatre *clusters* (*extended Hox*, *ParaHox*, *EHGbox* i *NKL*) de gens amb homeobox que agrupen tots els gens de la superclasse Antp (Fig. 11). El *cluster extended Hox* conté les diferents classes de gens dels *clusters* Hox (HOM-C de *Drosophila*), la classe *Even-skipped* (*Evs*) i la classe *Mox*. El *ParaHox* conté les classes *Caudal* (*Cdx*), *Xlox* i *Gsx*. El *cluster EHGbox* conté les classes *Engrailed* (*En*), *HB-9* i *Gbx*. Últim i més extens és el *cluster NKL*, que conté les classes *Msx*, *Dlx*, *Not*, *Barx*, *Hex*, *Hmx*, *Emx*, *Vax*, *Tlx*, *Lbx*, i *NK1*, *NK2*, *NK3*, *NK4* i *NK6*.

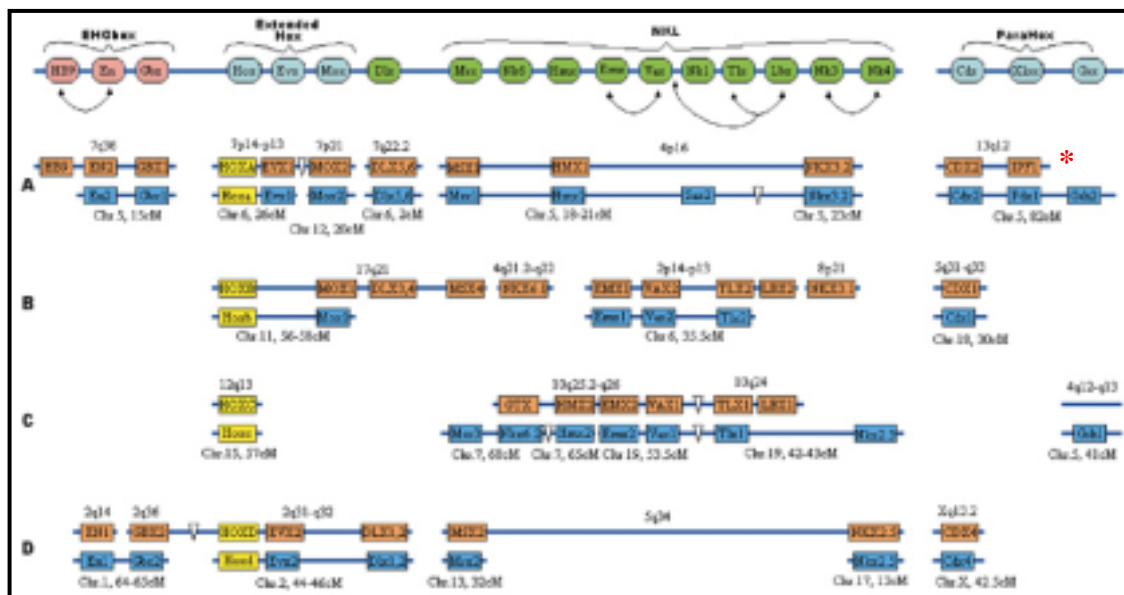


Figura 11. Els quatre *clusters* de gens de la superclasse Antp són esquemàticament representats a la part superior de la figura. A la part inferior, apartats A, B, C i D, es representen les posicions cromosòmiques dels gens de ratolí (en blau) i humans (en taronja). Els clusters Hox (groc) estan representats com si es tractés d'un únic gen, per tal de simplificar l'esquema. Adaptat de Pollard i Holland, 2000. Cal afegir el gen *Gsh1* al cromosoma 13 humà (asterisc vermell).

Dades de lligament entre gens de diferents clusters o entre gens adjacents a aquests han permès establir relacions de lligament entre els clusters en sí. Per exemple, al genoma humà el gen *Msx4* (del tipus NKL) està lligat als gens *Dlx3* i *Dlx4* (tipus NKL), que al seu torn estan lligats al gen *Mox1* (tipus *extended Hox*) i al mateix *cluster* HoxB. En canvi, els seus paràlegs *Msx1* i *Msx2* estan lligats a gens dels tipus NKL. Això podria indicar que el *cluster extended Hox* i el *cluster NKL* estaven originàriament lligats. També s'ha proposat lligament entre els gens *En* i *Gbx* (ambdós del tipus EHGbox) amb els gens *Evs* i *Hox* (els dos del tipus *extended Hox*) (Matsui i col., 1993). Aquest fet indicaria que originàriament tots els clusters de gens de la superclasse Antp, a excepció del *ParaHox*, ocupaven regions genòmiques veïnes. L'escenari evolutiu proposat per Pollard i Holland es presenta a la figura 12. Després d'una sèrie de duplicacions en tàndem, un gen amb homeobox ancestral (groc) dona lloc a un *cluster* de 3 gens (*Archehox*), que són els ancestres dels gens de tipus EHGbox, *extended Hox* i NKL (vermell, blau i verd respectivament). Cadascun d'aquests 3 gens pateix duplicacions en tàndem independents donant lloc a 3 *proto-clusters* a la mateixa regió genòmica. Una duplicació discreta del *cluster* ProtoHox produeix l'aparició del *cluster* ParaHox en una altra regió del genoma. Els *clusters extended Hox* i NKL

pateixen duplicacions en tàndem extra. Més tard, durant l'evolució dels vertebrats i molt probablement per DGS, apareixen els 4 clusters de cadascun dels quatre tipus.

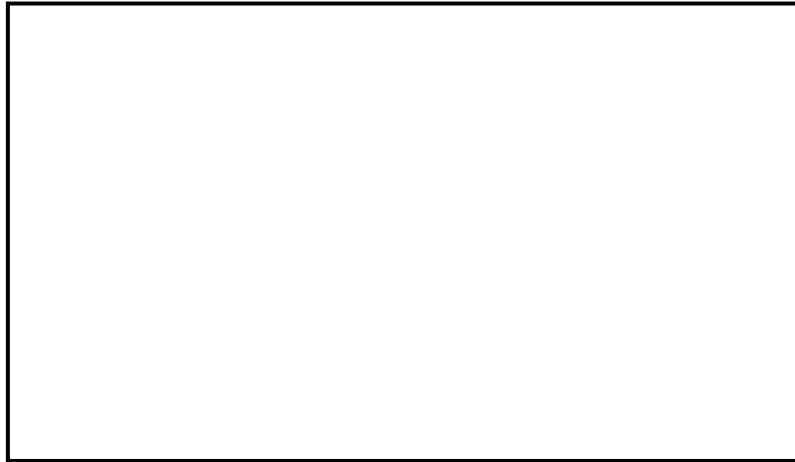


Figura 12. Gènesi i evolució dels gens de la superclasse Antp. Extret de Pollard i Holland, 2000.

5.2. Els gens Hox

Els gens Hox comprenen tres classes de gens amb homeobox (labial [lab], Antennapedia [Antp] i AbdominalB [AbdB]) (Bürglin, 1995), la funció dels quals és la d'especificar la identitat posicional de diferents estructures al llarg de l'eix anteroposterior de la majoria de metazous (McGinnis i Krumlauf, 1992; Slack i col., 1993). Aquests gens tenen dues característiques que s'han mantingut al llarg de l'evolució: (i) estan agrupats al genoma en forma de *clusters* i (ii) presenten el fenomen de la colinearitat pel qual la seva posició relativa dins el *cluster* mimetitza el seu lloc d'expressió al llarg de l'eix antero-posterior de l'embrió (Carroll, 1995). Es divideixen en tres grups: anteriors o 3' (Hox1 a 3 [altres autors consideren només els grups 1 i 2 com a gens anteriors, mentre que els gens Hox3 formarien un altre grup]), centrals (Hox 4 a 8) i posteriors o 5' (Hox 9 a 13), en funció de la seva seqüència, la seva posició relativa dins el *cluster* i el seus dominis d'expressió al llarg del desenvolupament (Schubert i col., 1993).

Els gens Hox no formen un grup monofilètic, sinó que aquest grup està format per ells mateixos més els gens ParaHox (Fig. 13; Kourakis i Martindale, 2000). L'existència d'un *cluster* format per gens ParaHox ha estat documentada en l'amfiox i no ha estat fins fa poc que s'ha pogut identificar *in silico* la seva existència en el genoma humà. L'estudi de l'expressió d'aquests gens en l'amfiox demostra la seva disposició colinear anàloga a la que presenten els gens Hox; és a dir, el gen ParaHox anterior (*Gsx*) s'expressa en territoris més anteriors que el gen que ocupa la regió central al cluster (*Xlox*) i aquest, al seu torn, s'expressa en territoris més anteriors que el gen ParaHox posterior (*Cdx*). Aquest fet ha permès suggerir que aquests dos *clusters* són evolutivament germans, és a dir, que provenen de la duplicació d'un únic *cluster* ancestral anomenat ProtoHox (Brooke i col., 1998; veieu discussió). Així, el *cluster* ProtoHox contindria un màxim de quatre gens: un d'anterior (del qual deriven els gens Hox anteriors i el gen ParaHox anterior), un de posterior (del qual deriven els gens Hox posteriors i el gen ParaHox posterior), un del grup paràleg (PG) 3 (del que deriven els gens Hox del PG3 i el gen *Xlox*) i possiblement un gen ProtoHox central (del que deriven els gens Hox centrals (PG4-8) i un gen ParaHox central que o bé s'ha perdut o no forma part del *cluster* ParaHox.

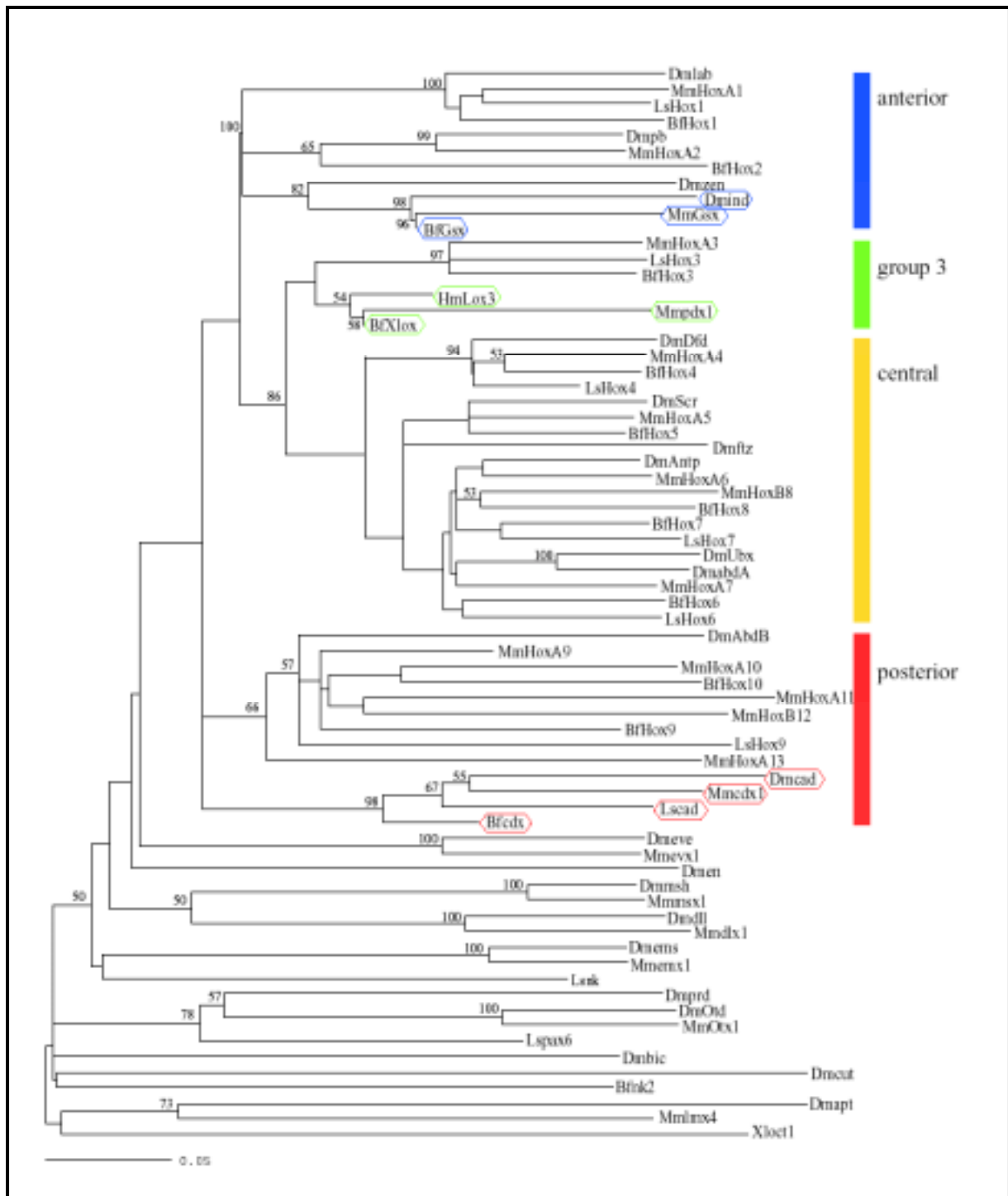


Figura 13. Arbre filogenètic deduit pel mètode neighbor-joining que mostra les relacions filogenètiques dels gens Hox i ParaHox (en rectangles), construït amb l'homeodomini de les proteïnes de ratolí (Mm), *Drosophila* (Dm), amfiox (Bf) i nemertins (Ls). A partir de Kourakis i Martindale, 2000.

5.2.1. Els *clusters* Hox

Com ja s'ha esmentat, els primers gens amb homeobox aïllats van ser precisament els gens homeòtics (Hox) de *D. melanogaster*, que estan disposats en dues regions cromosòmiques: el complex Antennapedia (ANTP-C) i el complex Bithorax (BX-C), que conjuntament reben el nom de complex Homeòtic (HOM-C; Akam, 1989) i actualment el nom genèric de *cluster* Hox igual que en els altres organismes on s'ha aïllat. El *cluster* de *D. melanogaster* està molt

desorganitzat: està trencat, conté gens amb homeobox sense funció homeòtica i conté altres gens que ni tan sols tenen homeobox (McGinnis i Krumlauf, 1992). Aquesta desorganització sembla ser un caràcter derivat, ja que en altres artròpodes com els escarabats s'ha caracteritzat l'existència d'un únic *cluster* (Beeman, 1987). Posteriorment es van aïllar els clusters d'altres organismes com per exemple el del nemàtode *Caenorhabditis elegans* (Kenyon i Wang, 1991) que també està força degenerat, en el sentit de que només conté 4 gens. A més, s'han trobat dos nous gens Hox posteriors (*nob-1* i *php-3*) que no formen part del *cluster* (van Auken i col., 2000). En mamífers hi ha 4 *clusters* Hox (Boncinelli i col., 1988): s'han aïllat els mateixos 39 gens en ratolí i humà, que en base a similitud de seqüència s'agrupen en 13 grups paràlegs (PG). Els gens del mateix PG estan relacionats per duplicació d'un únic cluster ancestral (Holland i Garcia-Fernández, 1996). Cadascun dels *clusters* (anomenats A, B, C i D) està compost per una determinada selecció dels 13 PGs. Aquestes dades van posar de manifest dues grans diferències:

1. L'existència d'un únic complex en organismes invertebrats respecte l'existència de múltiples clusters en mamífers.
2. L'existència de múltiples gens del tipus posterior en mamífers respecte l'existència d'un únic gen Hox posterior en invertebrats (*Abd-B*).

Una fita important en la recerca dels gens Hox la representa l'aïllament de l'únic *cluster* del genoma de l'amfiox, en el que es troba un representant de cadascun dels PGs de vertebrats, almenys fins al PG10 (Garcia-Fernández i Holland, 1994). Aquest treball permet limitar evolutivament les dues diferències anteriors. És a dir, pel que fa referència a l'existència de múltiples clusters, suposa una duplicació d'un únic cluster ancestral molt a prop de l'origen dels vertebrats. Aquest cluster ancestral ha de ser curiosament molt similar al de l'amfiox. També implica que les duplicacions en tàndem a partir d'un únic gen tipus posterior o *Abd-B* que donen lloc a múltiples PGs del tipus posterior en mamífers, tenen lloc abans de l'origen dels vertebrats. Treballs més recents en equinoderms (Martinez i col., 1999) i urocordats (Di Gregorio i col., 1995) limiten aquestes duplicacions a l'origen dels deuteròstoms, ja que ambdós grups tenen ja múltiples gens Hox posteriors. Actualment, també s'accepta l'existència de múltiples gens Hox posteriors en protòstoms (de Rosa i col., 1999).

5.2.2. Homologia funcional dels clusters Hox

En la majoria dels *clusters* Hox analitzats, com l'orientació de la transcripció dels gens és la mateixa (a excepció del gen *Dfd* de *Drosophila*) es pot establir una direcció 5'-3' del *cluster* en conjunt. Com s'ha esmentat, una característica d'aquests gens és la colinearitat espacial, propietat per la qual els gens situats a la part 3' dels *clusters* s'expressen en territoris més anteriors que els gens situats a la part més 5' (Fig. 14). Per tant, no només l'homologia de seqüència i la seva ordenació als complexos indica que els gens Hox inicialment trobats en *Drosophila* i en mamífers són homòlegs, sinó que la propietat de la colinearitat ho reforça (Graham i col., 1989; Duboule i Dollé, 1989). Estudis de l'expressió dels gens Hox en altres espècies indiquen que la colinearitat espacial no és un fenomen exclusiu de mamífers i *Drosophila*, sinó que pot ampliar-se a altres organismes.

Posteriorment, estudis funcionals dels gens Hox tant en ratolí com en *Drosophila* fan palesa la seva homologia a nivell funcional. Així, de manera similar als mutants homeòtics de *Drosophila*, ratolins *knockout* per a algun gen Hox mostren *a grosso modo* transformacions homeòtiques apreciables bàsicament a la columna vertebral (per canvi de la fórmula vertebral)

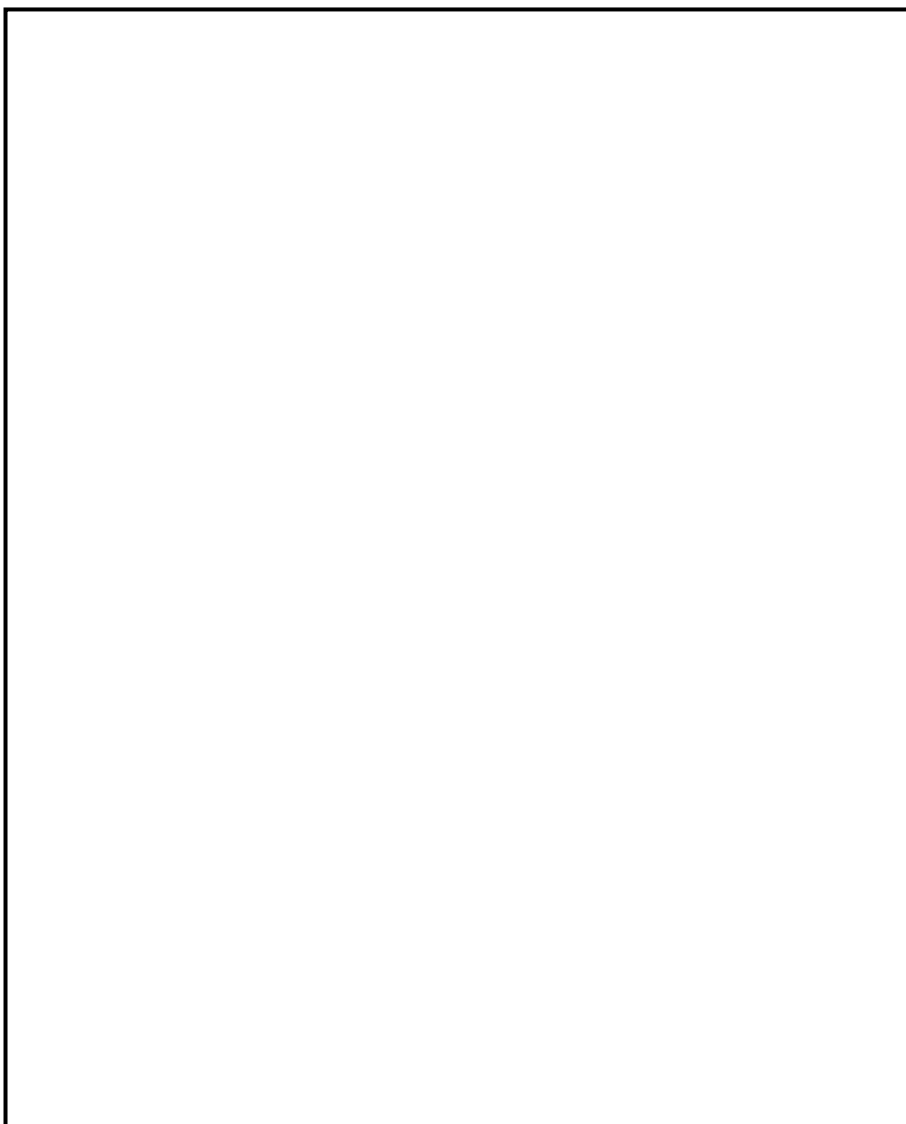


Figura 14. La colinearitat espacial. En un esquema dels clusters de mamífers i *Drosophila* es representa com els gens que ocupen les posicions més 3' del complex (en verd) actuen en parts més anteriors de l'embrió en desenvolupament, mentre que els gens més 5' (en groc) s'encarreguen de les regions més posteriors. A partir de Veraksa i col., 2000.

(Duboule, 1995). Una altra propietat funcional dels gens Hox que està conservada en mosques i ratolí fa referència a la jerarquia funcional entre els membres del complex, per la qual els productes dels gens que ocupen posicions més 5' del complex són prevalents sobre els productes de gens més 3' (Duboule i Morata, 1994). Aquest fet -que rep el nom de supressió fenotípica en *Drosophila* i prevalença posterior en vertebrats- explica un fet característic de l'expressió dels gens Hox: encara que la majoria s'expressen des d'un límit anterior ben definit fins a quasi la part més posterior de l'organisme, la seva manca de funció només té conseqüències a la part més anterior del seu domini d'expressió, allà on no es solapen amb cap gen Hox posterior a ells.

5.2.3. Per què mantenir un *cluster*?

En vertebrats, l'existència de mecanismes de regulació compartits entre diferents gens Hox tals com l'existència d'*enhancers* comuns a 2 gens Hox contigus o l'existència d'*enhancers* globals, semblen ser els responsables del manteniment d'un *cluster*. És a dir, cal mantenir la disposició en *cluster* dels gens Hox per a la seva correcta regulació (Duboule, 1998). En canvi, en *Drosophila*

no és necessària la disposició en *cluster* dels gens homeòtics per a la seva correcta expressió. Mutants naturals amb translocacions de gens homeòtics fora del *cluster* presenten una expressió normal dels gens translocats i, en diferents espècies de *Drosophila*, el *cluster* està trencat per diferents llocs (Mann, 1997). Per tant, no sembla haver cap motiu evident per a que els gens homeòtics de la mosca es mantinguin agrupats al genoma, fins i tot s'ha dit que el seu *cluster* està en "vies de desintegració".

Un altre fenomen que pot forçar la disposició en *clusters* dels gens Hox és la colinearitat temporal, que fa referència a l'expressió ordenada en el temps dels membres del complex. En alguns organismes tals com vertebrats (Simeone i col., 1990; Izpisua-Belmonte i col., 1991), insectes de banda germinal curta (Akam, 1994), sangoneres (Shankland, 1994) i amfiox (Wada i col., 1999), els gens Hox situats a la part 3' del complex a més d'expressar-se en territoris més anteriors de l'embrió ho fan abans en el desenvolupament que els gens més 5'. Per al manteniment de colinearitat temporal és primordial l'existència del *cluster*. Així, la transposició d'un transgèn amb les regions *cis* reguladores del gen *Hoxd-9* a una regió més 5' del *cluster* provoca un retard en la seva activació (van der Hoeven i col., 1996). Aquest retard va ser interpretat com a una manca d'accessibilitat causada, per exemple, per la configuració de la cromatina que inhibia les correctes interaccions reguladores. La validesa d'aquesta hipòtesi es va examinar mitjançant la transposició del mateix transgèn a diferents regions cada vegada més cap a 5' del *cluster* HoxD, fins trobar una localització en la qual aquest retard de l'expressió desapareixés. Un cop delimitada aquesta regió, es va delectonar i es va observar l'activació prematura dels gens del *cluster*. Així, aquest element, era el responsable del silenciament primerenc dels gens situats a la part més 5' del *cluster*, potser mitjançant la unió de proteïnes responsables del tancament de la cromatina (Kondo i Duboule, 1999).

Com l'expressió dels gens Hox en vertebrats està lligada a zones amb alta proliferació (a la part més posterior de la línia primitiva durant la gastrulació primer i al *tail bud* més tard, als esbossos de les extremitats, etc.) s'ha establert un lligament entre un fenomen recurrent com és la divisió cel·lular, amb un de lineal com és la morfogènesi: cèl·lules amb un alt índex mitòtic poden "obrir" els clusters Hox, i per tant activar gens Hox més posteriors progressivament en estructures més posteriors (Duboule, 1994). Així, aquesta propietat (la colinearitat temporal) és la que explicaria i regiria en últim terme el fenomen de la colinearitat espacial.

5.2.4. Gens Hox i evolució morfològica

El fet de que s'hagin caracteritzat gens Hox en la majoria de metazous, junt amb el fet de que la seva funció sigui la de diversificar estructures al llarg de l'eix antero-posterior en la majoria de filums on s'ha pogut determinar, va fer pensar que la comparació d'aquests gens entre diferents espècies podria servir d'indicador d'evolució morfològica. Així, el posterior estudi de l'expressió i/o funció d'aquests gens en diferents organismes ha fet possible aquesta comparació. Des d'aquesta aproximació teòrica inicial fins ara, s'han documentat casos en que tant la duplicació i divergència de gens Hox, com canvis en la seva regulació, canvis en la regulació dels seus gens diana, o més recentment, canvis en la seva seqüència codificant, han estat implicats en l'evolució de la diversitat dels animals. Així, l'estudi comparatiu dels patrons d'expressió dels gens Hox ha servit com a eina per a definir connexions plausibles i reals entre gens del desenvolupament i canvis al llarg de l'evolució. Com a norma general, a canvis més subtils són

atribuïbles diferències morfològiques entre espècies més properes, mentre que a canvis més dràstics són atribuïbles diferències entre espècies evolutivament més distants (Fig. 19).

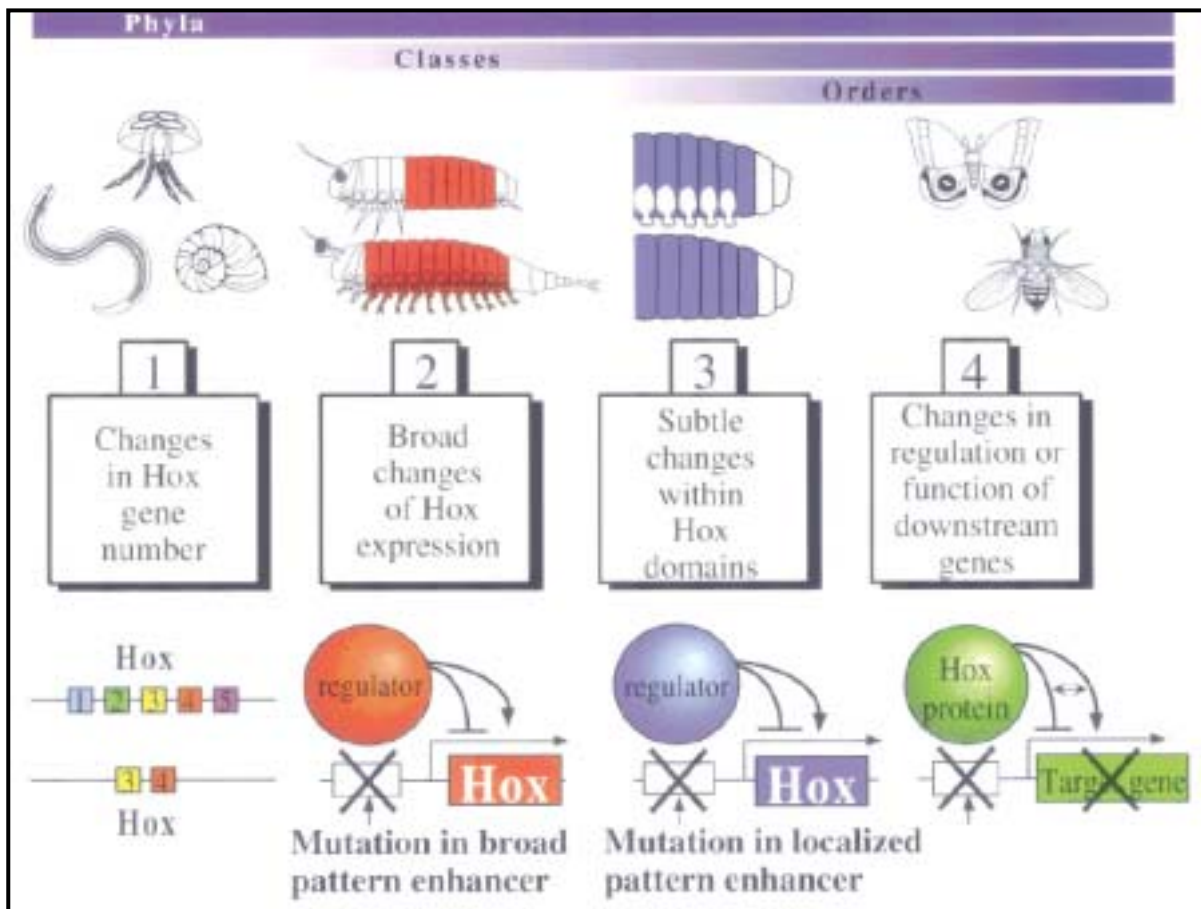


Figura 19. Canvis moleculars a nivell dels gens Hox que han pogut repercutir en variació morfològica al llarg de l'evolució dels animals. Canvis subtils (nivell 4) poden explicar diferències entre espècies més relacionades que canvis més dràstics (nivell 1). A partir de Gellon i McGinnis, 1998.

- Canvis en el nombre de gens

Com ja s'ha esmentat a la secció 3, la duplicació gènica seguida de divergència funcional té un paper cabdal en l'adquisició de noves funcions. Per tant, la duplicació dels *clusters* Hox en el llinatge dels vertebrats ha pogut afavorir l'adquisició de noves funcions per part dels nous gens duplicats, sense comprometre la funció ancestral. L'estudi de l'expressió dels gens del *cluster* Hox d'amfiox i d'ascídies, els altres dos subfílums que junt amb els vertebrats representen el fílum dels cordats, ha posat de manifest l'existència de llocs d'expressió únics a vertebrats. En vertebrats els gens Hox s'expressen de manera colinear en molts territoris com el sistema nerviós, derivats de les crestes neurals, el mesoderm somític i el de la placa lateral, els esbossos de les extremitats, etc. (Deschamps i col., 1999). En canvi en prevertebrats ho fan en un menor nombre de dominis. Concretament, en l'amfiox, els gens Hox més anteriors (amb l'excepció d'*AmphiHox2*) s'expressen de manera colinear en el sistema nerviós central (SNC) i no colinear en la part més posterior del mesoderm (Wada i col., 1999) igual que ho fan els gens Hox d'ascídies (Katsuyama i col., 1995; Gionti i col., 1998; Locascio i col., 1999). D'aquestes observacions podem concloure que la funció ancestral que aquests gens duïen a terme en l'ancestre comú dels cordats estava restringida al desenvolupament del SNC. La duplicació i posterior *de novo* utilització d'aquests duplicats en territoris que són únics als vertebrats com les crestes neurals o

les extremitats pot haver estat instrumental en l'aparició d'aquestes estructures. Cal destacar que hi ha dos menes d'adquisició de noves funcions: aquelles que han adquirit tots els gens Hox com, per exemple, en la morfogènesi dels derivats de les crestes neurals, i aquelles específiques de determinats gens com, per exemple, les que els gens del cluster HoxD tenen en la determinació del patró de l'eix proximo-distal de les extremitats. Les primeres queden rescatades per altres paràlegs en ratolins *knockout*, mentre que les segones, les específiques, queden al descobert. En general, sembla haver una relació directa entre el nombre de clusters Hox i la complexitat morfològica dins del fílum dels cordats: organismes més senzills tenen un menor nombre de clusters Hox (Carroll, 1995).

Malgrat això, l'existència de set *clusters* Hox en el peix zebra (Amores i col., 1998), sembla indicar que no sempre l'augment en complexitat gènica comporta un augment en complexitat morfològica. Aquest fet podria ser reconciliat amb observacions recents en les que es comprova que els gens *HoxB5a* i *HoxB5b* del peix zebra han patit un fenomen de subfuncionalització. Bruce i col. (2001) observen que aquests dos gens s'expressen en territoris específics, la suma dels quals representa el domini d'expressió del gen ortòleg de ratolí *Hoxb-5*. Mitjançant la sobre-expressió de cadascun dels gens paràlegs de peix zebra també demostren que la funció bioquímica dels dos és la mateixa. Amb això, argumenten que la manca de relació entre la complexitat morfològica i la complexitat gènica entre el peix zebra i els mamífers podria ser deguda a fenòmens de subfuncionalització en molts dels gens duplicats del peix que, malgrat forcen la seva conservació, no poden ser causa de cap innovació. Una altra explicació podria ser que aquests gens extra no estessin lligats a un augment de complexitat morfològica sinó que hagin estat instrumentals per a la radiació tan exitosa que han sofert els teleostis. Si aquests gens duplicats es perden de manera divergent (poblacions aïllades perden i fixen diferents gens) es va creant una barrera reproductiva que pot portar a l'especiació (Lynch i Conery, 2000).

El que encara està en discussió és en quin moment es va donar aquesta duplicació del *cluster* (de quatre a vuit) explicada, de manera més parsimoniosa, com a fenomen de poliploidia. Una fita important a aquesta controvèrsia la representen dos treballs: la revaloració del nombre de *clusters* A de fugu (que passa de tenir un a dos) (Amores i col., 1998), i la determinació del nombre de clusters (set) de medaka (Naruse i col., 2000). Així, la visió actual és que un fenomen de poliploidització va tenir lloc molt aviat després de la divergència dels peixos *ray-finned* i els peixos *lobe-finned* ja que tots els sarcopterigis examinats fins ara en tenen quatre (Taylor i col., 2001). Abans però, es pensava que aquest fenomen podia haver estat fins i tot específic del llinatge de peix zebra (Stellwag, 1999).

- Canvis extensos en els dominis d'expressió dels gens

Encara dins el fílum dels cordats, però ara fent referència a variabilitat morfològica entre diferents subclasses del mateix subfílum, trobem un exemple on el canvi en dels dominis d'expressió dels gens Hox es correlaciona amb una clara distinció morfològica. L'esquelet axial d'una serp, amb unes tres-centes vèrtebres similars, es diferencia clarament del de les aus en les que trobem, apart d'un menor nombre, tres tipus diferents de vèrtebres en quant a la seva morfologia. L'estudi comparatiu de la distribució de tres proteïnes Hox (*HoxB5*, *C6* i *C8*) va fer palesa la seva relació amb aquest canvi d'identitat del "tipus vertebral", doncs la seva distribució era diferent durant el desenvolupament del pollet i de la serp. En concret, l'expansió anterior dels dominis d'expressió de *HoxC8* i *HoxC6* al mesoderm somític i al lateral en la serp, es

correlaciona amb l'expansió de la identitat toràctica de les vèrtebres a l'esquelet axial (Fig. 20) (Cohn i Tickle, 1999). També s'ha documentat una variació equivalent a nivell dels dominis d'expressió dels gens Hox als somites que permet explicar el perquè del diferent nombre de vèrtebres de cada tipus en aus i mamífers (Burke i col., 1995).

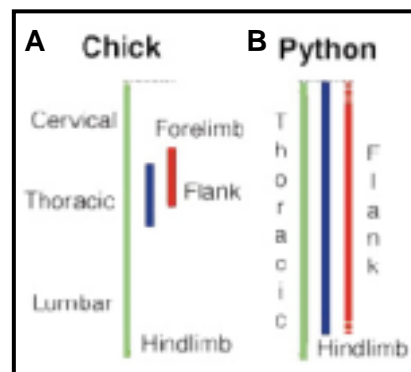


Figura 20. Esquema comparatiu de la distribució de les proteïnes HOXB5 (verd), HOXC6 (vermell) i HOXC8 (blau) en el mesoderm de l'embrió de pollet (A) i de serp (B). A partit de Cohn i Tickle, 1999.

- Canvis subtils en els dominis d'expressió dels gens

En insectes, les proteïnes Hox Ultrabithorax (Ubx) i Abdominal-A (Abd-A) tenen una gran capacitat repressora de l'aparició de potes a l'abdomen, per inhibició de l'expressió del gen *Distal-less (Dll)*. Contradictòriament, els pròlegs de la larva de papallones apareixen als segments abdominals en paral·lel a l'expressió dels dos gens abans esmentats. Un estudi més detallat va posar de manifest una petita diferència que reconciliava l'expressió d'*Ubx* i *Abd-A* amb la manca de repressió de formació dels pròlegs en papallones: es va observar que encara que ambdues proteïnes eren presents a l'abdomen en estadis primerencs, més tard quedaven absents de les poques cèl·lules que donaran lloc a aquests apèndix, caracteritzades per l'expressió del gen *Dll* (Warren i col., 1994). Així, la pèrdua d'expressió dels gens *Ubx* i *Abd-A* en un petit territori, és la responsable d'aquesta diferència morfològica (apèndix/no apèndix a l'abdomen) entre grups diferents d'insectes.

- Canvis en la regió codificant dels gens

Un altre nivell de modificació seria aquell que introdueix canvis en la regió codificant del gen fent que es produeixi una proteïna que manifesti una nova funció que sigui responsable d'una determinada característica d'un grup d'animals. Malgrat aquesta aproximació teòrica, no ha estat fins fa ben poc que s'ha documentat un cas en el que aquesta variació a nivell codificant no hagi compromès la funció Hox "canònica". Dos treballs complementaris intenten explicar com el grup d'insectes, amb sis potes, ha pogut divergir dels seus ancestres amb múltiples potes. Treballs fets amb *Drosophila* havien posat de manifest la gran capacitat repressora que té la proteïna Ubx sobre la formació de potes. Llavors, la manca d'expressió del gen *Ubx* als segments toràctics de la mosca permet la formació dels seus tres parells de potes, mentre que la presència d'Ubx als segments abdominals ho reprimeix. En canvi, les proteïnes Ubx d'altres artròpodes no manifesten aquest efecte repressor i, de fet, s'expressen a segments on es formaran potes. Galant i Carroll (2002) demostren que un domini a la part C-terminal de la proteïna Ubx és el responsable de la potent activitat repressora de la formació de potes que s'ha documentat en insectes. Aquest domini és present a insectes i no ho és a una espècie d'onicòfor. En la mateixa línia, però

aquesta vegada marcant diferències dins del mateix fílum artròpodes, Ronshaugen i col. (2002) demostren que la proteïna Ubx d'*Artemia* conté, en la mateixa regió C-terminal, un domini que inhibeix la funció repressora de la formació de potes que té Ubx en insectes.

- Canvis de regulació sobre els gens diana

Fa força temps que s'intueix que canvis en els gens diana de determinats gens Hox han pogut ser instrumentals en la producció de diferents morfologies, però no ha estat fins fa poc que les eines moleculars adequades han permès la seva demostració experimental. Es suggeria, com a mecanisme responsable de l'aparició diferencial d'un halteri al tercer segment toràctic (T3) d'una mosca, respecte l'aparició d'una ala en el segment homòleg de papallones, la regulació de diferents gens diana pel mateix gen Hox (*Ubx*). L'expressió d'*Ubx* a T3 inhibeix l'aparició d'una ala, mentre que la manca d'*Ubx* al segon segment toràctic (T2) de *Drosophila* permet la seva aparició (Carroll, 1995). L'estudi de l'expressió diferencial de gens en disc imaginal d'ala respecte disc imaginal d'halteri (i per tant putatiu gens diana d'*Ubx*) en *Drosophila*, va permetre a Weatherbee i col. (1998) la identificació de gens implicats en la via de diferenciació de l'ala, l'expressió dels quals estava directa o indirectament reprimida per Ubx. Posteriorment, es va estudiar l'expressió d'aquests gens que són reprimits per Ubx a disc imaginal d'halteri de *Drosophila*, en papallones. Això va permetre comprovar que molts d'aquests gens no estan inhibits per Ubx en les ales del T3 de papallones (Weatherbee i col., 1999). Així, aquest canvi a nivell de funció de Ubx ens explica diferències morfològiques entre espècies de dos ordres diferents.

5.2.5. Gènes i evolució dels clusters Hox

Després de l'aïllament del *cluster* Hox de *Drosophila* i els quatre de mamífers ja fa més de

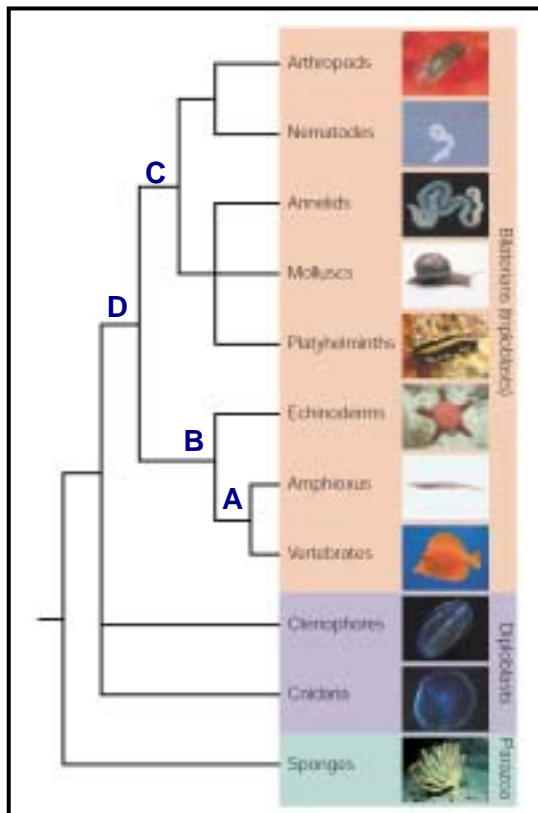


Figura 21. Arbre filogenètic dels metazous basat en DNA ribosomal i seqüències de gens Hox. A partir de Ferrier i Holland, 2001.

deu anys, s'han aïllat gens Hox o el(s) *cluster(s)* Hox d'altres organismes. Així, per comparacions del contingut gènic de cadascun podem hipotetitzar com hauria d'haver estat el *cluster* de l'organisme que ocupava un determinat node a l'escala evolutiva com, per exemple l'ancestre dels cordats (A en Fig. 21), dels deuteròstoms (B en Fig. 21), dels protòstoms (C en Fig. 21) i dels bilateris (D en Fig. 21). Quan és possible, és preferible la caracterització del *cluster* i no tan sols la dels gens doncs, moltes vegades, la simple comparació d'un gen Hox amb d'altres ja aïllats no resol la seva relació filogenètica; la posició relativa que aquest gen té dins el *cluster* també serveix com a caràcter filogenètic per establir la seva relació amb altres gens Hox. A mida que s'han anat aïllant gens Hox en nous organismes i a mida que treballs de filogènia molecular han refinat o fins i tot canviat la posició d'aquests organismes en l'arbre evolutiu, el contingut teòric en gens Hox d'aquests

clusters ancestrals ha anat canviant. Molt possiblement, seguirà canviant en els pròxims anys donat els enormes esforços que s'estan invertint en el clonatge dels gens Hox en organismes clau en l'escala evolutiva.

Dues diferències clares entre els *clusters* de vertebrats respecte els *clusters* dels invertebrats aïllats inicialment (*Drosophila* i *C. elegans*) feien referència a l'existència de múltiples *clusters* i múltiples gens del tipus posterior en vertebrats. Per resoldre com havia de ser el(s) *cluster(s)* de l'ancestre dels cordats (node A a Fig. 21) van ser d'importància vital la caracterització del nombre de *clusters* i de gens posteriors de diferents grups de prevertebrats i de vertebrats primitius. Així, la presència de (i) múltiples representants d'un mateix PG en lamprees (Sharman i Holland, 1998), (ii) un únic *cluster* en l'amfiox amb un mínim de dos gens posteriors (Garcia-Fernández i Holland, 1994) i, (iii) un mínim de tres gens Hox del tipus posterior en ascidis (Di Gregorio i col., 1995) van posar de manifest que el seu ancestre havia de tenir un únic *cluster* Hox amb múltiples gens del tipus posterior. Així, la duplicació dels *clusters* havia tingut lloc molt aviat en l'evolució dels vertebrats i la duplicació de la classe Abd-B ja era present en els cordats. Degut a la divergència aparent del *cluster* Hox d'ascidis, se suposa que el *cluster* de l'ancestre dels cordats havia de ser molt similar al dels amfiox.

Pel que fa referència a l'hipotètic *cluster* (únic) de l'ancestre dels deuteròstoms (node B en Fig. 21) han estat claus l'aïllament del complex d'equinoderms i la parcial determinació del nombre de gens d'una espècie d'hemicordats. La cerca de gens Hox per PCR mitjançant oligonucleòtids degenerats putativament capaços d'amplificar gens Hox del PG1 al 10 en una espècie d'enteropneusts (un grup d'hemicordats), va resultar en l'aïllament de nou fragments. Entre ells hi havia representants de tots tres grups (anteriors, centrals i posteriors), posant en evidència que a l'ancestre dels deuteròstoms ja havia tingut lloc la duplicació dels gens Hox posteriors (Pendelton i col., 1993). La cerca de gens Hox en diferents espècies d'equinoderms (a destacar Popodi i col., 1996; Martínez i col., 1999), va acabar-ho d'evidenciar: diferents espècies d'equinoderms, ja contenen múltiples gens Hox posteriors. La suma de totes les dades que es tenen d'aquest fílum recolzen la hipòtesi de que l'ancestre dels equinoderms tenia un únic *cluster* amb un mínim d'onze gens (tres anteriors, quatre medials i quatre posteriors), malgrat no es coneixen tots els ortòlegs de totes les classes d'equinoderms (Long i Byrne, 2001). Cal destacar que s'observa una clara relació d'ortologia entre els gens anteriors i medials d'equinoderms als corresponents PGs de vertebrats, mentre que aquesta relació no és clara pels gens Hox de tipus posterior. També cal destacar la duplicació específica d'un gen posterior en una espècie d'ofiuroid (Mito i Endo, 2000), la qual cosa fa que tingui sis gens posteriors mentre que els vertebrats només en tenen cinc. Així, el *cluster* de l'ancestre dels deuteròstoms havia de tenir un aspecte molt similar al de l'ancestre dels cordats i, per tant, molt similar al de l'amfiox.

Per tal d'inferir com era el complex de l'ancestre dels protòstoms (node C en Fig. 21) s'ha de comparar la composició en gens Hox de *Lophotrochozoa* i *Ecdysozoa*, que són les dues divisions incloses dins del grup dels protòstoms (Aguinaldo i col., 1997). De Rosa i col. (1999) recullen les dades que es tenen fins al moment i n'afegeixen de noves de dues espècies de *Lophotrochozoa* i d'una d'*Ecdysozoa*. Així, s'evidencia l'existència de gens centrals i posteriors específics de cadascun dels dos grups. Els *Lophotrochozoa* estan caracteritzats pels gens *Lox5*, *Lox2* i *Lox4* (com a centrals específics) i per *Post1* i *Post2* (com a posteriors específics) (Fig. 22). Els *Ecdysozoa* es caracteritzen per *Antp* i *Ubx/Abd-A (Ubd-A)* (centrals) i per *Abd-B* (com a

posterior) (Fig. 22). Basant-se en el nombre compartit de gens de cadascun dels grups hipotetitzen que el seu ancestre comú havia de tenir un cluster Hox amb un mínim de set gens.

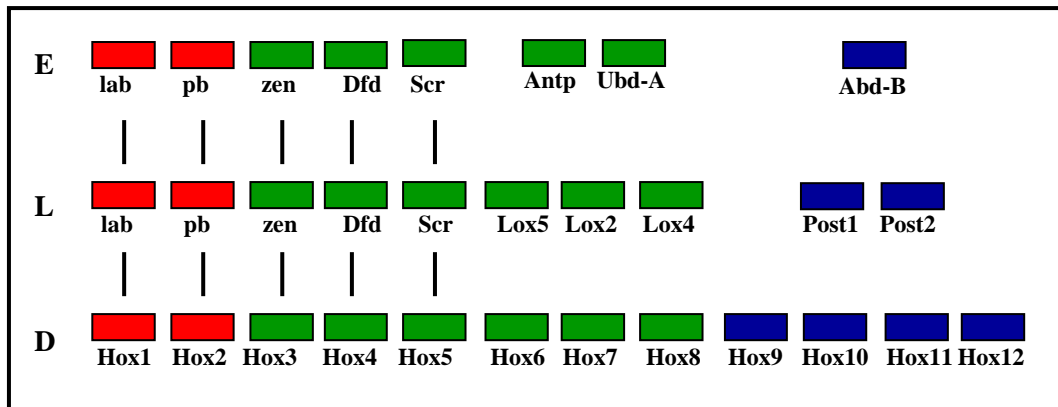


Figura 22. Nombre mínim de gens Hox a l'ancestre dels *Ecdysozoa* (E) dels *Lophotrochozoa* (L) i dels deuteròstoms (D). Els gens anteriors s'indiquen en vermell, els centrals en verd i els posteriors en blau. Les línies verticals marquen clares relacions d'ortologia.

Per a inferir el nombre de gens que havia de tenir l'ancestre dels organismes bilateris (node D en Fig. 21) cal comparar la composició en gens dels *clusters* de deuteròstoms i protòstoms. Entre aquests grups, només es poden establir relacions clares d'ortologia entre els cinc primers PGs, és a dir, *lab*/PG1, *pb*/PG2, *zen*/PG3, *Dfd*/PG4 i *Scr*/PG5; per tant, el nombre del que partim ja és cinc. Si suposem que no es poden establir relacions d'ortologia clares entre la resta de gens centrals i posteriors perquè provenen de duplicacions independents, a aquest número se li ha d'afegir un mínim d'un grup central més i un grup posterior (Kourakis i Martindale, 2000). Però si suposem que no podem establir aquestes relacions degut a fenòmens de major divergència d'aquests gens, l'ancestre comú als bilateris podria tenir un nombre mínim de deu gens Hox (als cinc clars ortòlegs hem d'afegir un mínim de tres gens centrals i dos gens posteriors) (de Rosa i col., 1999).

5.3. La classe Even-skipped (Evx)

La classe Evx està inclosa al grup *extended Hox* de la superclasse Antp de gens amb homeobox, doncs els gens Evx estan lligats al *cluster* Hox en algunes espècies com al corall *Acropora formosa* (Miller i Miles, 1993), a una espècie de tauró (Kim i col., 2000), a peix zebra (Sordino i col., 1996; Amores i col., 1998), a ratolí (Dush i Martin, 1992) i a humans (D'Esposito i col., 1991; Faiella i col., 1992). A més, filogenèticament són molt propers; concretament, els gens Evx es troben en una posició basal als gens Hox i ParaHox en arbres filogenètics (Gauchat i col., 2000). Tant el lligament dels gens Evx als *clusters* Hox com la seva posició basal a gens Hox i ParaHox (veieu Fig. 13) han suggerit a alguns autors que Evx podria representar el grup germà del gen *UrProtoHox* (Kourakis i Martindale, 2000) és a dir, el gen que per duplicació en tàndem dona lloc al cluster ProtoHox amb 4 gens i la duplicació d'aquest dona lloc als clusters Hox i ParaHox (Brooke i col., 1998). Si això és cert, es podria suposar l'existència d'un gen molt similar a Evx del tipus ParaHox (veieu discussió).

Hi ha tota una sèrie de característiques comunes als membres d'aquesta classe i que no són presents en gens amb homeobox d'altres classes, com ara:

- (i) la presència d'un intró (excepte en el cas de *Drosophila*) entre els residus 46 i 47 del domini homeòtic. Altres gens amb homeobox que tenen un intró dins l'homeobox, el solen tenir entre la posició 44 i 45 (Bürglin, 1995).
- (ii) en general, el domini homeòtic es situa a la part més carboxi-terminal (C-terminal) de la proteïna, mentre que en el cas de les proteïnes EVX aquest està situat a la part més amino-terminal (N-terminal).
- (iii) a diferència d'altres proteïnes amb domini homeòtic en les que l'homologia de seqüència està bàsicament restringida a l'homeodomini, en les proteïnes EVX hi ha altres regions molt similars, com l'extrem C-terminal, ric en amino àcids prolina, alanina i glutamina. Aquest domini és necessari i suficient per a la repressió dels gens diana de la proteïna EVX (Han i Manley, 1993; Briata i col., 1997).

5.3.1. Els gens *Evx* en invertebrats

El primer gen clonat de la classe *Evx* -que dona nom a la classe- fou el gen *even-skipped* (*eve*) de *Drosophila* (Macdonald i col., 1986) per hibridació creuada amb una sonda de l'homeobox de gens homeòtics. El seu fenotip va ser descrit al 1980 per Nüsslein-Volhard i Wieschaus i classificat com a gen de segmentació del tipus *pair-rule*, degut a que produïa deleccions dels segments imparells. L'estudi de la distribució del seu mRNA (Macdonald i col., 1986) i la localització de la proteïna EVE (Frasch i col., 1987) van fer palès el dinàmic i complex patró d'expressió d'aquest gen. El gen *eve* mostra una expressió transitòria en forma de set bandes (Fig. 23) a l'estadi de blastoderm cel·lular (que posteriorment esdevindran els parasegments imparells de l'embrió) que es resolten durant la gastrulació en un patró de 14 bandes. És en aquest "estadi" de 14 bandes en el que l'expressió d'*eve* i del gen de polaritat segmental *engrailed* (*en*) coincideixen en l'espai, marcant el compartiment anterior de cadascun dels parasegments. Aquests dos gens estan íntimament relacionats, ja que en absència de la funció EVE, el patró en 14 bandes del gen *en* es perd en els segments medials. Aquestes bandes d'expressió d'*eve* es perden un cop acabada la gastrulació amb l'inici de l'extensió de la banda germinal. És en aquest moment en que es detecta un nou lloc d'expressió a la part posterior de l'embrió, al primordi proctodeal, que és l'únic lloc on *eve* és actiu un cop la banda germinal està completament extesa. En començar la segregació dels neuroblasts en aquest estadi, l'expressió d'*eve* es fa palesa a petits conjunts de cèl·lules mare ganglionars de cada parasegment. L'estudi de mutants termosensibles ha revelat que la seva funció és necessària per a la correcta morfologia axonal de dos tipus particulars de neurones (aCC i RP2) (Doe i col., 1988).

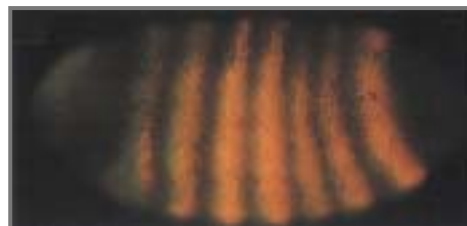


Figura 23. Doble immunofluorescència de la proteïna EVE (vermell) i Hairy (verd) en el blastoderm cel·lular de *Drosophila melanogaster*. Extret de Pankratz i Jäckle, 1993.

El clonatge i estudi de l'expressió d'ortòlegs del gen *eve* en altres insectes ha permès establir quines d'aquestes funcions són comunes a tots i quines són úniques a *Drosophila*. Primer, es va clonar l'ortòleg del gen *eve* en un insecte més primitiu -*Schistocerca*- el qual, a diferència de la mosca, té un desenvolupament del tipus de banda germinal curta. Aquest gen es comença a expressar a l'inici de la gastrulació, en un domini posterior de l'embrió. A mida que l'embrió s'elonga, l'expressió es manté a la regió més posterior del mesoderm. Pel que fa referència al

sistema nerviós central (SNC), el gen *eve* d'ambdós insectes s'expressen en patrons indistingibles als mateixos tipus neuronals (Patel i col., 1992).

Patel i col. (1994) van anar més enllà en l'estudi de l'evolució de la funció que aquests gens havien tingut en insectes per tal d'establir l'evolució de la formació del patró dels insectes i veure si la classificació d'aquests -segons el seu tipus de banda germinal- podia predir la mecànica del seu desenvolupament. Amb aquest objectiu, estudien la distribució de la proteïna EVE en tres espècies de coleòpters (situats evolutivament entre *Schistocerca* i *Drosophila*); un de banda germinal curta (*Tribolium*), un d'intermèdia (*Dermestes*) i un de llarga (*Callosobruchus*) amb un anticòs monoclonal generat contra la proteïna EVE d'*Schistocerca*. En tots tres, la proteïna EVE presenta un patró tipus *pair-rule*. La única diferència és el nombre relatiu de bandes que es formen abans respecte després de la gastrulació, encara que la resolució de 7 bandes en 14 no es fa exactament de la mateixa forma. Per tant, la designació en tipus de banda germinal no es correlaciona amb un determinat mecanisme de formació del patró. En el mateix estudi i en un de posterior de l'expressió del gene *eve* de *Tribolium* (*Tc eve*) (Brown i col., 1997) també es troba l'associació que aquests gens tenen a dominis posteriors i amb el SNC.

S'ha clonat un representant de la classe *Evx* al corall *A. formosa* anomenat *eveC* (Miller i Miles, 1993), que està lligat a un gen tipus *Antp* o *Hox* central anomenat *antpC*. Malauradament, no se'n coneix l'expressió.

També s'ha caracteritzat l'únic membre de la classe *Evx* del nemàtode *C. elegans* (*vab-7*), la funció del qual és necessària pel correcte desenvolupament d'estructures posteriors i per a una funció del SNC relacionada amb la locomoció. S'expressa per primera vegada a 4 cèl·lules posteriors a l'estadi de 100 cèl·lules, que donaran lloc als músculs de la paret dorsal-posterior, a l'epidermis i a dues neurones (Ahringer, 1996).

5.3.2. Els gens *Evx* en vertebrats

S'han clonat dos gens de la classe *Evx* en humans (D'Esposito i col., 1991; Faiella i col., 1991) i ratolí (Bastian i Gruss, 1990), tres en peix zebra (Joly i col., 1993; Thäeron i col., 2000; van der Hoeven i col., 1996) i un en *Xenopus* (Ruiz i Altaba i Melton, 1989).

L'únic gen *Evx* de *Xenopus*, *Xhox3*, presenta dues etapes d'expressió. Durant l'estadi de gàstrula i de nèurula el seu mRNA es troba al mesoderm de manera gradual al llarg de l'eix antero-posterior (essent més abundant a la part posterior de l'embrió) principalment en teixits mesodèrmics i ectodèrmics. Durant l'etapa més tardana, a l'estadi de *tail bud*, es troba concentrat als dos pols de l'embrió; a la part anterior del SNC [específicament a la frontera entre el cervell mitjà i el posterior (MHB)] i al *tail bud*. L'anteriorització dels embrions causa una disminució de l'expressió del gen, mentre que la manipulació contrària, la posteriorització, provoca un augment del seu mRNA, implicant així una altra vegada a la classe *Evx* amb una funció d'especificació de la part més posterior de l'embrió (Ruiz i Altaba i Melton, 1989).

En ratolí i en humans els dos gens *Evx* (*Evx1* i *Evx2*) es troben lligats als *clusters* *HoxA* i *HoxD* respectivament. El primer està a una distància de 45 Kb del seu gen *Hox* veí, *Hoxa13*, mentre que només 8 Kb separen a *Evx2* del gen *Hoxd13* (Dush i Martin, 1992; D'Esposito i col., 1991; Faiella i col., 1992). El gen *Evx1* es comença a expressar just després de la implantació de l'embrió a l'úter (5 dies *post coitum* [dpc]) a la capa d'endoderm visceral que envolta l'embrió en

forma de cilindre (*egg cylinder*). D'acord amb aquest patró d'expressió primerenc, ratolins homozigots nuls per al gen *Evx1* moren just després de la seva unió a la membrana de l'úter sense mostrar cap mena de creixement cel·lular, diferenciació o morfogènesi (Spyropoulos i Capecchi, 1994). Just abans i durant la gastrulació, les cèl·lules que expressen el gen -que són aquelles properes o dins la línia primitiva- ho fan amb un patró gradual (hi ha un decreixement dels nivells de missatge des de la part posterior a la part anterior). Aquestes cèl·lules són les que donaran lloc al mesoderm i ectoderm de la part posterior de l'embrió. A 10 dpc, quan el *tail bud* substitueix la línia primitiva com a font de nou mesoderm, *Evx1* s'expressa en aquest territori més posterior de l'embrió (Dush i Martin, 1992). Més endavant al desenvolupament, trobem l'expressió típica dels gens *Evx* al SNC: concretament, en determinades interneurons que estan en el procés de generació dels seus axons al llarg del tub neural i més rostralment amb límit anterior al MHB (Bastian i Gruss, 1990). També el trobem expressat de manera transitòria als esbossos de les extremitats (Dollé i col., 1994).

El gen *Evx2* de ratolí, en canvi, presenta un patró d'expressió que no està relacionat amb altres gens de la seva classe, sinó que és més semblant al dels gens *Hox* posteriors. Així, *Evx2* respon a la colinearitat espacial observada pels gens *Hox* i s'expressa en estadis més tardans que el gen *Hox* veí, *Hoxd13*. Un cop expressat -9,5 dpc- ho fa en àrees de l'epiteli de la cloaca i al mesoderm que l'envolta (del que s'originarà l'esbós del genital). Un dia més tard es detecta la seva expressió als esbossos de les extremitats (en mesènquima comprès dins el territori d'expressió del gen *Hoxd13*) i als genitals (Fig. 24A). Aquests dos llocs d'expressió, cloaca i esbossos de les extremitats, són característics dels gens posteriors dels clusters *HoxA* i *HoxD*, però no dels gens *Evx*. Així, homozigots nuls per al gen *Evx2* mostren defectes, encara que subtils, en l'extremitat (Hérault i col., 1996). Més tard però, sí que trobem un tret característic dels gens de la classe *Evx* en el patró d'expressió d'*Evx2*: la seva presència al SNC (Fig. 24B). En

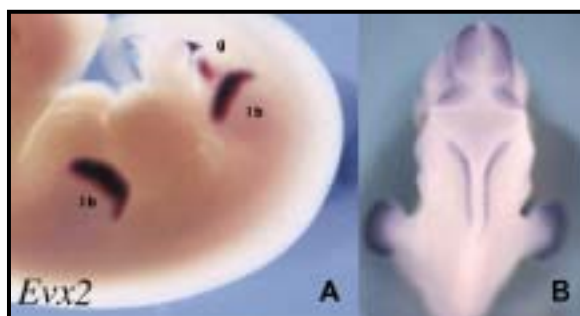


Figura24. A) Expressió de *Evx2* en l'embrió de 10,5 dpc de ratolí. Se senyala l'expressió als esbossos de les extremitats (lb) i a l'esbós del genital (g). B) Vista dorsal d'un embrió de 12 dies de ratolí en el que s'observa la seva expressió al SNC). Cortesia de D. Duboule i M. Kmita.

aquest territori, la única diferència entre els dos paràlegs de ratolí és el seu límit anterior d'expressió: *Evx2* s'expressa en territoris més rostrals a la MHB (Dollé i col., 1994). Els mateixos autors proposen una hipòtesi que permet explicar el perquè d'aquest patró tan diferent d'aquest gen respecte altres membres de la seva classe: degut a la seva proximitat als gens *Hox* posteriors del *cluster* *HoxD*, aquest gen entra dins la seva regulació global. El seu patró d'expressió més primerenc durant la gastrulació queda silenciada per la repressió global que els gens *Hox* més posteriors tenen en estadis primerencs del desenvolupament i, després, es passen a expressar a territoris típics de gens *Hox* degut a l'acció d'enhancers únics per a tots aquests gens *Hox* més posteriors que provoquen la seva expressió en aquests territoris. L'expressió al SNC, més tardana, ja ve dictada per regions *cis* reguladores pròpies del gen *Evx2* i té lloc quan la part posterior del complex ja és accessible a la maquinària de transcripció. L'element responsable del

silenciament de la part més 5' (posterior) del cluster HoxD ha estat mapat entre 16 i 28 Kb a 5' del gen *Evx2* (Kondo i Duboule, 1999). Així, la hipòtesi anterior pren forma, doncs el gen *Evx1*, a 45 Kb de *Hoxa13*, escapa d'aquest element "silenciador", mentre que *Evx2*, a només 8 Kb de *Hoxd13*, està dins la zona de cromatina no accessible a la maquinària transcripcional en etapes primerenques del desenvolupament.

El primer gen de la classe *Evx* aïllat a peix zebra va ser el gen *eve1*, que s'expressa a un domini ventro-lateral durant la gastrulació, a la part més posterior de l'embrió durant la somitogènesi i al *tail bud* durant l'elongació. A diferència d'altres gens *Evx*, no es troba al SNC (Joly i col., 1993). Un altre dels paràlegs, *evx1*, manifesta una expressió dinàmica durant la neurogènesi, en que s'expressa en determinats conjunts d'interneurones just abans de que emetin els seus axons, amb un límit anterior al MHB. També s'expressa durant la formació de l'intestí posterior i el sistema urogenital (Thäeron i col., 2000) i durant el desenvolupament i regeneració de les aletes (Borday i col., 2001). El tercer i últim representant de la classe *Evx* al peix zebra, *Evx2*, presenta un patró d'expressió molt similar al del seu ortòleg de ratolí (pel que fa referència a la colinearitat espacial), també degut a la proximitat d'aquest gen al cluster Hoxd. Es diferencien en que l'ortòleg de peix zebra no es detecta a la cloaca i sí es troba expressat al *tail bud*, potser degut a que la distància entre aquest gen i el Hox veí és major (Sordino i col., 1996).

5.3.3. Evolució de la funció dels gens *Evx*

El fet de que s'hagin aïllat representants de la classe *Evx* en un ampli ventall d'organismes i s'hagi determinat el seu patró d'expressió i/o funció, permet proposar un escenari evolutiu de la funció d'aquesta classe, en el que podem distingir entre les funcions comunes que aquests gens duen a terme en tots els organismes fins al moment estudiats i les funcions úniques que aquests gens han adquirit en determinats grups. Aquesta evolució de la funció dels gens *Evx* està resumida en la taula II.

Taula II. Evolució de l'expressió dels gens de la classe *Evx*

organisme	pair-rule	territoris posteriors	SNC	MHB	tail bud	apèndixs
<i>Schistocerca</i>	-	✓	✓	-	-	-
escarbats	✓	✓	✓	-	-	-
<i>Drosophila</i>	✓	*	✓	-	-	-
<i>C. elegans</i>	-	✓	✓	-	-	-
Peix zebra	-	✓	✓	✓	✓	✓
<i>Xenopus</i>	-	✓	✓	✓	✓	??
ratolí	-	✓	✓	✓	✓	✓

En vermell s'indica manca d'expressió (-) i en verd presència (✓). Els dominis d'expressió de vertebrats fan referència a algun dels seus gens paràlegs. El primer rectangle indica la funció derivada que aquests gens han adquirit en alguns insectes, el segon fa referència als trets comuns de l'expressió dels gens *Evx* al llarg de l'escala evolutiva i el tercer a dominis d'expressió únics de vertebrats. Els interrogants indiquen que mitjançant l'estudi realitzat a *Xenopus* no es pot afirmar ni negar si hi ha expressió del gen als apèndixs.

D'aquesta taula resum en podem treure les següents consideracions evolutives:

- (i) El patró *pair-rule* que els gens *Evx* tenen en algun ordre d'insectes és un caràcter derivat adquirit a l'ancestre comú d'escarbats i dípters (Patel i col., 1992, 1994).
- (ii) Els trets comuns d'expressió a tots els gens *Evx* són la seva presència a estructures posteriors de l'embrió [l'asterisc a *Drosophila* indica que trobem aquesta restricció posterior com a expressió al primordi del proctodeu (Goto i col., 1989), o potser simplement ha perdut secundàriament aquest domini] i durant el desenvolupament del SNC.
- (iii) Hi ha funcions úniques dels gens *Evx* a vertebrats, com la que duen a terme a la MHB (potser lligat a les propietats organitzadores que aquesta regió duu a terme en vertebrats), durant l'elongació de l'embrió a partir del *tail bud* [que és una sinapomorfia dels cordats (Gee, 1996)] i durant el desenvolupament de les extremitats [per "segrestament" del paràleg *Evx2* pels elements reguladors dels gens posteriors del cluster *HoxD* (Dollé i col., 1994)].

5.4. La classe Mox

La classe Mox ha estat classificada dins la superclasse Antp de gens amb homeobox degut a l'existència d'un intró conservat entre els residus 44 i 45 de l'homeodomini (Bürglin, 1995). Malgrat això, la posició relativa d'aquesta classe dins la superclasse Antp encara està en discussió. Alguns autors, basant-se en relacions filogenètiques, l'han situat com a molt propera a gens del *cluster Hox* i a gens del *cluster ParaHox* (Balavoine, 1996), on la classe Mox representaria el gen ParaHox central escapat del complexe al llarg de l'evolució. Altres autors la situen estretament relacionada amb els gens del *cluster Hox*, formant amb aquests i els gens *Evx*, el *cluster extended Hox* (Pollard i Holland, 2000), basant-se en dades de lligament al genoma humà i al de ratolí entre els gens Mox i els *clusters Hox*. Per exemple, el gen *Mox1* i el *cluster HoxB* són al cromosoma 17 humà i al cromosoma 11 de ratolí (veieu Fig. 11).

Representants de la classe Mox han estat clonats en un ampli ventall d'organismes, des de diblàstics a vertebrats. En concret, s'han publicat 4 gens Mox d'invertebrats [*Cnox5* del cnidari *Hydra magnipapillata* (Naito i col., 1993), *butonless (btn)* de *Drosophila* (Chiang i col., 1994), *Hrox1* del molusc *Haliothis rufescens* (Degnan i col., 1997), i el gen *GtDutarh2* de planària (Lukianov i col., 1994)]. En vertebrats hi ha dos gens de la classe Mox (*Mox1* i *Mox2*), que han estat clonats en ratolí (Candia i col., 1992), humans (Futreal i col., 1994; Grigoriou i col., 1995), pollet (Rallis i col., 2001), rata (Gorski i col., 1993), *Xenopus* (Candia i Wright, 1995) i peix zebra (Neyt i col., 2000). La característica comú d'aquests gens és la seva associació a derivats mesodèrmics.

5.4.1. Els gens Mox en organismes invertebrats

Durant el desenvolupament de *Drosophila*, el gen *btn* s'expressa a les cèl·lules medials dorsals (de l'anglès *dorsal medial (DM) cells*), que es distribueixen al llarg de la línia medial dorsal de pont mesodèrmic que connecta les dues mitges parts del mesoderm. En el mutant de manca de funció de *btn*, l'inicial determinació de les cèl·lules DM té lloc normalment, però la posterior diferenciació es veu afectada i, per tant, les cèl·lules DM es perden (Chiang i col., 1994).

En el cas del gasteròpode *Haliothis*, s'ha analitzat l'expressió del gen *Hrox1* per RT-PCR quantitatiu. El missatger es detecta per primera vegada a la larva trocòfora que acaba d'eclosionar i es manté constant i en màxim nivell durant la morfogènesi de la larva de trocòfora

a velíger, que és precisament quan està tenint lloc la diferenciació muscular (Degnan i col., 1997). Malauradament, no es tenen dades d'expressió dels gens *Mox* en els altres invertebrats abans esmentats.

5.4.2. Els gens *Mox* en vertebrats

En ratolí, els gens *Mox* estan lligats a diferents processos al llarg del desenvolupament, presentant expressió solapant en alguns i específica en d'altres. En concret, *Mox1* comença a expressar-se durant la gastrulació -7 dpc- principalment al mesoderm paraxial pressumptiu. L'expressió de *Mox2*, en canvi, no es detecta fins a que comença la somitogènesi (8 dpc). Durant aquest procés, els dos gens *Mox* estan coexpressats als somites epitelials i *Mox1*, però no *Mox2*, es detecta també en la part més anterior de mesoderm presomític (PSM), abarcant una longitud equivalent a 2 o 3 somites. Durant la diferenciació inicial dels somites en que perden el seu caràcter epitelial i es subcompartimentalitzen, els dos gens es troben coexpressats a l'escleròtom, mentre que el dermomiòtom tan sols expressa *Mox1* (Candia i col., 1992). Als esbossos de les extremitats també hi ha una expressió diferencial dels gens *Mox*: només *Mox2* s'expressa als mioblasts que es delaminen del dermomiòtom per anar a colonitzar el mesènquima de l'esbós de l'extremitat (Candia i Wright, 1996) (Fig. 25). Així, s'ha hipotetitzat que els gens *Mox* podrien estar involucrats en les primeres etapes d'especificació dels somites, abans de la diferenciació terminal dels seus derivats. Les dades de manca de funció dels gens *Mox* no han pogut, de moment, validar aquesta hipòtesi, ja que els mutants tenen afectat el desenvolupament d'estructures derivades dels somites, però la somitogènesi sembla produir-se normalment. Ratolins homozigots nuls per al gen *Mox2* presenten defectes en la musculatura de les extremitats, caracteritzats per una reducció global de la massa muscular i l'eliminació de determinats músculs (Mankoo i col., 1999), mentre que ratolins homozigots nuls per a *Mox1*, presenten defectes vertebrals (citats a Stamatakis i col., 2001). Segurament, en les primeres etapes de somitogènesi cadascun dels paràlegs pot suplir la funció de l'altre, fenomen típic de gens duplicats amb redundància parcial i, per tant, caldrà esperar a la generació del doble mutant per validar la hipòtesi de la funció dels gens *Mox* en les primeres etapes de la determinació dels somites basada en dades d'expressió.

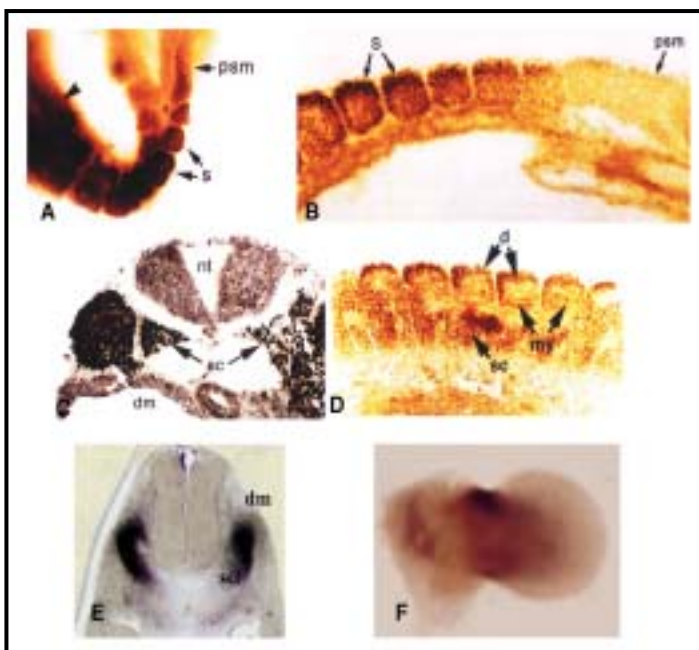


Figura 25. Expressió dels gens *Mox1* i *Mox2* de ratolí. A) Expressió de *Mox1* en els somites ja formats i en el PSM. B) L'expressió de *Mox2* als somites ja formats. C) Detall de l'expressió inicial del gen *Mox1* en tots els compartiments del somita en diferenciació en un tall transversal. D) Tall sagital que mostra l'expressió de *Mox2* en tots els compartiments del somite en diferenciació amb excepció del miòtom. E) Expressió de *Mox2* en els mioblasts que es delaminen del dermomiòtom per anar a colonitzar el mesènquima de l'esbós de l'extremitat. F) Localització del mRNA de *Mox2* a la regió més dorsal i a la més ventral de l'esbós de l'extremitat. d, dermòtom; dm, dermomiòtom, my, miòtom; nt, tub neural; psm, mesoderm presomític; s, somita; sc, escleròtom. Adaptat de Candia i Wright, 1996 i de Rallis i col., 2001.

En pollet, el gen *cMox2* s'expressa de manera similar al seu trans-ortòleg de ratolí durant la somitogènesi i migració dels mioblasts a l'esbós de l'extremitat (Rallis i col., 2001). En *Xenopus* només s'ha trobat un gen *Mox*, *XMox2*, i encara que per seqüència pertany clarament al grup *Mox2* de vertebrats, el seu patró d'expressió és molt més semblant als gens *Mox1*, doncs es comença a expressar durant la gastrulació al mesoderm posterior, i durant la somitogènesi al PSM i als somites (Candia i Wright, 1995).

5.5. La classe *empty spiracles* (*Emx*)

La classe *Emx* pertany a la superclasse *Antp* de gens amb homeobox, ja que comparteix amb algun altre dels seus membres l'existència d'un intró dins l'homeobox entre els residus 44 i 45 de l'homeodomini i d'un hexapèptid conservat a 5' de l'homeodomini (Bürglin, 1995). A més d'aquests dos dominis proteics, 14 amino àcids de la regió N-terminal de la proteïna en conformen un tercer -que rep el nom de domini Hep- ja que va ser caracteritzat per primera vegada a les proteïnes *H2.0* de *Drosophila*, a les *engrailed* i a les portadores de domini *paired* (Allen i col., 1991). Més tard s'ha comprovat que aquest domini és present a altres classes de gens amb homeobox (com *Gbx*, *Msh*, *NKs*, *gsc*, etc.) i fins i tot a altres factors de transcripció amb altres dominis com la família *Snail-Slug* amb dits de Zinc o la família *HNF3* amb dominis *forkhead* (per a una llista completa veure Williams i Holland, 2000).

La relació de la classe *Emx* amb altres membres de la superclasse *Antp* està més clara que en el cas abans esmentat de la classe *Mox*. Filogenèticament, és molt propera a altres classes com *Dll*, *Msh*, *Tlx*, *NK1*, *NK2*, etc. (Bürglin, 1995), és a dir, amb altres membres del grup de gens *NKL*. Que a més algun dels membres d'aquest grup de gens es trobin lligats a gens *Emx* en regions sintèniques al genoma humà i de ratolí, reforça aquesta relació de parentiu (Pollard i Holland, 2000; veieu Fig. 11).

5.5.1. El gen *empty spiracles* (*ems*) de *Drosophila*

Els gens del *cluster Hox* de *Drosophila* actuen com a gens selectors assignant identitats úniques als segments més posteriors del cap, als toràcics i als abdominals. Queden per identificar doncs, gens que puguin estar duent a terme funcions anàlogues als *Hox* en regions més anteriors a aquelles identificades pels gens del *cluster HOM* més anteriors (*Deformed* i *labial*). La troballa de tres gens [que de regions anteriors a posteriors són *orthodenticle* (*otd*) (Finkelstein i col., 1990), *empty spiracles* (*ems*) (Dalton i col., 1989) i *buttonhead* (*btn*) (Finkelstein i Perrimon, 1990)] necessaris per al correcte desenvolupament dels segments més anteriors del cap va resoldre aquesta qüestió. La funció d'aquests gens és requerida per a l'establiment de blocs contigus de segments (com a gens *gap*) i també per a donar identitat segmental al cap (com a gens selectors de manera similar als gens homeòtics) (Finkelstein i Perrimon, 1990). Cinc al·lels del mutant *ems* van ser descrits per Jürgens i col. al 1984. Aquests mutants estan caracteritzats per defectes cefàlics (en concret manca dels segments intercalar, antenal i pre-antenal, i la no invaginació del cap) i per l'absència dels *Filzkörper* (estructures cuticulars internes) als espiracles posteriors que fa que aquests semblin estar buits, la qual cosa dona nom -espiracles buits- al gen.

Dalton i col. (1989) utilitzen l'homeobox del gen *eve* per tal d'identificar altres gens amb homeobox al genoma de la mosca per hibridació creuada. Aquesta aproximació els porta a la

identificació de 7 grups de gens amb homeobox. El clon E4, situat a l'interval citogenètic 88A1-88A4, conté el gen *ems*. Un altre clon -E5- molt similar a E4 i contigu al genoma, representa un duplicat específic del llinatge de *Drosophila* d'*ems*. Mitjançant la mateixa aproximació, Walldorf i Gehring (1992), identifiquen el clon W13 per hibridació creuada amb l'homeobox del gen *muscle specific homeobox (msh)*, que representa el gen *ems*. El gen *ems* té dues fases d'expressió (Dalton i col., 1989; Walldorf i Gehring, 1992):

1. La primera fase, específica de regió cefàlica, consisteix en una sola banda anterior als estadis de blastoderm sincicial i cel·lular (Fig. 26) i durant la gastrulació, aquesta banda es troba just al límit anterior del front cefàlic. L'expressió del gen *ems* durant aquesta primera fase és depenent del producte matern *bicoid (bcd)*: a més concentració de proteïna BCD, més posterior esdevé la banda d'expressió.



Figura 26. Expressió primerenca detectada per hibridació *in situ* del gen *ems* a l'embrió de *Drosophila*. Anterior és cap a l'esquerra i posterior cap a dalt. Extret de Grossniklaus i col., 1994.

2. La segona fase, de patró metamèric, consisteix en parcel·les bilaterals en cadascun dels segments de la mosca amb expressió d'*ems* en cèl·lules ectodèrmiques i neurals al voltant de les fenedures traqueals, essent el segment d'expressió més elevada, el 8è abdominal (A8), allà on el formaran el *Filzkörper*. L'expressió en A8 està dirigida pel gen homeòtic més posterior.

Cal destacar que no tots els llocs d'expressió del gen queden afectats i són visibles al fenotip del mutant. En concret, només trobem fenotip en el cap i en el desenvolupament dels espiracles posteriors d'A8, mentre que el gen s'expressa en tots els segments del cos en un determinat moment del desenvolupament. Macías i Morata (1996) intenten explicar aquest fenomen en termes de supressió fenotípica a nivell dels membres del cluster HOM; aquests gens estarien inhibint l'acció d'*ems* a cadascun dels segments on s'expressen. Una explicació alternativa podria ser que l'altre duplicat (*E5*), pogués estar substituint la funció d'*ems* al mutant. Curiosament, *E5* i *ems* comparteixen aquest domini d'expressió reiterat, i no els altres específics d'*ems* (Dalton i col., 1989).

5.5.2. La classe *Emx* en vertebrats

S'han aïllat homòlegs del gen *ems* de *Drosophila* en un ampli ventall d'organismes des de cnidaris a humans. Els primers estudis es van fer en vertebrats i demostraven que, igual que *ems*, els gens de vertebrats estaven implicats en el desenvolupament de les parts més anteriors del sistema nerviós central (SNC).

En ratolí s'han aïllat dos gens semi-ortòlegs a *ems* que s'anomenen *Emx1* i *Emx2*, l'expressió dels quals està restringida a determinades regions del cervell anterior en desenvolupament, en concret al telencèfal i als bulbs olfactoris. El mRNA d'*Emx2* ja és detectable a 8,5 dpc a regions dorsals del neuroectoderm anterior. Un dia més tard, la seva expressió és clarament detectable a la part ventral del diencèfal presumptiu i a la part dorsal del telencèfal presumptiu. També es detecta la seva expressió a altres zones fora del SNC com, per exemple, al sistema urogenital. El gen *Emx1* es comença a expressar una mica més tard (9,5 dpc) a la regió dorsal del tub neural

anterior que donarà lloc al telencèfal. La seva zona d'expressió està continguda dins la zona d'expressió de l'altre paràleg (Simeone i col., 1992a). El tipus de patró d'expressió aniuat en el temps i en l'espai que presenten els gens *Emx* junt amb els dominis també aniuats dels semi-ortòlegs de ratolí del gen *otd* de *Drosophila* (*Otx1* i *Otx2*), un altre gen *gap* cefàlic, és evident a 10 dpc, moment del desenvolupament en el que té lloc l'especificació de les principals regions del cervell. Aquest fet va fer que Simeone i col. (1992b) proposessin que aquests gens podien estar implicats en l'establiment i/o la senyalització dels límits i la identitat d'aquestes regions del cervell anterior (Fig 27).

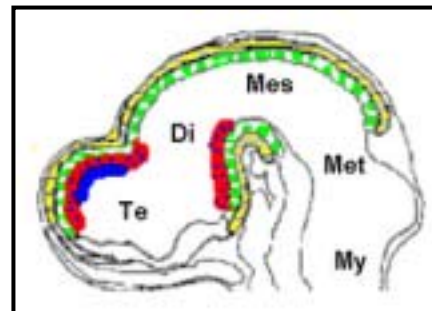


Figura 27. Esquema del cervell de ratolí on es mostra l'expressió dels gens de la classe *Emx* i *Otx* (blau, *Emx1*; vermell, *Emx2*; verd, *Otx1*; groc, *Otx2*). Adaptat de Simeone i col., 1992b.

El fenotip dels ratolins *knockout* per a cadascun dels gens està d'acord amb el seu patró d'expressió al llarg del desenvolupament, encara que no és massa sever segurament degut al rescat de la manca de funció d'un dels paràlegs per l'altre. Els homozigots nuls per al gen *Emx1* són viables i fèrtils, però els manca el *corpus callosum*, que és el principal tracte de fibres que uneix els dos hemisferis cerebrals (Qiu i col., 1996). Els homozigots mutants per al gen *Emx2* moren al cap de poques hores d'haver nascut degut a severes deficiències urogenitals que tenen. El fenotip més evident en el SNC, és palès en l'enorme reducció de l'escorça cerebral i la manca del *dentate gyrus* (Yoshida i col., 1997).

Representants de la classe *Emx* s'han aïllat en altres vertebrats com peix zebra (Morita i col., 1995) i *Xenopus* (Pannese i col., 1998). Hi ha dos gens *Emx* en cadascun d'ells, i la seva expressió en regions cefàliques és molt similar a l'abans descrita per als gens *Emx* de ratolí. Per tant, sembla que la funció dels gens amb homeobox de la classe *Emx* s'ha conservat al llarg de l'evolució des de *Drosophila* fins a vertebrats.

5.5.3. La classe *Emx* en altres espècies

A part de *Drosophila* i alguns vertebrats, s'han trobat representants de la classe *Emx* en altres organismes al llarg de l'escala evolutiva, malgrat la seva expressió i funció no està tan ben determinada com en els casos anteriors. Per exemple, hi ha un ortòleg del gen *ems* en l'amfiox *B. floridae* (*AmphiEmxA*; Williams i Holland, 2000) i sembla estar present al genoma de *C. elegans* (The *C. elegans* Sequencing Consortium, 1998). Malauradament, no es disposa de dades d'expressió per a cap dels dos gens.

Un altre dels organismes en els que s'ha trobat un gen de la classe *Emx* és en l'hidrozou *Hydractinia symbiolongicarpus* (*Cn-ems*). Aquest s'expressa en teixits endodèrmics de l'hipostom, específicament en cèl·lules de caràcter glandular i epitelial (Mokady i col., 1998). Aquest domini d'expressió suggereix als autors que l'associació dels gens del grup *Emx* amb estructures anteriors precedeix l'evolució de la simetria bilateral. Encara que aquesta afirmació sigui atractiva, cal prendre-se-la de manera prudent, ja que altres autors defensen que l'hipostom dels diblàstics no és homòleg al pol anterior dels triblàstics (Martindale i Henry, 1999).

Al genoma de la lamprea *Lamprera japonica* hi ha un ortòleg d'*ems* anomenat *LjEmx* (Myojin i col., 2001). L'expressió de *LjEmx* al llarg del desenvolupament és molt dinàmica: primer es troba tot al llarg del tub neural i mesènquima que l'envolta, on és regulat negativament més tard per passar a expressar-se als derivats cefàlics de les crestes neurals i a l'ectoderm que cobreix el telencèfal, però no en aquest. Així, l'expressió de la classe *Emx* al agnats sembla haver divergit de la de gnatòstoms, malgrat aquesta afirmació s'ha de prendre cautelosament, perquè no es demostra que no existeixi un altre paràleg al genoma de la lamprea, que s'expressi de manera "canònica" a la part dorsal del telencèfal.

Malgrat la repetida associació dels gens *Emx* a estructures nervioses anteriors, l'ortòleg de l'ascidi *Halocynthia roretzi* (*Hremx*; Oda i Saiga, 2001) no s'expressa en aquest territori. *Hremx*

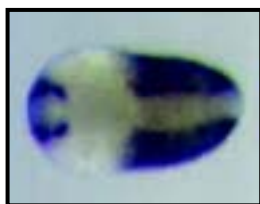


Figura 28. Expressió de *Hremx* en una visió dorsal d'una nèurula de *H. roretzi*. Anterior és a la dreta. Extret de Oda i Saiga, 2001.

s'expressa a cèl·lules ectodèrmiques anteriors i posteriors a ambdós costats de la nèurula i la larva (Fig. 28). Els autors proposen dos escenaris per a explicar la manca d'expressió d'aquest gen en regions anteriors del SNC. En el primer, pressuposen que aquest domini d'expressió que era present a l'ancestre comú de *Drosophila* i vertebrats s'ha perdut en ascidis, mentre que en el segon, aquesta expressió característica dels gens *Emx* en mosca i vertebrats en la part més anterior del SNC s'hauria adquirit independentment en aquests dos llinatges.

6. EL DOMINI bHLH

Igual que el domini homeòtic tractat al capítol anterior, el domini bàsic hèlix-volta-hèlix (de l'anglès *basic helix-loop-helix* [bHLH]) és un domini d'unió a DNA. Aquest domini, identificat per primera vegada als factors de transcripció murins E12 i E47 (Murre i col., 1989), consta d'un domini bàsic de 12 residus seguit per dues hèlix α de 15 residus separades per un loop de longitud variable (entre 10 i 17 aminoàcids). Des d'aleshores fins ara s'han descrit més de quatre-centes proteïnes amb aquest domini en diferents organismes, des del llevat a l'home, que es classifiquen en famílies i grups d'ordre major en funció de la seva similitud de seqüència dins del domini bHLH i l'existència d'altres dominis. Aquest grup de proteïnes està involucrat en esdeveniments cel·lulars clau com la proliferació, la determinació i la diferenciació cel·lular, i en processos del desenvolupament com la neurogènesi, la miogènesi, l'hematopoesi o el desenvolupament del pàncreas (Massari i Murre, 2000).

L'estructura tridimensional del domini bHLH d'E47 unit al DNA va demostrar com es fa aquesta unió i, per tant, com les proteïnes amb domini bHLH duen a terme la seva funció com a factors de transcripció. Les proteïnes amb domini bHLH formen dímers (homo i/o heterodímers) mitjançant el domini HLH i la capacitat d'unió al DNA la confereix el domini bàsic. En el cas de E47, es va resoldre que els homodímers formen un feix de quatre hèlix paral·lel, que així permet que els dominis bàsics interaccionin amb el solc major del DNA on reconeixen unes caixes de seqüència consens de sis parells de bases. Apart del domini bàsic, determinats residus del *loop* i de la segona hèlix també són capaços d'interaccionar amb el DNA (Ellenger i col., 1994).

6.1. Classificació de les proteïnes amb domini bHLH

S'han fet diferents classificacions de les proteïnes amb domini bHLH, cosa que no ha resultat ser massa senzilla degut a la heterogeneïtat que presenten els seus membres i la nomenclatura sovint erràtica. L'alliberació de les dades del genoma totalment seqüenciat de *Drosophila*, *C. elegans* i humana ha permès cercar-hi tots (o la majoria) dels gens que codifiquen

Group	Family	basic	Helix1	Loop	Helix2
A	NeuroD	RRVKANAREERTRMHGLNDALDMLRRVMP	CYSKT-----	QK--	LSKIETLRLARNYIWAIS
	Ngn	RRLKANMRERMRMHNLNAALDALREVL	PTFPED-----	AK--	LTKIETLRFAHNYIWAIT
	Ato	RRLAANAREERRRMHGLMHAFDQLRMV	IPSFNND-----	KK--	LSKYETLQMAQIYINAIAS
	Beta3	LRLNINAREERRRMHDLNDALDELRAVI	PYAHSPS-----	VRK--	LSKIATLRLAKNYILMQA
	Hist	RRLESNERERQRMHKLNMAFQALREVI	PHVRAD-----	KK--	LSKIETLTLAKNYIKSIT
	Paraxis	QRQAANAREERDRSQSWMTAFTALRTL	IPTEPVD-----	RK--	LSKIETLRLASSVIAHLA
	HyoR	QRNAANAREERARMRVLSKAFSRLKTS	LPPWPPD-----	TK--	LSKLDLTLRLASSVIAHR
	Hand	RRGTANRKRERRTQSINSAFAELRECI	PWVPAD-----	TK--	LSKIETLRLATSVIAYIM
	PTF1	LRQAANWRERRRMQSINDAFEGLRSHI	PTLPYE-----	KR--	LSKVDLTLRLAIGVINFLS
	Twist	QRVMANWRERQRTQSLMEFAALRKTI	PTLPSE-----	K--	LSKIQLTKLAARYIDFLY
	NACL	YRTAHATRRERIRVEAFMLAFAELRKL	LPTLPD-----	KK--	LSKIEILRLAICYIYSYLM
	SCL	RRVFTNSFRERMRQQHVMGAFaelRKL	LPTHPD-----	RK--	LSKMEVLRRLAMKYIGFLV
	Hesp	QRQASEREKLRMRTLARALHELRRFL	PPSVAP-----	TGQNLTKIETLRLAIRYIGHLS	
	HyoD	RRKAATMRERRRLSKVMEAFETLKRCT	SSNPN-----	QR--	LPKVEILRNAIRYIEGLQ
	AS	AVARRNERERMRVKLWNLGFATLREHV	PNGAAN-----	KK--	MSKVETLRSAVQYIRALQ
E12/E47	RRMANAREERLRVRDINEAFKELGRMV	QLHLKSD-----	KP--	QTKLLILHQAVAVILEQQ	
B	Hyc	RRRNWILERQRNDLRSSFLTLRDHVPE	LVK-----	NEK--	AAQVVLKPKATEVWHAQ
	Max	KRAHVALERKRDRDKDSFHSLRDSVP	SLQ-----	GEK--	ASPAQLDKKATEVIQYMR
	USF	RRAQNEVERRRRDKLNWIVQLSKTI	PDCHAD-----	NSKTGASGGGLSKACDWIRELR	
	Sat1	QPQDEIIAERKRREKLSQRFIALSAL	VPLGKK-----	MDKASVLLGEAIKYLKQMQ	
	Hyc7E	EPLNWEAERQRREKLNQRFYALRAV	VPNVSK-----	MDKASVLLGDASVYINELR	
	R	SIKNWMSERRRREKLNEMFLILKSV	VPSIHK-----	VDKASVLLAETIAYLKEDE	
	GBOB	ATDSLSLAERVRRREKLSERMKLLQ	ALVPGCDKV-----	TGSAVLLDEIINWVQSDQ	
	HITF	NHNLKSFVERRRRFRNDRIKELGTL	IPKSNPD-----	MR--	WNVGTLLKASVDYIRKQ
	TF4	RRRAQTQAEQKRDAIKRGYDDLQTI	VPTCQQD-FSIGSQK--	LSKAVLQKTIIDYIQFLH	
	RTG3P	KREFNAVERRRRELEKQKIKELGQ	LVPSSLNYDDLKQIK--	PNVGIILDRTEVYLQY--	
	AP4	RREIANSNERRRMQSINAGFQSLKTL	IPHTDG-----	EK--	LSKAAVLLQQTAEVYFSE
	SREBP	KRTAFNAIEKRYRSSLNDKIVELKD	LWVGTEAK-----	LNKSAVLRKAIDYIRFLQ	
	Pho4	KRESKHAEQARRNRLAVALHELASLI	PAEWKQ-----	NVSAAPSATTVEAACRYIRHQ	
	Mad	NRTSNELEKMRRAHLRNCLDGLKAI	VPLNQDAT-----	RHTTLGLLTOARALIENLKH	
	SRC	CDTLASSTEKRREQENKYLEGLAELL	SANISDIDS----	LSVKPDCKLKKTVDQIQIK	
C	ARNT	ARENHSEIERRRRNMAYITELSDMV	PTCSALA-----	RK--	PDRLTILRMVAVSHMELR
	Bmal	AREAHSQIEKRRRDKMNSFIDELAS	LVPTCNAMS-----	RK--	LDRLTVLRMAVQHMKTLR
	Trh	MKEKSRDAARSRRGKEMFEFYELAK	LPLPAAIT-----	SQ--	LDRASIIIRLTTISYLMKRD
	Sim	MKEKSKNAARTRRERESEFYELAK	LPLPSAIT-----	SQ--	LDRASIIIRLTTISYLMKRV
	PAS1	MKEKSRNAARWRRGKEMLEFRELAK	LPLPGATS-----	SQ--	LDRASIVRLSVTYLRLRR
	HIF	MKEKSRDAARSRRSKETEYFYELA	HELPLPHSVS-----	SH--	LDRASIMRLAISFLRTHK
	AHR	AEGIKSNPSSKRHRDLNTELDHLAS	LPLFSPDII-----	SK--	LDRLSVLRLSVSYLRAKS
	Clock	KRASRNKSEKRRRQFMWLKELSSM	LPGNTRK-----	MDKTIIVLEKVIQFLQKHM	
D	Emc	-----	SKLKELVPSIPQN-----	KK--	VTMEILQHVIVYILDLQ
E	Gridlock	RKKRGLIEKRRRDRIMVSLSELRR	LVPTAFKQ-----	GSAKLEKAEILQMTVDHLKMLQ	
	E(spl)	RKLSKPLMEKRRRARIMVSLQLRSL	LERHYSHQIR--	KR--	LEKADILELSVKVMRSLQ
	Hairy	RKSSKPLMEKRRRARIMVSLQLKTL	LILDALKDSS--	RHSK--	LEKADILEMTVKHLRMLQ
F	COE	-----	ALNEPTLDYGFQRLMKVIRHPGDP-----	ER--	LPMEVILKRAADLVEALY

Figura 29. Alineament de les 44 famílies de proteïnes amb domini bHLH agrupades en els grups d'ordre major (A-F). En negre (100%) i gris (80%) es mostren els residus conservats per tots els grups, mentre que els residus específics de grup es mostren en colors específics per a cada grup. Adaptat de Ledent i Vervoort, 2001.

per a proteïnes amb domini bHLH i fer-ne una classificació. S'han trobat 58 gens que codifiquen per a proteïnes amb domini bHLH al genoma de *Drosophila*, 39 al de *C. elegans* i 125 a l'humà (Ledent i Vervoort, comunicació personal). Aquestes proteïnes s'han agrupat en famílies sempre que almenys un ortòleg en dos dels organismes abans esmentats quedessin agrupats de manera monofilètica amb un valor de *bootstrap* superior al 50%. Així, s'han definit 44 famílies en animals, que s'han classificat en sis grups (grup A, B, C, D, E i F) en funció de similitud de seqüència i de la presència d'altres dominis conservats apart del domini bHLH (Fig. 29) (Ledent i Vervoort, 2001). Apart de compartir propietats estructurals, les famílies de cada grup tenen també propietats bioquímiques similars.

De la comparació del nombre de representants de cadascuna de les famílies en els genomes de *Drosophila*, *C. elegans* i l'humà, cal destacar (Ledent i Vervoort, comunicació personal):

- (i) 43 de les 44 famílies definides tenen almenys un representant en protòstoms (*Drosophila* i/o *C. elegans*) i almenys un en deuteròstoms, la qual cosa indica que aquestes famílies ja eren presents al seu ancestre comú.
- (ii) ha existit un enorme increment de la complexitat de les famílies en el llinatge de vertebrats. Així, per la majoria de famílies trobem un representant en *Drosophila* i/o *C. elegans* i més d'un al genoma humà. Només 8 famílies tenen més d'un membre en *Drosophila* i només 3 tenen més d'un membre en *C. elegans*.
- (iii) s'han perdut molts gens en el llinatge de *C. elegans*: 14 famílies no tenen cap representant al seu genoma.

6.2. Proteïnes amb domini bHLH del grup E

Al llarg d'aquesta Tesi s'ha treballat amb gens de la família hairy (grup E), així que serà l'únic grup explicat una mica en detall. Tres famílies de proteïnes -la família Gridlock/Hey, la família Enhancer of Split (E[Spl]), i la família Hairy- formen el grup E. Les tres tenen la característica de ser inhibidors actius de la transcripció i de presentar, a més del domini bHLH, dos dominis conservats: el domini *orange*, que els dóna especificitat i un tetrapèptid conservat a l'extrem C-terminal, necessari i suficient per a permetre la seva interacció amb el co-repressor general Groucho (Dawson i col., 1995). Aquest tetrapèptid és WRPW, en el cas de les famílies Hairy i E(spl), i YRPW per a la família Gridlock/Hey. A diferència dels grups anteriors, homodímers del grup E reconeixen una caixa al DNA de seqüència nucli CACGCG o CACGAG, anomenades caixes N. El fet de que reconeguin aquesta caixa, ve determinat per la presència d'un residu Prolina (P) al domini bàsic (Sasai i col., 1992).

Els homodímers de proteïnes del grup E reben el nom de repressors actius perquè regulen negativament l'expressió dels seus gens diana mitjançant la seva unió a llocs específics del DNA (caixes N) i inhibeixen la transcripció a través d'un domini diferent al d'unió al DNA anomenat domini de repressió. En el cas de la família Hairy, s'ha vist que els homodímers units al DNA, són capaços d'interaccionar amb el co-repressor general Groucho. Per a aquesta interacció, són necessaris el domini WRPW C-terminal de Hairy, i les set repeticions WD40 a la part C-terminal de les proteïnes Groucho (Jiménez i col., 1997). Recentment s'ha demostrat que les proteïnes Groucho interaccionen funcionalment amb la proteïna Rpd3 (Chen i col., 1999). Com que Rpd3 té

activitat histona deacetilasa, aquesta interacció indicaria que la inhibició mediada per homodimers del grup E involucraria processos de modificació de la cromatina (Fig. 30).

Fig. 30. Mecanisme d'inhibició activa mediada per la interacció d'homodimers del grup E al co-repressor Groucho i d'aquest a Rpd3 amb activitat histona deacetilasa. Adaptat de Massari i Murre, 2000.



6.3. La família hairy

S'han trobat representants de la família hairy en diversos organismes, d'invertebrats a vertebrats, que estan implicats en processos del desenvolupament molt dispars: és una característica d'aquesta família gènica la seva elevada pleiotropia.

En invertebrats s'han clonat i s'ha estudiat l'expressió i/o funció dels gens hairy de *Drosophila*, *Tribolium* i *C. elegans* (el gen *lin-22* de *C. elegans* ha estat classificat com a hairy o com a E[spl] segons l'autor; veure més endavant). En vertebrats s'han aïllat representants de la família hairy a peix zebra, pollet, ratolí i humà. Cal destacar l'elevada similitud de seqüència entre els gens hairy i els gens E(spl), cosa que ha provocat que a mida que s'anaven clonant representants de cadascuna d'aquestes dues famílies a vertebrats se'ls posés el nom de HES (de l'anglès *Hairy/Enhancer of Split related* en el cas de ratolí i humà) o de HER (de l'anglès *Hairy/Enhancer of Split Related* en el cas de peix zebra), seguits d'un número que, molt probablement, fa referència a l'ordre de clonatge. Aquest fet fa que la nomenclatura no sigui massa clara en el cas de vertebrats.

6.3.1. Els gens hairy de *Drosophila*

El primer representant de la família hairy aïllat va ser el gen *hairy* de *Drosophila*. Actualment es coneixen tres representants de la família en aquest mateix organisme: *hairy*, *deadpan* i *similar to deadpan*.

La funció del gen *hairy* (*h*) de *Drosophila* és necessària en dos processos diferents al llarg del desenvolupament. Primer, *h* actua com a gen *pair-rule* primari durant la segmentació de l'embrió. Més tard, entre els estadis de larva i de pupa, és necessari pel correcte desenvolupament del sistema nerviós perifèric (SNP) on inhibeix la producció ectòpica de pèls sensorials en diverses estructures de l'adult com el nòtum i l'ala. Així, en funció del tipus de mutació, els mutants presenten un fenotip variable des del típic de *pair-rule* a un fenotip de presència de pèls sensorials ectòpics, d'aquí el seu nom (pelut). L'expressió de *h* és inicialment ubíqua. Aquesta expressió primerenca es resol en un patró en forma de set bandes i un domini anterior-dorsal d'expressió al blastoderm cel·lular (Fig. 31) (Ingham i col., 1985). Com a gen *pair-rule* primari, el gen *hairy* és regulat directament per gens de tipus gap com *krüppel*, i regula l'expressió de gens *pair-rule* secundaris com *fushi-tarazu* (Carroll i col., 1988). L'expressió del gen decau ràpidament durant la gastrulació i l'extensió de la banda germinal.

El paper del gen *h* durant la neurogènesi ha estat estudiat a fons en el desenvolupament dels òrgans sensorials del SNP de *Drosophila*. Aquests òrgans estan formats per quatre cèl·lules (una neurona, una cèl·lula



Figura 31. Expressió primerenca del gen hairy al blastoderm cel·lular de *Drosophila* en forma de set bandes (1-7) i en el domini més anterior-dorsal (AD). Adaptat de Lardelli i Ish-Horowicz, 1993.

glial, la cèl·lula del pèl sensorial i la cèl·lula *socket*) que provenen d'un únic precursor neuronal anomenat precursor de l'òrgan sensorial (SOP). Molts gens que codifiquen per a proteïnes amb domini bHLH, entre ells el gen *h*, estan involucrats en el control de la formació del SOP. A mode de resum, la determinació de la SOP es pot dividir en quatre etapes (Fisher i Caudy, 1998):

- (i) Primer actuen els gens anomenats de pre-patró que limiten la zona d'expressió dels gens proneurals (Fig. 32A). Els gens de pre-patró inhibeixen l'expressió dels segons, que són reguladors positius necessaris per a la formació del SOP. El paper del gen *h* durant la neurogènesi és el d'actuar com a gen de pre-patró que inhibeix específicament l'expressió del gen proneural *achaete*, també amb domini bHLH.
- (ii) La segona etapa es defineix per l'expressió dels gens proneurals en agrupacions de cèl·lules (Fig. 32 B). Aquestes cèl·lules formen grups o *clusters* d'equivalència en els que totes les cèl·lules són competents a esdevenir SOP, però només una acabarà essent-ho.
- (iii) El tercer procés ve determinat per interaccions cèl·lula-cèl·lula que involucren la via de Notch. És un fenomen d'inhibició lateral la funció del qual és la de triar l'única cèl·lula que esdevindrà SOP. Els gens efectors de la via de Notch en aquest procés són els gens del complex E(spl), que inhibeixen l'expressió d'alguns gens proneurals i dels gens precursors en les cèl·lules del grup d'equivalència que no donaran lloc al SOP (Fig. 32 C).
- (iv) la única cèl·lula en la que els productes dels gens del complex E(spl) no han inhibit l'expressió dels gens precursors donarà lloc al SOP (Fig. 32 D). Un dels gens precursors és, sorprenentment, un dels gens de la família hairy, *deadpan* (veieu més endavant).

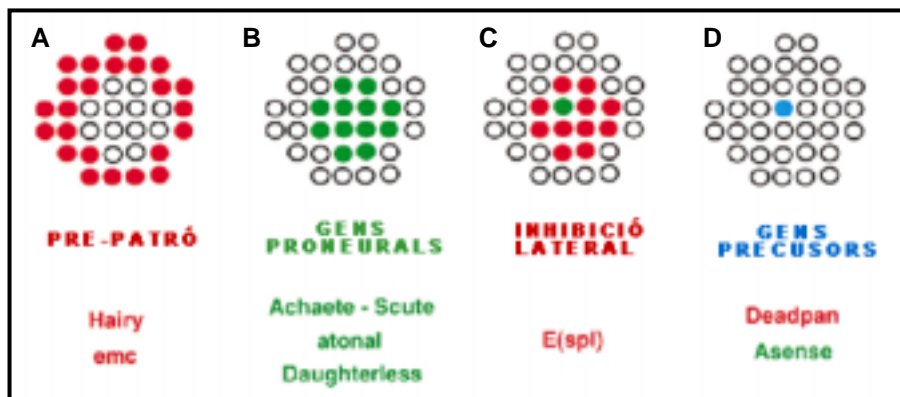


Figura 32. Paper de les proteïnes amb domini bHLH durant la formació dels SOP durant la neurogènesi de *Drosophila*. L'expressió dels gens de pre-patró (A) limita l'expressió dels gens proneurals als *clusters* d'equivalència (B). Fenòmens d'inhibició lateral dins els *clusters* (C) limiten l'expressió dels gens precursors a una única cèl·lula (D) que esdevindrà SOP. Tots els gens assenyalats codifiquen per a proteïnes amb domini bHLH. Adaptat de Fisher i Caudy, 1998.

El segon representant de la família, *deadpan* (*dpr*), està involucrat en dos processos al llarg del desenvolupament: la determinació del sexe i el control de la neurogènesi. En *Drosophila*, el sexe somàtic es determina pel quocient de cromosomes sexuals (X) respecte autosomes (A), anomenat quocient X:A. Un quocient d'1 resulta en el desenvolupament d'una femella, mentre que un quocient de 0,5 resulta en el desenvolupament d'un mascle. Aquest quocient es tradueix bioquímicament pels nivells de producte d'elements "del numerador" de gens lligats al X, en relació als nivells de producte "del denominador" produïts per gens autosomals. Es coneixen varis gens que codifiquen per a productes del "numerador" que,

curiosament, també tenen domini bHLH com *Daughterless* o *Scute*, mentre que *Dpn* és l'únic producte del denominador conegut. Un quocient de 0,5 fa que no hi hagi expressió del gen *Sex lethal*, cosa que desencadena el programa masculí (Younger-Shepherd i col., 1992). *dpn* també s'expressa durant la neurogènesi com a gen precursor (Fig. 32 D). El paper de *dpn* en aquest procés no està massa clar ja que *a priori* sembla que actua de manera oposada al seu paràleg *hairy*, com a regulador positiu de la diferenciació neuronal.

Fa menys de dos anys que l'alliberament de la seqüència completa del genoma de *Drosophila* ha permès la identificació del gen *similar to deadpan (side)*. Aquest gen, anomenat així per la seva elevada similitud de seqüència a *dpn*, està expressat en un conjunt de cèl·lules del SNC (Moore i col., 2000).

6.3.2. Ortòlegs de *hairy* en altres invertebrats

En l'embrió de banda germinal curta *Tribolium*, el gen *hairy* presenta dos dominis d'expressió. A l'estadi de blastoderm inicial trobem la seva expressió en dues bandes. Noves bandes d'expressió es van afegint de manera ordenada per la part posterior a mida que van apareixent nous segments per creixement de la part posterior de l'embrió. El segon domini d'expressió és no segmental i té lloc a la part més posterior del mesoderm (Sommer i Tautz, 1993). A diferència del gens de la família *hairy* de *Drosophila*, no s'ha trobat cap relació entre el gen *hairy* i el desenvolupament del SN en aquest coleòpter.

El gen *lin-22* de *C.elegans* no s'agrupa amb la resta de membres de la família *hairy* de proteïnes amb domini bHLH en un arbre filogenètic, malgrat ser la família amb la que comparteix més similitud de seqüència. Això ha fet que sigui considerat com a representant de la família *hairy* pels qui el van clonar (Wrischnick i Kenyon, 1997), mentre que és considerat com a membre de la família E(spl) per altres (Ledent i Vervoort, 2001). El seu estudi funcional revela característiques comuns amb el gen *hairy* de la mosca, però també amb els gens del tipus E(spl), així que no es resol aquesta qüestió. En concret, la presència de neurones extra en el nemàtode mutant tant podria ser conseqüència d'una manca de funció del tipus *hairy* com del tipus E(spl), ja que ambdues famílies estan involucrades en una inhibició de la diferenciació neuronal prematura.

6.4. Funcions de la família *hairy* en vertebrats

S'han clonat representants de la família *hairy* en molts vertebrats, des del peix zebra a l'humà. Com ja s'ha esmentat amb anterioritat, hi ha tres factors que dificulten la inclusió d'aquests gens en la família *hairy* o, pel contrari, en la família E(spl):

- I. els gens reben el nom de HES (de l'anglès *Hairy/Enhancer of Split related* en el cas de ratolí i humà) o de HER (de l'anglès *Hairy/Enhancer of Split Related* en el cas de peix zebra), seguits d'un número que, molt probablement, fa referència a l'ordre de clonatge i no a la família de la que formen part. Per exemple, els gens *her6* i *her9* de peix zebra són semi-ortòlegs del gen *HES1* de ratolí i pertanyen a la família *hairy*, mentre que el gen *her1* no ho és i pertany a la família E(spl). En principi, veient només el nom del gen tendríem a pensar que els gens *HES1* i *her1* pertanyen a la mateixa família, quan no és així. En els únics organismes on la relació està clara és en pollet i en *Xenopus*, ja que als gens de la família *hairy* se'ls ha anomenat *c/X-hairy* respectivament, seguits d'un número (1-2).

- II. altres vegades gens de famílies diferents porten a terme funcions més similars que gens de la mateixa família. Continuant amb l'exemple anterior, *HES1* sembla ser funcionalment més semblant a *her1* que els gens de la seva mateixa família *her6* i *her9*.
- III. la relació que tenen amb la via de Notch. Com hem vist, en *Drosophila* els gens del grup E(spl) actuen com a efectors de la via de senyalització de Notch durant el procés d'inhibició lateral en el desenvolupament dels SOP. Pel contrari, fins al moment no s'ha establert cap relació entre els gens de la família hairy de *Drosophila* (*h*, *dpp* i *side*) amb la via de Notch. En canvi, en vertebrats s'han descrit tant gens hairy com gens E(spl) com gens de la tercera família del grup E de gens amb doïmini bHLH, Gridlock/Hey, com a gens diana de la via de senyalització de Notch.

En vertebrats, els gens de la família hairy duen a terme múltiples funcions al llarg de quasi tot el desenvolupament. Això ha fet que cada grup de recerca s'hagi fixat i hagi estudiat a fons una d'elles. No hi ha dubte de que la més estudiada és la que aquests gens duen a terme durant la somitogènesi en relació amb la via de Notch.

6.4.1. Els gens hairy i la somitogènesi

La segmentació del cos dels vertebrats és evident a nivell de la columna vertebral, els músculs associats i el sistema nerviós perifèric. Aquest patró segmental es produeix gràcies a la producció seqüencial dels somites mesodèrmics durant el desenvolupament, procés que rep el nom de somitogènesi. Els somites apareixen bilateralment al costat del tub nerviós i la notocorda, com a esferes epitelials que broten de l'extrem més anterior del mesoderm presomític (PSM). La somitogènesi progressa d'una forma altament coordinada en el temps i en l'espai: es forma un nou parell de somites cada X temps que és específic de cada espècie a la part més anterior del PSM.

Ja fa molt de temps que models teòrics havien predit l'existència d'un rellotge molecular que actua a nivell del PSM responsable de la creació d'aquesta peridiocitat en la somitogènesi (Cooke i Zeeman, 1976), però no ha estat fins fa poc que proves moleculars han validat la seva existència. La primera prova va ser l'estudi de l'expressió d'un dels gens de la família hairy, *c-hairy1*, al PSM de pollet (Palmeirim i col., 1997). Aquest gen s'expressa d'una manera dinàmica i sincrònica bilateralment, en forma d'onada d'expressió que travessa tot el llarg del PSM en sentit postero-anterior cada 90 minuts, temps equivalent a la formació d'un somita. L'expressió d'aquest gen es pot dividir en tres etapes: en la primera, l'mRNA del gen apareix en un gran domini a la zona posterior del PSM; en la segona, el domini es fa més petit i en zones més anteriors del PSM; i en la tercera, l'expressió del gen queda restringida a una banda a la part anterior del PSM que representa la part posterior del somita que s'està formant (s0) (Palmeirim i col., 1997). Tres anys més tard, es troba un altre representant de la mateixa família, *c-hairy2*, que presenta una expressió dinàmica al PSM molt similar a la del seu paràleg (Jouve i col., 2000). L'única diferència entre aquests dos gens és la seva expressió en els somites ja formats: mentre que *c-hairy1* s'expressa al compartiment posterior, *c-hairy2* ho fa a l'anterior (Fig. 33). Aquestes oscil·lacions es detecten per primera vegada al territori on se suposa que estan les cèl·lules soca que donen lloc al mesoderm paraxial o PSM, indicant que comencen en tots els precursors de mesoderm paraxial tan bon punt aquests ingressen per la línia primitiva. Per tant, el nombre de cicles que ha patit una determinada cèl·lula està en relació amb el nivell axial que aquesta ocuparà un cop formi part d'un somita (Jouve i col., 2002).

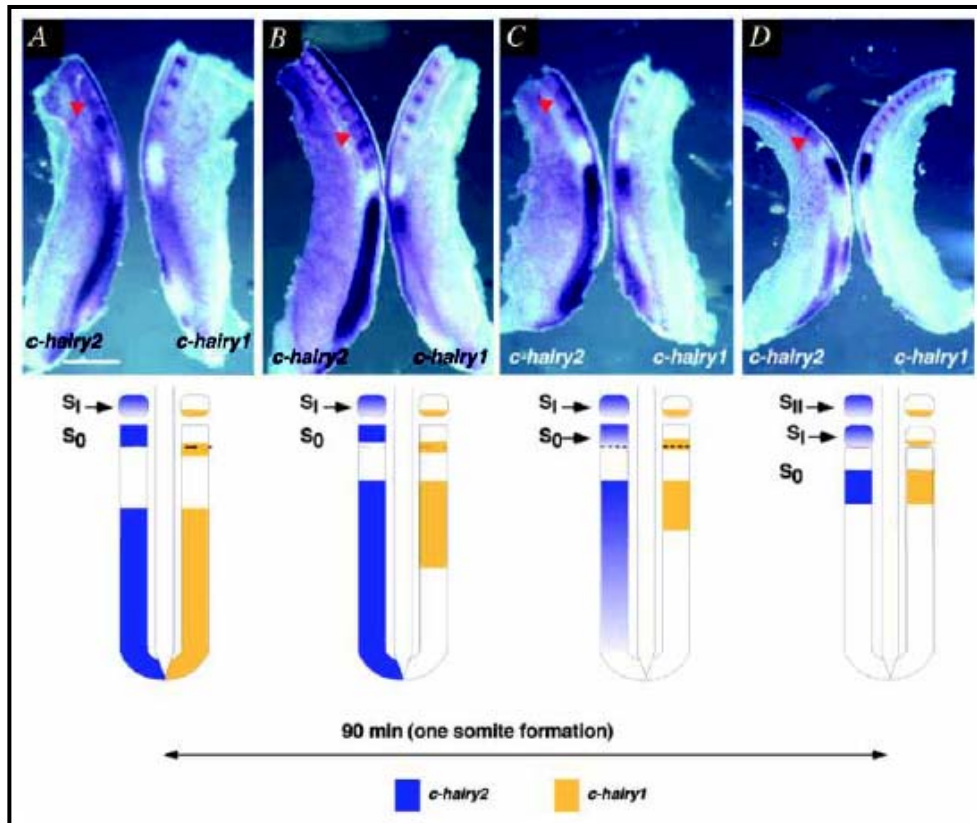


Figura 33. Expressió (superior) i esquema de l'expressió (inferior) dels gens *c-hairy1* (groc) i *c-hairy2* (blau) en l'embrió de pollet. Aquests gens s'expressen al PSM d'una manera cíclica amb una periodicitat que correspon al temps necessari per a la formació d'un somita. A partir de Jouve i col., 2000.

L'expressió cíclica d'aquests dos gens és una propietat intrínseca de les cèl·lules del PSM, ja que es manté en PSMs mantinguts *in vitro* (separats d'altres teixits com l'ectoderm) i a més, no depèn de cap senyal que difongui des de territoris posteriors perquè pot mantenir-se en PSMs mantinguts *in vitro* als que s'ha tallat la part més posterior. El fet de que la proteïna amb domini bHLH Hairy de *Drosophila* fos un potent repressor de la transcripció va fer pensar als autors que podia ser ella mateixa la responsable de l'expressió cíclica del gen i, per tant, el mateix gen *hairy* representaria el rellotge molecular lligat a l'elevada coordinació de la somitogènesi. El fet de que la inhibició de la síntesi proteica no tingui cap conseqüència sobre el comportament cíclic dels gens hairy en pollet, els situa per sota del rellotge i no com a components d'aquest (Palmeirim i col., 1997; Jouve i col., 2000). Així doncs, els components d'aquest rellotge molecular quedaven per identificar, malgrat el comportament dels gens *hairy* prefigurava la seva existència.

Des del descobriment del comportament cíclic d'aquests gens, s'ha descrit el comportament també cíclic d'altres gens de la mateixa família o de les altres dues famílies - E(spl) i Hey/Gridlock- que junt amb hairy formen el grup E de factors de transcripció amb domini bHLH. A la taula III es resumeixen els gens de la família hairy que s'han trobat en diferents vertebrats (indicant si mostren expressió cíclica o no al PSM) i gens d'altres famílies de la classe E que mostren expressió cíclica al PSM.

Taula III. Gens amb domini bHLH del grup E implicats en la somitogènesi

organisme	gen	família	dinàmica al PSM*	referència
pollet	<i>c-hairy1</i>	Hairy	✓	Palmeirim i col., 1997
pollet	<i>c-hairy2</i>	Hairy	✓	Jouve i col., 2000
pollet	<i>c-Hey2</i>	Hey/Gridlock	✓	Leimeister i col., 2000
ratolí	<i>HES1</i>	Hairy	✓	Jouve i col., 2000
ratolí	<i>HES7</i>	E(spl)	✓	Bessho i col., 2001
ratolí	<i>Hey2</i>	Hey/Gridlock	✓	Leimeister i col., 2000
<i>Xenopus</i>	<i>X-hairy1</i>	Hairy	-	Davis i col., 2001
<i>Xenopus</i>	<i>X-hairy2</i>	Hairy	-	Davis i col., 2001
peix zebra	<i>her6</i>	Hairy	-	Pasini i col., 2001
peix zebra	<i>her9</i>	Hairy	-	Leve i col., 2001
peix zebra	<i>her1</i>	E(spl)	✓	Holley i col., 2000

(*) En vermell s'indica manca d'expressió dinàmica al PSM (-) i en verd presència (✓).

D'aquesta taula les següents observacions mereixen especial consideració:

- I. el gen *c-Hey2* s'expressa d'una manera cíclica molt similar a com ho fan els gens hairy, i que s'ha demostrat la interacció entre les proteïnes c-hairy1 i 2 i la proteïna c-Hey2, fet que podria indicar que no estiguessin actuant de manera independent, sinó que fos una combinatòria de proteïnes amb domini bHLH la que produís una correcta regulació dels gens diana (Leimeister i col., 2000).
- II. el gen *HES7* de ratolí cicla de manera molt similar a *HES1* al PSM de l'embrió i el gen *HES1* no cicla en el *knockout* per *HES7*. Per tant, això indicaria que *HES7* està per sobre de *HES1* (Bessho i col., 2001).
- III. fins al moment no s'ha trobat cap gen de *Xenopus* que cicli de manera anàloga als gens que ho fan en altres vertebrats (Davis i col., 2001).
- IV. cap dels gens de la família hairy de peix zebra cicla, sinó que el gen que cicla és *her1* (Holley i col., 2000).

Una fita molt important en el mecanisme molecular de la somitogènesi va ser el descobriment d'expressió cíclica (gairebé idèntica a la dels gens abans esmentats) del gen *lunatic fringe* (*L-fng*) en pollet i ratolí (McGrew i col., 1998; Aulehla i Johnson, 1999), ja que permetia establir un vincle entre els gens "cíclics" i la via de senyalització de Notch. Els fenotips mutants dels components de la via de Notch la implicaven en la somitogènesi, i el fet de que la seva funció fos necessària per al comportament cíclic dels gens abans esmentats la situaven per sobre d'aquests, potser formant part del rellotge molecular encarregat d'establir la seva expressió periòdica (del Barco Barrantes, 1999; Jouve i col., 2000). A més, per la majoria d'aquests gens cíclics s'havia demostrat la seva posició com a gens diana de la via de Notch (Jarriault i col., 1995). Malgrat això, el fet de que els patrons d'expressió al PSM dels components de la via fossin estàtics no permetia trobar una relació clara entre la necessitat d'aquests gens i el comportament cíclic dels gens hairy entre d'altres. *L-fng* codifica per a una proteïna amb activitat glicosil transferasa que modula l'activitat de la via de Notch ja que regula l'afinitat del receptor Notch pels seus lligands (Moloney i col., 2000). Així, *L-fng* podria localitzar la via de senyalització de Notch a dominis discrets del PSM (Aulehla i Johnson, 1999). *L-fng* no presenta

una expressió dinàmica en peix zebra, sinó que és un dels lligands, *DeltaC*, el que l'exhibeix, lligant així de nou la via de Notch amb el rellotge molecular responsable de la segmentació periòdica del PSM (Jiang i col., 2000).

Els gens hairy de vertebrats també estan involucrats en la regionalització antero-posterior dels somites. A diferència de la formació del patró dorso-ventral i medio-lateral dels somites que té lloc després de la seva segmentació, la regionalització antero-posterior té lloc a nivell del PSM. L'adquisició de la identitat anterior o posterior per part de cèl·lules situades a la part més anterior del PSM es fa palesa per l'expressió en forma de bandes de molts gens, la majoria dels quals relacionats una altra vegada amb la via de senyalització de Notch, entre els que hi ha els gens *hairy*. Pel que fa referència als gens de la família hairy, l'expressió diferencial en compartiments somítics anteriors i posteriors que comença a nivell del PSM es manté en tots els somites ja formats. Curiosament, malgrat que a nivell de seqüència el gen *c-hairy2* és més similar a l'únic gen de ratolí *HES1*, el segon és més similar a *c-hairy1* a nivell d'expressió: *HES1* i *c-hairy1* es detecten al compartiment posterior, mentre que l'mRNA de *c-hairy2* és present a l'anterior (Jouve i col., 2000).

Com es pot veure, el procés de somitogènesi en vertebrats és molt complex. Cada vegada s'estan descobrint més gens implicats que semblen tenir en comú algun tipus de relació amb la via de Notch, ja siguin els propis components de la via o gens diana. Així, en una revisió actual es parla de l'estreta interrelació de més de vint gens, i això tenint en compte només el procés de segmentació (Saga i Takeda, 2001). A mode de resum, només comentar un dels models teòrics proposats i intentar exposar com algunes dades moleculars dels últims anys el recolzen. Cooke i Zeeman (1976) proposen el model del rellotge i l'onada de maduració (de l'anglès *clock-and-wavefront model*) segons el qual, existeix un rellotge cel·lular capaç d'interaccionar amb una onada que va progressant en direcció antero-posterior. Així, la producció de les fronteres entre somites contigus té lloc quan cèl·lules que es troben en un estat permissiu del cicle que dicta el seu rellotge intern es troben amb una onada de maduració que progressa en sentit antero-posterior al PSM. Com a resultat, els somites es formen en el mateix sentit antero-posterior a una freqüència temporal dictada pel rellotge i a una freqüència espacial definida per la velocitat de progressió de l'onada de maduració. Aquest model explicaria el control de la periodicitat i la naturalesa autònoma cel·lular de l'oscil·lació dels gens "cíclics". La recent troballa de que la via de senyalització de factors de creixement fibroblàstic (FGF) -en particular la de FGF8- divideix el PSM en dues regions, una de més anterior (I) que és competent a la segmentació i que no expressa *FGF8*, i una altra de posterior (II) no competent i que expressa *FGF8*, representaria l'onada de maduració antero-posterior, ja que el límit d'expressió d'*FGF8* va progressant en el mateix sentit a mida que avança la somitogènesi (Dubrulle i col., 2001). Així, quan les cèl·lules passen el punt de transició (surten de la regió II per entrar a la I), poden aturar aquesta oscil·lació i expressar els gens necessaris per a la segmentació. Si només aquelles que han experimentat un determinat nombre de cicles d'expressió dels gens cíclics a la regió II són capaces d'iniciar un programa de segmentació, s'assegura la peridiocitat espacial en la formació de les fronteres entre somites contigus.

6.4.2. Altres funcions de la família hairy en vertebrats

Una característica dels gens de la família hairy és la seva elevada pleiotropia. Per tant, malgrat l'aspecte més estudiat ha estat la seva implicació en la somitogènesi, estan involucrats

en molts altres processos, tals com el desenvolupament del sistema nerviós i el de determinats derivats endodèrmics i mesodèrmics. En general, estan involucrats en vies de diferenciació cel·lular que contenen altres proteïnes amb domini bHLH. En aquestes vies, la combinatòria de proteïnes repressores [com les de la família hairy o E(spl)] amb altres proteïnes activadores (com MyoD, Mash/Scute o Neurogenines), produeix que una cèl·lula triï el seu destí entre seguir proliferant o bé començar a diferenciar-se. En alguns d'aquests processos, també s'ha involucrat a la via de Notch com a responsable de l'activació dels gens de la família hairy o de la família E(spl).

El gen *her9* de peix zebra s'expressa de manera específica durant el desenvolupament del sistema nerviós bàsicament al cervell anterior i mitjà i, de manera transitòria, al cervell posterior (Leve i col., 2001). La presència del seu mRNA es fa evident al final de la gastrulació, en la part més anterior de la placa neural i, acabada la gastrulació, la seva expressió apareix en forma de bandes longitudinals al llarg de la placa neural anterior. Més tard, durant la somitogènesi, l'expressió de *her9* es detecta al tub neural i al mesoderm de la línia mitja i, en nivells baixos, als somites ja formats. Després la seva expressió al tub neural és clarament observable en forma de bandes transversals a la part anterior del tub neural i al cervell posterior, deixant un buit d'expressió a la MHB (Fig. 34A). En estadis més tardans, la seva expressió es detecta en zones de la pituitària anterior i als ulls, així com a l'hipotàlam anterior. L'altre representant de la família hairy de peix zebra -*her6*- s'expressa en somites i, a més a més, durant la segmentació del cervell posterior de manera dinàmica i en forma de bandes (Pasini i col., 2001) (Fig. 34B). D'un patró molt complex se'n deriva que després de 24 hores de desenvolupament, tots els rombòmers han expressat el gen *her6*, malgrat ho fan en diferents estadis i a diferents nivells. El seu mRNA es detecta també a la notocorda encara que a nivells molt baixos.



Figura 34. Expressió dels gens *her9* (A) i *her6* (B) en el sistema nerviós de l'embrió de peix zebra. L'expressió de *her9* es detecta tot al llarg del cervell anterior, mitjà i posterior a excepció del buit que deixa a la MHB. L'expressió de *her6* és molt dinàmica durant el desenvolupament del cervell posterior. En aquesta imatge es detecta la seva expressió a la part més anterior del cervell, així com als rombòmers 3 i 5 del cervell posterior. A partir de Leve i col., 2001 i de Pasini i col., 2001.

En *Xenopus*, l'expressió del gen *Xhairy2* està restringida a alguns derivats del neuroectoderm i, en el mesoderm, al PSM i als pronefros. L'altre gen, *Xhairy1*, s'expressa de manera idèntica que *Xhairy2* al PSM i, encara que també s'expressa a teixits nerviosos, la seva expressió no és solapant (Davis i col., 2001). En embrions de *Xenopus* s'ha demostrat que la via de Notch produeix la inhibició del gen *XMyoD* a través de la transactivació de *Xhairy1*. Així, s'ha hipotetitzat que la funció d'aquest gen seria la de mantenir els precursors miogènics en un estat indiferenciat mitjançant la repressió de *XMyoD* (Umbhauer i col., 2001).

Els dos gens *hairy* de pollet també s'expressen durant el desenvolupament del sistema nerviós i en derivats mesodèrmics als esbossos de les extremitats (Jouve i col., 2000). A la regió posterior del tub neural, l'expressió d'aquests dos gens està regionalitzada i és complementària:

c-hairy1 s'expressa a la placa del terra i al terç més dorsal incloent la placa del sostre, mentre que *c-hairy2* ho fa a la zona mitja del tub (Fig. 35).

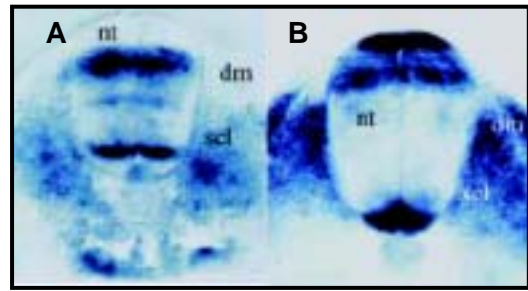


Figura 35. Expressió dels gens *c-hairy2* (A) i *c-hairy1* (B) en el sistema nerviós de pollet. En talls transversals s'observa el patró restringit i complementari d'aquests dos gens. Adaptat de Jouve i col., 2000.

El gen *HES1* de ratolí ha estat indiscutiblement el gen hairy més estudiat pel que fa referència a altres funcions diferents a la que duu a terme en la somitogènesi. S'expressa a nivells basals a gairebé tot l'embrió, però és en cèl·lules epitelials com les del neuroepiteli, en derivats mesodèrmics -com els músculs- o a l'intestí, on ho fa a alts nivells (Sasai i col., 1992). Durant el desenvolupament del sistema nerviós, la seva expressió comença tot al llarg de la placa neural, fins a quedar més restringida a determinades zones del tub neural de l'embrió més tardà. Una anàlisi més detallada permet la identificació d'aquestes zones com a regions on les cèl·lules neuroepitelials estan proliferant, en les que la funció de *HES1* és la d'inhibir la seva prematura diferenciació. Així, la sobre-expressió d'aquest gen provoca una inhibició de la diferenciació glial i neuronal (Ishibashi i col., 1994). Per altra banda, la seva manca de funció produeix defectes severos en la neurulació i el tamany de les vesícules telencefàliques que són més petites. També hi ha sobre-expressió de gens neurals de diferenciació (com el gen amb domini bHLH *Mash-1*) i de marcadors de diferenciació terminal de neurones. Una altra característica és l'aparició prematura de neurones postmitòtiques (Ishibashi i col., 1995). La via de Notch també està per sobre de *HES1* en aquesta inhibició que té sobre la diferenciació dels precursors neuronals (Ohtsuka i col., 1999). És en el sistema nerviós on s'han descrit funcions per a *HES1* que són molt similars a les que el gen *hairy* de *Drosophila* fa durant el desenvolupament dels òrgans sensorials (secció 6.3.1). Per exemple, en el desenvolupament de l'epiteli olfatiu la proteïna HES1 actua primer com a gen de pre-patró delimitant un domini d'expressió de gens pro-neurals per inhibició de *Mash1* (ortòleg d'*achaete* de *Drosophila*). Més tard però, actua també durant el procés d'inhibició lateral, que en *Drosophila* involucraria a gens de la família E(spl) i no al gen *hairy* (Cau i col., 2000).

El ratolí *knockout* per *HES1* també presenta una hipoplàsia pancreàtica produïda per una accelerada diferenciació dels precursors epitelials cap a cèl·lules endocrines postmitòtiques definides per l'expressió de glucagó. A més la seva manca de funció també està associada amb una sobre-expressió de molts gens amb domini bHLH com *Mash1*, *Ngn3* o *Neurod*, fet que comporta una prematura diferenciació de múltiples tipus cel·lulars endocrins a l'intestí i a l'estómac en desenvolupament (Jensen i col., 2000). De fet, en aquest procés *HES1* duu a terme una funció anàloga a la que té durant el desenvolupament del sistema nerviós, on és activat per la via de senyalització de Notch i la seva funció és la de contrarestar l'activitat estimuladora de la diferenciació d'altres proteïnes amb domini bHLH.

Objectius

El nostre grup està interessat en entendre com ha estat possible l'aparició d'algunes de les innovacions morfològiques que presenten els vertebrats. Com ja s'ha esmentat en la introducció, per a aquest propòsit cal escollir com a model un organisme al que li mancaràn aquestes innovacions, però que alhora sigui molt proper als organismes que les presenten. Proves de diferent tipus (morfològiques, moleculars, etc.) recolzen que aquest organisme és l'amfiox. D'altra banda, com la morfologia final dels organismes la dicten els gens que regeixen el seu desenvolupament, cal triar els gens apropiats a les innovacions que es volen estudiar. Com ja s'ha esmentat també, s'ha comprovat que el canvi en el nombre i/o regulació dels gens Hox ha estat instrumental en l'evolució de les diferents morfologies. Així, en un inici ens vam plantejar l'estudi de determinats gens del desenvolupament en l'amfiox, en concret el dels gens Hox que no havien estat caracteritzats, i el d'altres gens del grup *extended Hox*.

En concret, els objectius inicialment proposats van ser:

1. La caracterització de la part més posterior o 5' del *cluster Hox* de l'amfiox, per veure si existien més gens posteriors lligats al gen *AmphiHox10* mitjançant la tècnica de passeig cromosòmic amb l'ús de genoteques genòmiques de còsmids primer, i amb la construcció i posterior utilització d'una genoteca de fags després.
2. La caracterització d'una altra classe del grup *extended Hox* -la classe *Evx*- mitjançant tant l'aïllament dels clons genòmics i de cDNA, com l'estudi de la seva expressió al llarg del desenvolupament. Posteriorment vam voler determinar si estaven lligats al *cluster Hox* anteriorment caracteritzat, com passa en vertebrats i en cnidaris.
3. La caracterització de l'última classe del grup *extended Hox* -la classe *Mox*- mitjançant tant l'aïllament dels clons genòmics i de cDNA dels representants d'aquesta classe, com l'estudi de la seva expressió al llarg del desenvolupament.

Com el gen *Mox* d'amfiox va resultar estar íntimament lligat a la somitogènesi, és per aquest motiu que, *a posteriori* vam trobar interessant plantejar-nos un nou objectiu que va ser:

4. La caracterització dels gens de la família hairy degut a la seva implicació en la somitogènesi, mitjançant l'aïllament dels clons de cDNA dels representants d'aquesta classe i l'estudi de la seva expressió al llarg del desenvolupament.

Objectives

Our group is interested in understanding how vertebrates have acquired some morphological innovations that appear as vertebrate novelties. To this end, we have studied specific developmental genes in amphioxus, the closest living invertebrate relative to vertebrates. As mentioned in the Introduction, changes in the number and/or regulation of Hox genes are known to have been instrumental for the evolution of different morphologies. We thus decided to study these genes and other homeobox-containing genes that are closely related to the Hox genes, i. e. the rest of extended Hox genes.

Hence, we planned to characterise:

5. The posterior-most end (5' end) of the *B. floridae* Hox gene cluster, to ascertain whether there were more posterior Hox genes 5' to *AmphiHox10*.
6. The *Evx* class of *B. floridae* at both the genomic and the cDNA level, as well as their expression pattern along amphioxus development. Moreover, we wanted to analyse whether the *Evx* genes were linked to the 5' part of the Hox gene cluster.
7. The *B. floridae* *Mox* class genes at the genomic and cDNA level, as well as their expression pattern along amphioxus development.

Since the amphioxus *Mox* gene appeared to be closely related to the process of somitogenesis, we decided, in order to gain more insight into this process, to characterise:

8. The *B. floridae* hairy family. In particular, we decided to isolate the cDNA clones of the amphioxus hairy family members and study their expression pattern along amphioxus development.

Introducció als resultats

Els resultats d'aquesta Tesi Doctoral s'han dividit en cinc capítols, cadascun dels quals consta d'un article (dos publicats, dos en premsa i un en preparació). A més, s'ha afegit un annex a continuació dels capítols II i III en el que s'afegeixen resultats que fan referència a aquests dos capítols, que no estan publicats ni en preparació. Cal dir que els resultats no s'han estructurat seguint el mateix ordre que la introducció, per tal de donar-los una visió més general pel que fa referència al genoma de l'amfiox. També esmentar que alguns dels resultats van ser duts a terme en estreta col·laboració amb un investigador postdoctoral que formava part del grup, en Dave Ferrier. Així, l'apartat de resultats té la següent estructura:

- I. Caracterització de la Classe Mox de *B. floridae*
- II. Caracterització de l'extrem 5' del *cluster* Hox de *B. floridae*
- III. Caracterització de la Classe Even-skipped (Evx) de *B. floridae*
- Annex: Lligament dels gens Evx al *cluster* Hox de *B. floridae*
- IV. Evolució del genoma de l'amfiox. Un nou cas de duplicació: la Classe Empty spiracles (Emx) de *B. floridae*
- V. Caracterització de la família hairy de *B. floridae*

L'amfiox és un sistema model força utilitzat en el camp de l'Evo-Devo i, més específicament per estudiar com els vertebrats s'han originat a partir d'organismes invertebrats. El fet de que el seu genoma hagi escapat de les múltiples duplicacions gèniques/genòmiques que van ocórrer molt aviat durant l'evolució dels vertebrats, fa que l'amfiox sigui sovint triat com a model on estudiar la funció d'una família gènica de vertebrats mitjançant l'estudi del seu (en principi) únic gen "lliure de redundàncies".

Malgrat això, l'amfiox no és l'ancestre dels vertebrats, sinó tan sols un organisme morfològicament molt semblant a com aquest suposadament era, i el seu genoma ha divergit en el temps tant del genoma d'aquest ancestre com ho han fet els genomes dels vertebrats. Per tant, el genoma de l'amfiox també pot haver patit canvis (com duplicacions gèniques) durant aquests més de 500 milions d'anys d'evolució independent, encara que no hagi patit les duplicacions a gran escala que van ocórrer molt aviat durant l'evolució dels vertebrats. De fet, al llarg d'aquest treball s'han trobat més "excepcions" que "normes" pel que fa referència a la singularitat (existència de gens de còpia única) del seu genoma. Només en el primer capítol (la classe Mox) s'ha trobat un gen únic pro-ortòleg de varis gens de la mateixa família en vertebrats. En la resta de capítols es descriuen duplicacions específiques al genoma de l'amfiox. En especial en l'últim capítol (la família hairy) s'ha trobat un cas extrem de duplicació gènica en el genoma de l'amfiox, que té un mínim de sis gens hairy, mentre que els vertebrats només tenen dos o fins i tot tan sols un.

CAPÍTOL I

CARACTERITZACIÓ DE LA CLASSE MOX DE *B. floridae*

Minguillón, C. and Garcia-Fernández, J. (2002). *The single amphioxus Mox gene: insights into the functional evolution of Mox genes, somites and the asymmetry of amphioxus somitogenesis*. DEVELOPMENTAL BIOLOGY, en premsa.

Resum

Per PCR sobre cDNA d'embrions de *B. floridae* amb òligonucleòtids degenerats corresponents a les regions més conservades de l'*homeobox* es va amplificar el corresponent fragment del gen *Mox*. El crivellatge d'una genoteca genòmica i d'una genoteca de cDNA de larves de la mateixa espècie utilitzant el producte de PCR com a sonda va permetre l'aïllament i posterior caracterització de l'únic representant de la classe *Mox* al genoma de *B. floridae*, que hem anomenat *AmphiMox*. L'anàlisi filogenètic dels representants de la classe *Mox* en diferents espècies recolza la posició de l'amfiox com a grup germà dels vertebrats i la del gen *AmphiMox* com a pro-ortòleg dels dos gens *Mox* de vertebrats.

L'estudi de la seva expressió al llarg del desenvolupament va fer palesa la seva íntima associació amb la somitogènesi: *AmphiMox* s'expressa de manera molt dinàmica durant aquest procés. Breument, la seva expressió es detecta per primera vegada en aquella regió del mesoderm presomític (PSM) primer, i del *tail bud* després, que està donant lloc a un nou somita en sentit antero-posterior (com a mínim des del somita número 5). La seva expressió s'apaga un cop el somita s'ha acabat de formar. Aquest estudi va permetre:

- VI. ampliar el nombre de somites asimètrics en l'amfiox. En treballs morfològics clàssics es diu que l'asimetria inherent a la somitogènesi dels cefalocordats és visible a partir del novè parell de somites, o sigui només en els somites que es formen durant la segona fase de somitogènesi a partir del *tail bud*. Els nostres resultats van indicar que aquesta asimetria és present com a mínim des del cinquè parell de somites i, per tant, que no només la segona fase sinó que també la primera fase de la somitogènesi són inherentment asimètriques.
- VII. documentar l'existència d'una regió similar al PSM dels vertebrats només durant la fase primerenca de la somitogènesi de l'amfiox, doncs durant la fase tardana els somites apareixen directament del *tail bud*.

L'estudi comparatiu de l'expressió dels gens de la classe *Mox* de vertebrats respecte *AmphiMox*, i d'ambdós respecte el d'altres invertebrats, ens va permetre suggerir que:

- (i) la co-opció dels gens de la classe *Mox* en els dos subfílums de cordats amb somites (cefalocordats i vertebrats) en el procés de somitogènesi.
- (ii) la co-opció dels gens *Mox* després de les duplicacions que tenen lloc en el llinatge que dona lloc als vertebrats, en l'elaboració de derivats somítics com les vèrtebres o la musculatura de les extremitats.

CAPÍTOL II

CARACTERITZACIÓ DE L'EXTREM 5' DEL *CLUSTER* HOX DE *B. floridae*

Ferrier, D.E.K., Minguillón, C., Holland, P.W.H. and Garcia-Fernández, J. (2000). *The amphioxus Hox cluster: deuterostome posterior Flexibility and Hox 14*. EVOLUTION AND DEVELOPMENT 5, 284-293.

Resum

Mitjançant la tècnica de passeig cromosòmic vam caracteritzar la part més 5' o posterior del *cluster* Hox de *B. floridae*. En concret, vam documentar l'existència de quatre nous gens Hox (respecte els 10 gens que havien estat publicats [Garcia-Fernández i Holland, 1994]) que formen part del complex. Malgrat el *cluster* de l'amfiox havia estat definit com a "arquetípic" respecte els de vertebrats, trobem un gen "extra" al *cluster*, al que anomenem *AmphiHox14*. A la resta de gens els anomenem d'*AmphiHox11* a *AmphiHox13*, en funció a la seva posició relativa al complex i no perquè filogenèticament puguin ser assignables a un determinat grup paràleg de vertebrats.

Intentem diferenciar entre dos escenaris:

- (i) que els gens Hox posteriors provenen de duplicacions independents que han tingut lloc durant l'evolució dels vertebrats i de l'amfiox.
- (ii) que realment podem establir relacions d'ortologia 1 a 1 entre un determinat PG de vertebrats i un determinat gen de l'amfiox.

Després de dur a terme diferents aproximacions vam veure que cap és prou clara com per a triar un dels dos escenaris anteriors. Així, mentre els gens Hox de protòstoms poden ser assignables a determinats grups posteriors (*Post1*, *Post2*, o *Abd-B*), aquest fenomen no és possible entre l'amfiox i els vertebrats. A més, observem que aquest fet (la no clara assignació) té lloc entre tots els deuteròstoms en general, fet al que anomenem flexibilitat posterior dels deuteròstoms (*deuterostome posterior flexibility*), fet que reflexa que els gens Hox anteriors tenen una major restricció que els posteriors al canvi, potser lligat amb una major plasticitat de les estructures posteriors de l'embrió respecte les anteriors.

Aportació personal al treball

L'autora d'aquesta Tesi va construir la genoteca en bacteriòfags a partir de DNA genòmic d'un sol individu adult de *B. floridae* sobre la qual es van realitzar els cinc darrers passos del *chromosome walking*. La construcció d'aquesta genoteca es va fer necessària ja que treballàvem amb una genoteca de còsmids que presentava tres problemes en quant al mapatge dels clons positius, punt clau per a l'elecció del clon òptim per a fer el següent pas:

- l'elevada mida de l'insert
- el fet de que la genoteca estigués feta a partir de DNA de més d'un individu feia que tinguéssim dificultats al mapar els positius degut a l'elevat grau de polimorfisme de les poblacions naturals de *B. floridae*.

- els clons es trobaven a la Universitat de Reading (Anglaterra), la qual cosa feia que a vegades tardéssim en rebre el positiu que demanàvem, o fins i tot a vegades havia arribat un positiu equivocat.

L'autora també va col·laborar en la completa caracterització (mapatge, seqüenciació, etc.) de l'últim gen aïllat, *AmphiHox14*, així com en l'elaboració del manuscrit.

CAPÍTOL III

CARACTERITZACIÓ DE LA CLASSE EVEN-SKIPPED (EVX) DE *B. floridae*

Ferrier, D.E.K., Minguillón, C., Cebrián, C. and Garcia-Fernández, J. (2001). *Amphioxus Evx genes: implications for the evolution of the Midbrain-Hindbrain Boundary and the chordate tailbud*. DEVELOPMENTAL BIOLOGY 237, 270-281.

Resum

Vam aïllar i caracteritzar a nivell genòmic i de cDNA els dos representants de la classe Evx de *B. floridae*, als que anomenem *AmphiEvxA* i *AmphiEvxB*. L'anàlisi filogenètica d'ambdós respecte els gens de vertebrats va demostrar que es tracta d'una duplicació independent d'aquesta classe al genoma de l'amfiox: així, ambdós gens s'agrupen junts com a trans-ortòlegs dels gens Evx de vertebrats. L'estudi de: (i) la seva seqüència (en comparació amb la dels gens de la mateixa classe d'altres organismes) i, (ii) el seu patró d'expressió, ens permeten diferenciar els dos gens com a:

- prototípic: *AmphiEvxA*
- derivat: *AmphiEvxB*

L'estudi de l'expressió del gen *AmphiEvxA* va fer palesa la seva implicació en diversos estadis i processos al llarg del desenvolupament tals com la gastrulació, l'especificació de la part més posterior de l'embrió, la neurogènesi i el desenvolupament del *tail bud*, totes elles característiques típiques dels gens de la classe Evx en vertebrats.

El fet de que representants d'aquesta classe hagin estat aïllats en molts organismes al llarg de l'escala evolutiva ens va permetre, mitjançant un estudi comparatiu de la seva expressió, proposar un escenari evolutiu per a la classe Evx en el que:

- (i) identifiquem patrons únics a determinats grups com és el cas del patró *pair-rule* en alguns insectes.
- (ii) Identifiquem patrons comuns a tots els organismes bilateris com l'expressió durant la gastrulació o el desenvolupament del sistema nerviós
- (iii) identifiquem patrons comuns als cordats com l'expressió durant el desenvolupament del *tail bud* fent palès, a nivell molecular, que aquest territori representa una sinapomorfia dels cordats.
- (iv) identifiquem patrons únics a vertebrats que podrien estar lligats a l'aparició d'innovacions morfològiques com l'expressió a la frontera entre el cervell mitjà i el posterior, que seria posterior a la duplicació dels gens de la classe Evx en vertebrats, fet que suggereix que l'amfiox té un cervell bipartit (veieu però, capítol V).

Aportació personal al treball

L'autora d'aquesta Tesi va realitzar el pas de *chromosome walking* des del gen *AmphiEvxA* a *AmphiEvxB*, la qual cosa va permetre l'aïllament i caracterització del gen *AmphiEvxB* a nivell genòmic. També va aïllar i caracteritzar el cDNA del mateix gen, i estudiar la seva expressió al

llarg el desenvolupament per hibridació *in situ whole mount*. Malgrat els resultats d'expressió no es van incloure en aquest article, es van afegir a un altre publicat posteriorment (veieu capítol IV de Resultats). També va col·laborar en la preparació del manuscrit.

ANNEX I

LLIGAMENT DELS GENS EVX AL *CLUSTER* HOX DE *B. floridae*

Minguillón, C., Ferrier, D.E.K and Garcia-Fernández, J. (no publicat)

Resum

S'ha continuat la caracterització de l'extrem 5' del *cluster* Hox d'amfiox mitjançant la tècnica de *chromosome walking* (passeig cromosòmic) sobre la genoteca genòmica construïda en bacteriòfags λ FixII a partir de DNA d'un sol individu (capítol II).

En concret, s'ha continuat el *walking* a partir de l'extrem 5' del gen *AmphiHox14* (capítol II), per veure si existia lligament amb els gens de la classe *Evx* (capítol III). S'han fet 10 nous passos del passeig, que cobreixen un total d'un total d'un total de 87 Kb. Es va realitzar una hibridació amb un oligonucleòtid degenerat que reconeix la tercera hèlix de l'homeobox (anomenat SO2) sobre els fags positius. Aquesta hibridació va revelar l'existència de bandes positives de fags del segon, tercer, quart, cinquè, vuitè, novè i desè passos de *walking*. El posterior subclonatge i seqüenciació de les bandes positives va demostrar que no es tractaven de cap nou gen Hox, ni tampoc de cap dels dos gens *Evx*, sinó de seqüències no codificants només parcialment similars a l'oligonucleòtid SO2, que van ser detectades degut a les condicions de restrictivitat (força permissiva) que es van fer servir.

En aquesta secció també es discuteix sobre l'existència de lligament entre els gens *Evx* i els del complexe Hox en l'amfiox. En base a l'existència d'elements repetitius, de seqüències similars a la de l'oligonucleòtid SO2 i al patró d'expressió dels gens *AmphiEvxA* i *B*, es recolza la hipòtesi de que no estan lligats, o si ho estan, no ho és estretament. També es discuteix sobre la necessitat d'altres estratègies per a determinar (en cas de que existeixi) el lligament.

ANNEX I. ARE HOX AND EVX GENES LINKED IN THE *B. floridae* GENOME?

After the unexpected finding of a fourteenth Hox gene in the amphioxus genome, we continued the chromosome walking experiments to isolate further Hox genes, if any, and, especially, to reach the end of the *B. floridae* Hox cluster defined by the linkage with the Evx genes. This genomic walking procedure was performed using the genomic library constructed in the λ FixII phage with DNA of a single individual as described in chapter II

Ten further walking steps covering a genomic region of about 87 Kb were performed. The positive phages obtained in each walking step were hybridised with a degenerate oligonucleotide (SO2) recognising the third helix of the homeobox (Garcia-Fernández and Holland, 1994). Bands of phages from the second, third, fourth, fifth, eighth, ninth and tenth step gave a positive signal for the SO2 hybridisation. Subsequent subcloning and sequencing of these bands revealed that none of them represented Hox or Evx genes. In addition, the phage contig that covered the whole 87 Kb genomic region was hybridised with the most extreme bands of the contig that contained both amphioxus Evx genes (chapter III), and no hybridising bands were detected. Thus, the Evx genes are not in proximity of 90 Kb to the amphioxus Hox genes cluster. Moreover, there is not any further Hox gene within this genomic region (Fig. 1).



Figure 1. Map of the 87 Kb genomic region analysed upstream to *AmphiHox14* (right), and contig of about 40 Kb containing the amphioxus Evx genes (left). The lines and numbers below refer to the phage (for the Hox region) and cosmid (for the Evx region) names and relative lengths. The arrows above the gene names indicate their transcription orientation.

Evx genes are closely linked to the Hox gene cluster in a range of organisms, namely mammals (Dush i Martin, 1992; D'Esposito i col., 1991; Faiella i col., 1992), zebrafish (Amores et al., 1998) and a cnidarian species (Miller and Miles, 1993). Hence, this linkage may represent the ancestral condition in the Animal Kingdom. The examination of the amphioxus Hox gene cluster up to the *AmphiHox10* gene revealed its astonishing resemblance to the deduced putative cluster of the ancestor of vertebrates. It has a single cluster containing one representative of each vertebrate paralogous group (Garcia-Fernández and Holland, 1994). However, the 5' part of its Hox cluster is rather unusual: it has an extra Hox gene, most probably originated by independent duplication in the amphioxus lineage (chapter II), and no Evx gene closely linked to its 5' extreme (chapter III, and this section).

Two lines of evidence suggest that *Evx* genes are not closely linked to the Hox cluster in the *B. floridae* genome:

- (iii) Too many positive clones (hundreds) were obtained in the seventh and eighth step of the genomic walking from the *AmphiHox14* gene. This may be indicative of repetitive elements within the genomic region used as a probe for the screening. It is believed that the clusters in vertebrates are devoid of such DNA elements (Hart et al., 1985). Hence, it may indicate that we are already walking outside of the cluster.
- (iv) The expression pattern of *AmphiEvxA* resembles that of other *Evx* genes and not that of the posterior group of Hox genes. The vertebrate *Evx1* genes show a "canonical" *Evx*-type expression pattern, e.g. at the onset of gastrulation at the posterior end of the developing embryo, and later in the central nervous system (e. g. Bastian and Gruss, 1990). In contrast, the expression patterns of zebrafish and the mouse *Evx2* genes are reminiscent of those of the posterior Hox genes and obey the temporal colinearity of the Hox cluster genes (Dollé et al., 1994; Sordino et al., 1996). The mouse *Evx2* gene and the HoxD cluster (8 Kb) are more tightly linked than the *Evx1* gene and the HoxA cluster (45 Kb), which may account for this Hox-type regulation of the *Evx2* gene. Thus, the fact that *AmphiEvxA* is expressed as a "canonical" *Evx* gene during amphioxus embryogenesis points to a loose linkage to the posterior-most Hox gene.

The amphioxus Hox cluster is about three times as long as the average mammalian Hox clusters. Hence, the amphioxus *Evx* genes may be linked to the cluster at about 135 Kb (three times 45 Kb; the distance between the mouse *Evx1* and the HoxA cluster) of the posterior-most Hox gene, i.e. at least six additional walking steps are needed to establish such a linkage using the λ FixII genomic library. If we consider that the amphioxus genome will be hopefully sequenced in slightly more than one year, other approaches are necessary to establish this linkage more rapidly. These strategies may include the use of large-insert libraries, e.g. those made with cosmids, PACs (P1 phage artificial chromosomes) or, preferably, BACs (bacterial artificial chromosomes). However, the use of such libraries delays the mapping of the positive clones obtained in each step, which hinders the process of genomic walking itself. The FISH (fluorescence *in situ* hybridisation) technique on metaphasic chromosomes could then be used to test whether it is worth continuing with genomic walking experiments, by ascertaining whether an *Evx*-signal and a posterior Hox-signal colocalise in the same chromosome. Only in this case, in particular if both signals are really close to each other (which may not be easily discernible by metaphase FISH and may require interphase FISH), is chromosome walking helpful.

However, what will this linkage data tell us? If Hox and *Evx* genes are not linked in the amphioxus genome, it will certainly represent a derived condition, as these genes are linked in vertebrates and cnidarians (although the cnidarian data is restricted to a single reference [Miller and Miles, 1993] that was not further continued, and to a congress-talk around for 2 years that have not been peer-reviewed published yet [Martindale *et al.*, unpublished]). Otherwise, we will obtain an expected result. Nevertheless, it would ensure that we have reached the end of the Hox cluster and allow us to ascertain whether there are further Hox genes between *AmphiHox14* and the *Evx* genes.

CAPÍTOL IV

EVOLUCIÓ DEL GENOMA DE L'AMFIOX. UN NOU CAS DE DUPLICACIÓ: LA CLASSE EMPTY SPIRACLES (EMX) DE *B. floridae*

Minguillón, C., Ferrier, D.E.K., Cebrián, C. and Garcia-Fernández, J. (2002). *Gene duplications in the prototypical cephalochordate amphioxus*. GENE 287, 121-128.

Resum

Un dels gens *Emx* de l'amfiox (*AmphiEmxB*) va ser aïllat per hibridació creuada en un *screening* de la genoteca genòmica utilitzant com a sonda un fragment de PCR que contenia la regió més conservada de l'homeobox del gen *AmphiMox* (capítol I). Un dels clons genòmics positius per a *AmphiEmxB* va ser utilitzat com a sonda pel crivellatge d'una genoteca de cDNA d'embrions de *Branchiostoma floridae*, la qual cosa va permetre l'aïllament del cDNA complet d'*AmphiEmxB*. La posterior seqüenciació dels clons genòmics i de cDNA ens va permetre la caracterització del gen sencer. Vam trobar interessant determinar el patró d'expressió del gen al llarg del desenvolupament, per determinar si estava lligat al desenvolupament de les estructures més anteriors del cervell com ho fan els gens de la classe *Emx* en altres organismes o, tal com ocorre en ascidis, no es fa palesa aquesta funció. Però malgrat es van repetir diverses vegades hibridacions *in situ whole mount*, aquestes mai van arribar a ser prou clares i repetitives, i s'ha preferit no incloure-les en el treball.

Donat que un altre gen de la mateixa classe havia estat aïllat amb anterioritat (*AmphiEmxA*, Williams i Holland, 2001) ens trobàvem una altra vegada davant d'un cas de duplicació en el genoma de l'amfiox. Així, vam presentar els resultats d'aquest gen junt amb els que ja havien estat presentats [*Hox* (capítol II) i *Evx* (capítol III)] en un article en el que també fem una cerca bibliogràfica dels casos de duplicació que havien estat documentats fins a la data en l'amfiox.

Així, en l'article:

- (i) presentem tota una colla de gens duplicats independentment al genoma de diferents espècies d'amfiox (tres casos documentats pel nostre grup),
- (ii) observem que en principi no hi ha cap relació clara entre el tipus de gen (respecte per a quin tipus de proteïna codifica) i el seu destí després de la duplicació,
- (iii) i discutim els resultats en relació al popular estat "pre-duplicatiu" del genoma de l'amfiox.

Aportació personal al treball

Pel que fa referència al *cluster Hox* i als gens *Evx*, aquesta ja s'ha especificat als capítols II i III respectivament. En el cas del gen *AmphiEmxB*, l'autora d'aquesta Tesi va aïllar els clons genòmics, els de cDNA, i va fer l'anàlisi filogenètic. També va dur a terme la cerca bibliogràfica dels gens duplicats al genoma de l'amfiox i va participar en l'elaboració del manuscrit.

CAPÍTOL V

CARACTERITZACIÓ DE LA FAMÍLIA HAIRY EN L'AMFIOX *B. floridae*

Minguillón, C., Panopoulou, G.D. and Garcia-Fernández, J. (2002). *The amphioxus hairy genes: differential fate after duplication*. En preparació.

Resum

En aquest capítol es presenta l'aïllament i caracterització (a nivell de cDNA) dels sis representants de la família hairy de *B. floridae* als que anomenem *AmphihairyA* fins a *AmphihairyF*. Estudis filogenètics demostren que tots els gens hairy de l'amfiox queden agrupats com a trans-ortòlegs dels gens hairy de vertebrats, fet que indica que provenen de fenòmens de duplicació independents. A més, amplifiquem la regió genòmica corresponent al putatiu intró 2 més exons flanquejants de dues seqüències, probablement pseudogens, que malgrat tenir la més elevada similitud a representants de la família hairy són molt divergents, a les que anomenem *Amphihairy-like1* i *2*. Experiments d'RT-PCR demostren que dos dels sis gens, *AmphihairyE* i *AmphihairyF*, no s'expressen a nivells detectables durant l'embriogènesi i tampoc en l'adult. En canvi, l'estudi de l'expressió dels altres quatre gens, *AmphihairyA*, *B*, *C* i *D*, per hibridació in situ *whole mount*, mostra que tenen un patró força específic i complementari. Breument, el gen *AmphihairyA* s'expressa en teixits endodèrmics i en la part més posterior del sistema nerviós central; *AmphihairyB* en la part més anterior del sistema nerviós central i en somites; *AmphihairyC* en la part mitja del sistema nerviós central, en l'endoderm i els somites; i *AmphihairyD* en en la part mitja del sistema nerviós central, en l'endoderm i la notocorda. Així, la suma dels seus patrons d'expressió recorda, *a grosso modo*, l'expressió de l'únic gen hairy de ratolí *HES1*. Aquest fet ens fa proposar que la família hairy ha patit un fenomen de subfuncionalització (degeneració diferencial de les regions *cis* reguladores) després de les duplicacions en el llinatge de l'amfiox, la qual cosa assegura el seu manteniment al genoma de l'amfiox com a gens duplicats.

L'estudi de l'expressió dels gens hairy durant el desenvolupament de l'amfiox ens permet observar que l'asimetria del procés de la somitogènesi és present des de la formació del primer parell de somites (l'expressió del gen *AmphiMox* ens havia permès observar aquesta asimetria tan sols a partir del cinquè parell de somites; capítol I). També observem que malgrat els quatre primers somites apareixen de cop, la seva determinació és progressiva (en sentit antero-posterior). L'estudi comparatiu de l'expressió dels gens *Amphihairy* respecte els gens hairy de vertebrats i protostoms ens permet concloure que moltes funcions que els gens hairy duen a terme en vertebrats i no en protòstoms, no representen innovacions dels vertebrats, sinó que ja eren dutes a terme per l'únic gen hairy de l'ancestre dels cefalocordats més vertebrats. També, que el comportament cíclic que exhibeixen els gens hairy d'amniotes representa una sinapomorfia d'aquest grup.

Aportació personal al treball

Tota la recerca experimental que es presenta en aquest treball ha estat realitzada per l'autora d'aquesta Tesi Doctoral, a excepció del clonatge i seqüenciació d'*AmphihairyF* que va ser realitzat per la segona autora, Georgia Panopoulou. Ens va enviar el cDNA corresponent i vam analitzar la seva expressió per hibridació *in situ whole mount* i RT-PCR.

Discussió general

Els resultats presentats en aquesta Tesi Doctoral es discuteixen als articles inclosos als capítols de Resultats. Així doncs, aquesta secció és tan sols una discussió general en la que s'intenta relacionar els diferents gens que s'han estudiat i fer una explicació global del procés que ha estat analitzat en més profunditat: la somitogènesi. Per finalitzar es contemplen algunes dificultats que poden sorgir de l'utilització de l'amfiox com a sistema model al laboratori.

En concret, aquesta secció s'ha dividit en quatre apartats:

1. Evolució del grup *extended Hox* de gens amb homeobox

Tres de les quatre classes de gens amb homeobox que han estat objecte d'estudi al llarg d'aquest treball pertanyen al grup *extended Hox* (Hox, Evx i Mox). Mitjançant l'anàlisi filogenètic d'aquestes classes es proposen dos escenaris evolutius que podrien explicar l'origen i evolució dels gens *extended Hox* i ParaHox.

2. La somitogènesi de l'amfiox i les innovacions dels vertebrats

L'estudi del gen Mox i de la família hairy de l'amfiox ens ha permès l'estudi acurat del procés de somitogènesi en cefalocordats, així com l'estudi de determinades sinapomorfies de vertebrats referents al mateix procés. Així, en aquesta discussió general es fa una explicació global de la somitogènesi en l'amfiox posant especial ènfasi a les diferències observades entre l'amfiox i els vertebrats.

3. Precaucions a l'hora d'interpretar els resultats obtinguts de l'estudi dels gens de l'amfiox

En aquest apartat s'adverteix d'alguns aspectes que han de tenir-se en consideració i que sovint es passen per alt a l'hora d'interpretar els resultats obtinguts de l'estudi dels gens de l'amfiox.

4. Problemes amb l'ús de l'amfiox com a sistema model

La curta durada de l'època reproductiva de l'amfiox i l'enorme despesa que representa la seva recol·lecció als Estats Units representen algunes de les limitacions que es deriven de l'ús de *B. floridae* com a organisme model al laboratori. En aquesta darrera secció es presenten algunes alternatives que podrien solucionar les anteriorment esmentades limitacions.

General discussion

The main issues of this Thesis have been discussed in the Results sections (included within the manuscripts). This last part is merely a general discussion aimed at linking some of the genes that have been the subject of study in this Thesis. In addition, the use of amphioxus as a model system to study vertebrate innovations is assessed. The section is organised as follows:

1. Possible evolutionary scenarios explaining the relationship between the extended Hox group of homeobox-containing genes (Hox, Evx and Mox) and the ParaHox genes, on the basis of phylogenetic trees and linkage data in a range of organisms.
2. Regarding somitogenesis, the developmental process that has been deeply analysed through the study of Mox and hairy genes, we have found some cases of co-option that may be linked to the acquisition of vertebrate novelties. We have also shed some light on the amphioxus somitogenesis process itself.
3. Cautions to be kept in mind when interpreting the results obtained from the study of amphioxus genes.
4. Limitations inherent in the use of amphioxus as a model system in the laboratory.

1. EVOLUTION OF THE EXTENDED HOX GROUP OF HOMEBOX GENES

The homeobox-containing genes that have been the subject of study in this Thesis belong to the Antp superclass of homeobox genes (Bürglin, 1995). All are easily classifiable into a specific class of genes, namely Hox, Evx, Mox and Emx. However, there is not a widely accepted consensus view about the relationship between these classes, especially among the members of the extended Hox group, which is formed by the clustered Hox genes and the Mox and the Evx classes (Pollard and Holland, 2000). These genes have been unified into the extended Hox group, owing to their linked disposition in the genome of certain organisms, e.g. Evx genes are closely linked to the 5' end of the Hox gene cluster in most vertebrates and in a cnidarian species (Dush and Martin, 1992; Miller and Miles, 1993; Amores *et al.*, 1998). Similarly, Mox genes map close to the opposite extreme of the HoxA and HoxB cluster in the human genome (<http://genome.ucsc.edu>). These linkage data prompted some authors to suggest that Mox and Evx genes originated during the tandem duplication events that produced the ancestral Hox cluster gene (Pollard and Holland, 2000). In a phylogenetic tree, Hox genes alone do not form a monophyletic clade, but a clade containing both Hox and ParaHox genes. Evx genes fall basal to the clade of Hox and ParaHox genes (Kourakis and Martindale, 2000), and Mox genes are thought to fall in a position closely related to the Hox and ParaHox genes (Balavoine, 1996). The Mox gene has often been referred to as a ParaHox gene, more precisely as the missing ParaHox gene related to the central group (PG4 to PG8) of Hox genes (Balavoine, 1996; Gauchat, 2000).

Moreover, the views (Hox- or ParaHox-related) of the evolutionary relationship between the Mox and the Hox and ParaHox genes are contradictory. If Mox genes are derived from the tandem duplication of a particular Hox gene (and thus linked to the Hox gene cluster), they cannot be considered ParaHox genes, because these have not evolved from the duplication of a

Hox gene. Besides, Hox and ParaHox gene clusters are regarded as sister clusters, as they may have resulted from the duplication of the same ProtoHox cluster, which gave rise to both Hox and ParaHox clusters (Brooke *et al.*, 1998) (Figure 1). *Mox* is considered the missing central gene in the ParaHox gene cluster (question marks in Fig. 1; Balavoine, 1996). The hypothesis that *Evx* is a sister group of Hox plus ParaHox genes, based on its basal position to the Hox and ParaHox monophyletic group, disagrees with its origin by tandem duplication of a Hox gene. The basal position rather supports that an *Evx* gene was originally linked to the *UrProtoHox* gene (i. e. the single gene, which after tandem duplication gave rise to the ProtoHox cluster; Fig. 1).

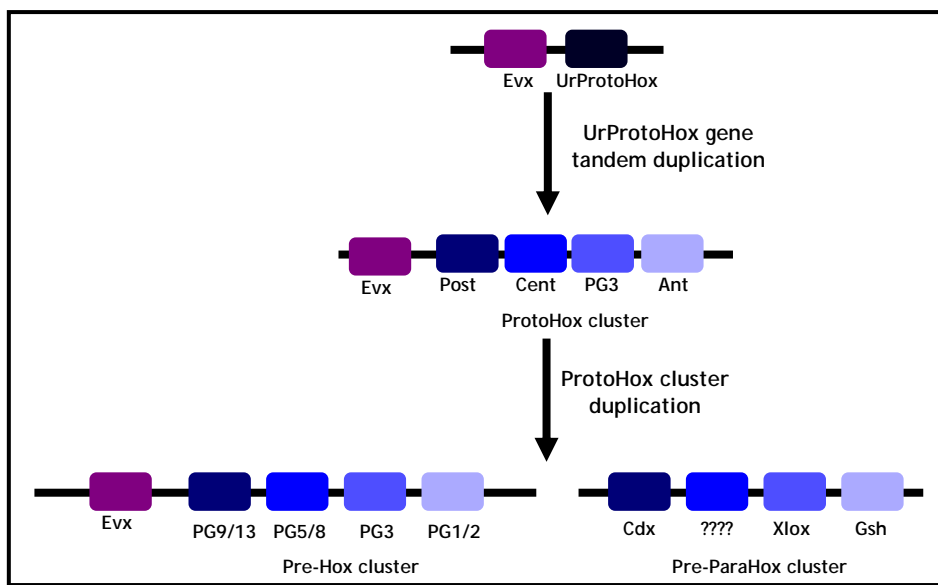


Figure 1. Origin of the twin clusters by duplication of an ancestral ProtoHox gene cluster. A single ProtoHox gene cluster composed of a maximum of four genes (a posterior one, a central one, a group 3 gene and an anterior one) gives rise after duplication to two clusters, namely the Hox and ParaHox gene clusters. Each of the resultant clusters has four members (one of each group present in the ProtoHox gene cluster). Subsequent evolution by tandem duplication in the case of the Hox cluster or central gene loss in the case of the ParaHox cluster gives rise to the actual Hox and ParaHox gene clusters with 13 and 3 genes in mammals, respectively. Modified from Brooke *et al.*, 1998.

All these disagreeing points of view led us to construct phylogenetic trees and search for data backing up the proposed evolutionary relationships between the extended Hox group and ParaHox genes of homeobox-containing genes, and discuss scenarios that may not have been considered yet. I would like to highlight that the nature of the homeodomain proteins used for the analysis hinders the construction of such trees. As mentioned in the Introduction, the homeodomain is highly conserved, even among members of distinct classes of homeobox genes, and outside the homeodomain (at least for the classes considered here), the similarities between related proteins are restricted to scattered residues. Hence, the use of the whole protein sequence to reconstruct phylogenies may disrupt the alignment and the use of only the conserved homeodomain may not be enough informative. Thus, it is not easy to obtain a robust tree and each type of analysis (phylogeny method and part of the protein used, among others) may yield a distinct tree. Therefore, I would not undoubtedly sign for any of these scenarios. However, I would reject the hypothesis that *Mox* gene is the central ParaHox gene, as in none of the trees built is a sister group relationship between *Mox* and the central Hox genes found. Numerous phylogenetic trees were constructed (using the whole protein, the homeodomain plus

some flanking amino acids or only the homeodomain sequences, following various phylogenetic methods). Nevertheless, as mentioned above, owing to the nature of the molecules under analysis, the trees obtained did not have particularly high bootstrap values. Accordingly, I shall not include all these phylogenetic trees, but only one of them and two consensus trees showing the topology that has been obtained in most of the trees. The interpretation of these three trees led us to consider two novel scenarios for the evolutionary relationship among the extended Hox group genes (at least for the relationship between the Mox and the other classes).

Figure 2 shows one of the most frequent unrooted phylogenetic trees, regardless of the part of the molecule used. In this particular tree, the homeodomain plus ten flanking residues in each side were subjected to neighbour-joining analysis using the ClustalX program. The robustness of the tree was assessed by bootstrap analysis (1000 replicates). Note:

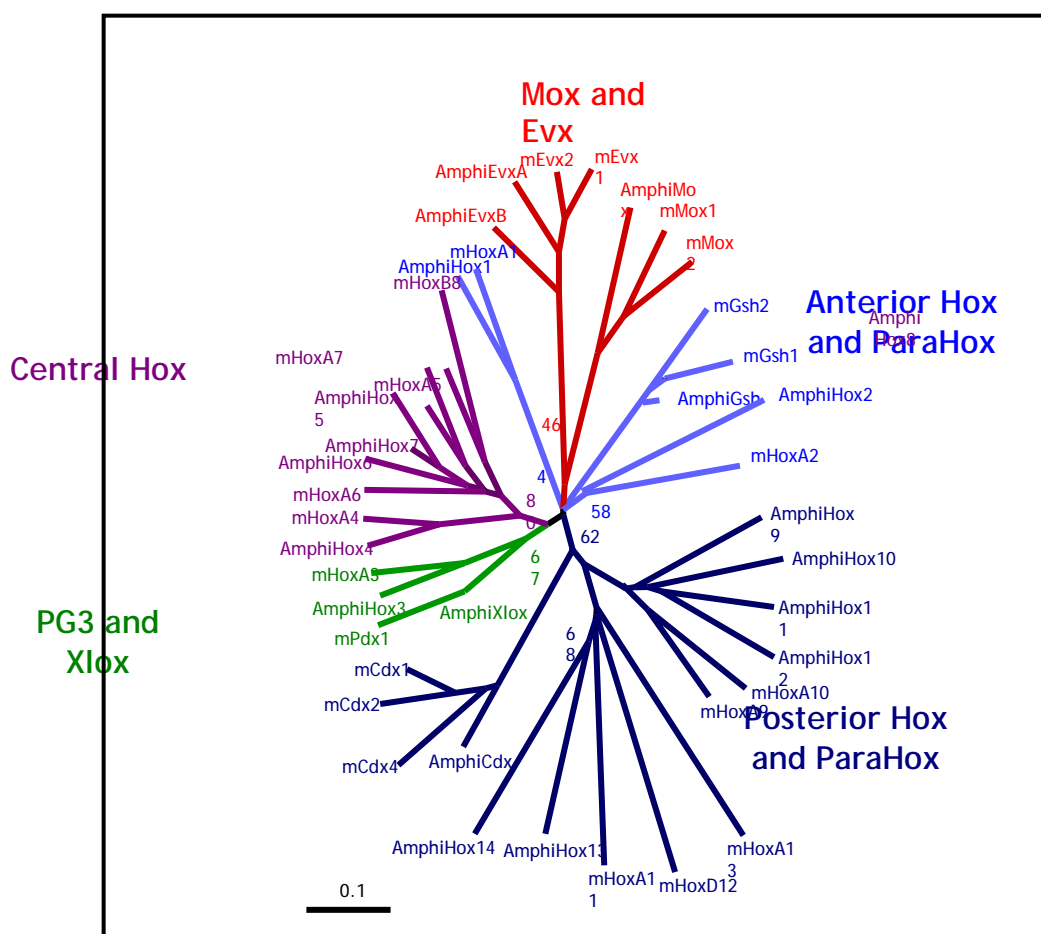


Figure 2. Neighbor joining phylogenetic tree of the ParaHox, Mox and Evx genes of both *B. floridae* and mouse, and the amphioxus Hox and selected mouse Hox genes. Only selected bootstrap values (discussed in text) are shown. Five groups of major order are shown by a color code: red, Mox plus Evx; dark blue, posterior group of Hox plus ParaHox genes; light blue, anterior group of Hox plus ParaHox genes; green, group 3 of Hox plus ParaHox genes; purple, central group of Hox genes.

5. The grouping of Mox and Evx class genes (red in Fig. 2). Although the bootstrap value that supports this relationship is only 46%, lower values have been reported elsewhere. Thus, not only Evx but also Mox genes are equally related to both Hox and ParaHox genes in this analysis.

6. The lack of a ParaHox-type central gene. Only Hox genes are grouped within the central group (80%). Hence, Mox is not the central ParaHox gene according to this analysis (or in any of the trees obtained).
7. The consistent grouping of the already proposed relationship for the Hox and ParaHox genes, i. e. *Cdx* is the posterior ParaHox gene more closely related to the posterior group of Hox genes (62%); *Xlox/Pdx1* is the ParaHox gene more closely related to group 3 of Hox genes (67%) and *Gsh* is the ParaHox gene more closely related to the anterior group of Hox genes (58%). However, Gsh genes seem to be more closely related to PG2 than to PG1 of Hox genes, although the bootstrap value that supports the divergence of the PG1 group from the other two groups is negligible (4%). Moreover, the sister group relationship between these anterior Hox and ParaHox genes is the lowest found in other analysis (Brooke *et al.*, 1998).
8. The sister group relationship of the amphioxus genes *AmphiHox13* and *14* (68%) suggests that they represent specific duplicate genes in the *B. floridae* genome (chapter II).

The Evx and Mox classes were closely related in all unrooted trees. When the trees were rooted with a member of the NKL group of homeobox genes (e. g. the Emx class), two major tree topologies were obtained (Fig. 3). Again, none of the trees revealed a close relationship between Mox and the central Hox genes. Both consensus trees show similar basal positions for the Evx and Mox classes with respect to the monophyletic Hox and ParaHox group, but they differ in the relationship between these two basal clades (Mox and Evx). In some trees, they appear as sister groups (Fig. 3A), whereas in others, the Mox class is basal to the Hox and ParaHox genes and the Evx class is located as the sister group of the former group (i. e. Mox, Hox and ParaHox) (Fig. 3B). A possible evolutionary scenario is discernible for each of the above consensus trees. The first hypothesis has been envisioned by Kourakis and Martindale (2000), who nonetheless did not identify the Mox gene.

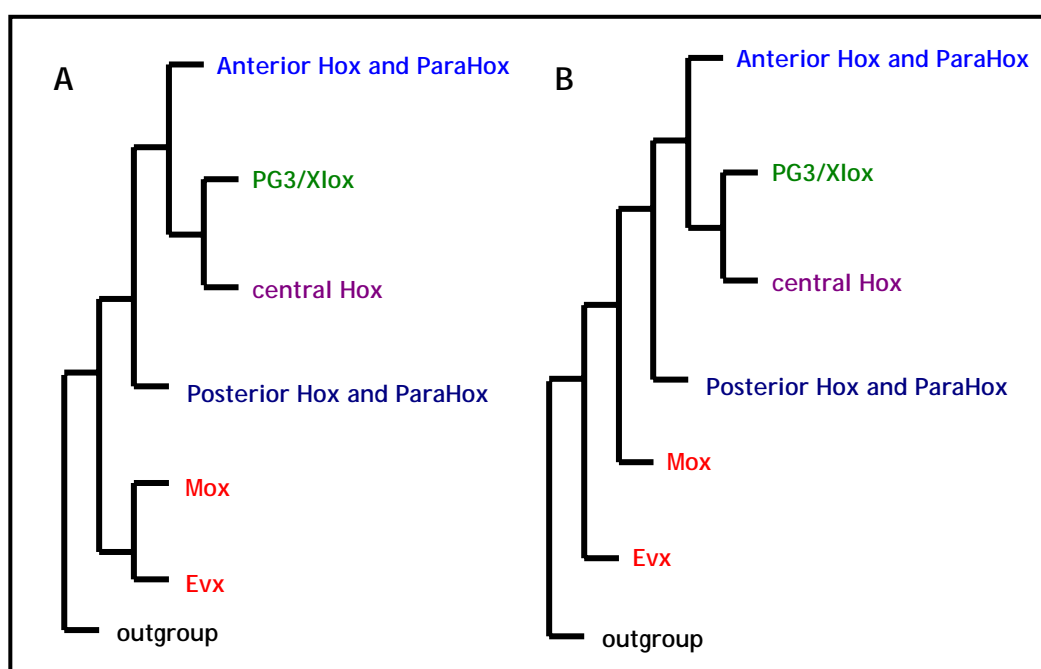


Figure 3. Consensus phylogenetic trees showing the relationship among the members of the Extended-Hox and ParaHox group of homeobox genes when rooted with a NKL-type gene. A) Sister group relationship between Mox and Evx classes. B) Basal position of Mox genes with respect to Hox and ParaHox genes, and that of Evx genes with respect to all the former. The colour code is the same as in figure 2.

The first scenario proposed (Fig. 4) is based on the *Evx* and *Mox* sister group relationship (Fig. 3A), which suggests that they result from the duplication of a single gene (here referred to as *ProtoMoeve* gene). The basal position of the *Evx*/*Mox* clade with respect to the *Hox*/*ParaHox* group suggests that the *ProtoMoeve* gene was initially linked to the single *UrProtoHox* gene (1 in Fig. 4). After the tandem duplication of the *UrProtoHox* gene, the *ProtoMoeve* gene remains linked to the *ProtoHox* cluster, both forming the *ProtoMoeveHox* cluster (2 in Fig. 4). If this complex is then tandemly duplicated giving rise to the twin *Pre-extHox*/*ParaHox* cluster, it will contain four *Hox* and *ParaHox* genes (from the former *ProtoHox* cluster), one *Mox* and one *Evx* gene (from the duplication of the *ProtoMoeve* gene) (3 in Fig. 4). If this twin cluster is broken between the *Mox* and the *ParaHox* genes, two separate clusters appear: the *Pre-ParaHox* and the *Pre-extended Hox* (with *Mox*, *Evx* and *Pre-Hox* genes) which follow their independent evolution (4 in Fig. 4). The former loses the central gene, whereas the latter expands by internal tandem

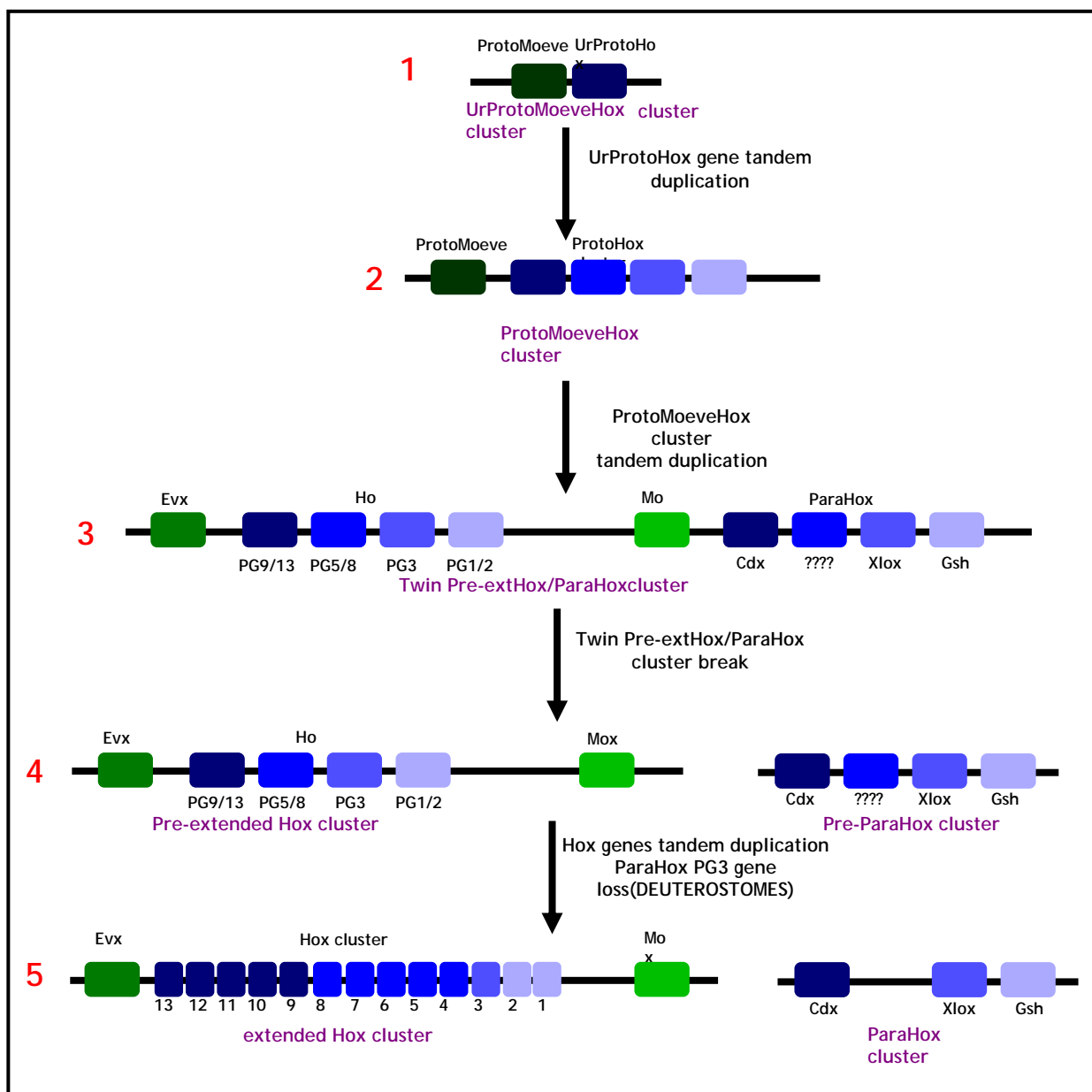


Figure 4. First scenario proposed for the origin and evolution of the extended Hox and ParaHox genes based on the sister group relationship between *Mox* and *Evx* genes.

duplication (5 in Fig. 4). This reasoning may not only explain the phylogenetic relationship among the genes, but also reconcile the linkage (*Evx* 5' to the Hox cluster and *Mox* 3' to the Hox cluster [Fig. 5]) with phylogenetic data. Note that the order of the twin Pre-extHox/ParaHox cluster split and the further evolution of the Hox and ParaHox clusters can be inverse (step 5 before step 4 in Fig. 4), and that the ProtoMoeveHox cluster may undergo a non-tandem duplication (step 2 to 3 in Fig. 4). Nevertheless, this would require further steps to explain the *Mox* gene linkage to the Hox cluster genes.

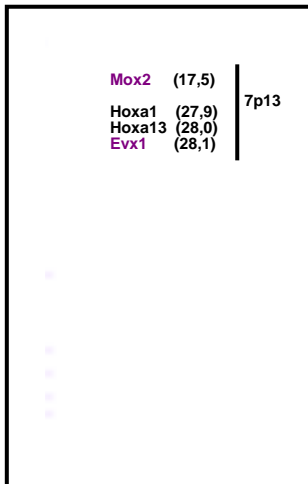


Figure 5. Chromosomal locations of *Hox*, *Mox* and *Evx* genes in the short arm of human chromosome 7.

The second scenario (Fig. 6) is based on the consensus tree shown in figure 3B. It differs from the above scenario in that a single gene (the *ProtoMoeve* gene in the first scenario depicted in Fig. 4) gives rise to the *Mox* and *Evx* genes. Here, *Evx* and *Mox* classes are also basal with respect to Hox and ParaHox genes, but not sufficiently related to have arisen by the duplication of the same gene. On the contrary, the closer relationship between *Mox* and Hox and ParaHox suggests that they come from the duplication of a single gene (here referred as the *UrProtoMoHox* gene) that was linked to the *Evx* gene, both forming the *UreveProtoMoHox* cluster (1 in Fig. 6). The tandem duplication of the *UrProtoMoHox* gene gives rise to both the *Mox* and *UrProtoHox* genes (2 in Fig. 6). The further tandem duplication of the *UrProtoHox* gene yields the ProtoHox cluster genes that are between *Evx* and *Mox* genes (3 in Fig. 6). If a further non-tandem duplication restricted to the ProtoHox cluster genes occurs, we end up with the Pre-extended Hox and Pre-ParaHox clusters in distinct chromosomes (4 in Fig. 6). To achieve the current composition of these two complexes, we need the central gene loss in the ParaHox cluster and internal tandem duplications in the Hox one (step 5, as required in the first scenario proposed in Fig. 4).

As both evolutionary schemes involve the same steps and reconcile the linkage with phylogenetic data, they are in principle equally probable. Nevertheless, a higher number of trees support the second one (about 70%). As the *Mox* linkage to the Hox cluster has only been reported for mammals, it may represent the derived condition. The linkage of the *Mox* to ParaHox cluster genes would strongly support the first scenario (Fig. 4). Figure 4 has been drawn to reconcile the linkage with phylogenetic data. Nevertheless, if the *Mox*/ParaHox linkage holds true and so mammals represent the derived condition, the Twin Pre-extHox/ParaHox cluster split between *Mox* and Hox genes. We tested this hypothesis in the amphioxus genome by hybridising a contig of bacteriophages and cosmids that covers the amphioxus ParaHox cluster genomic region

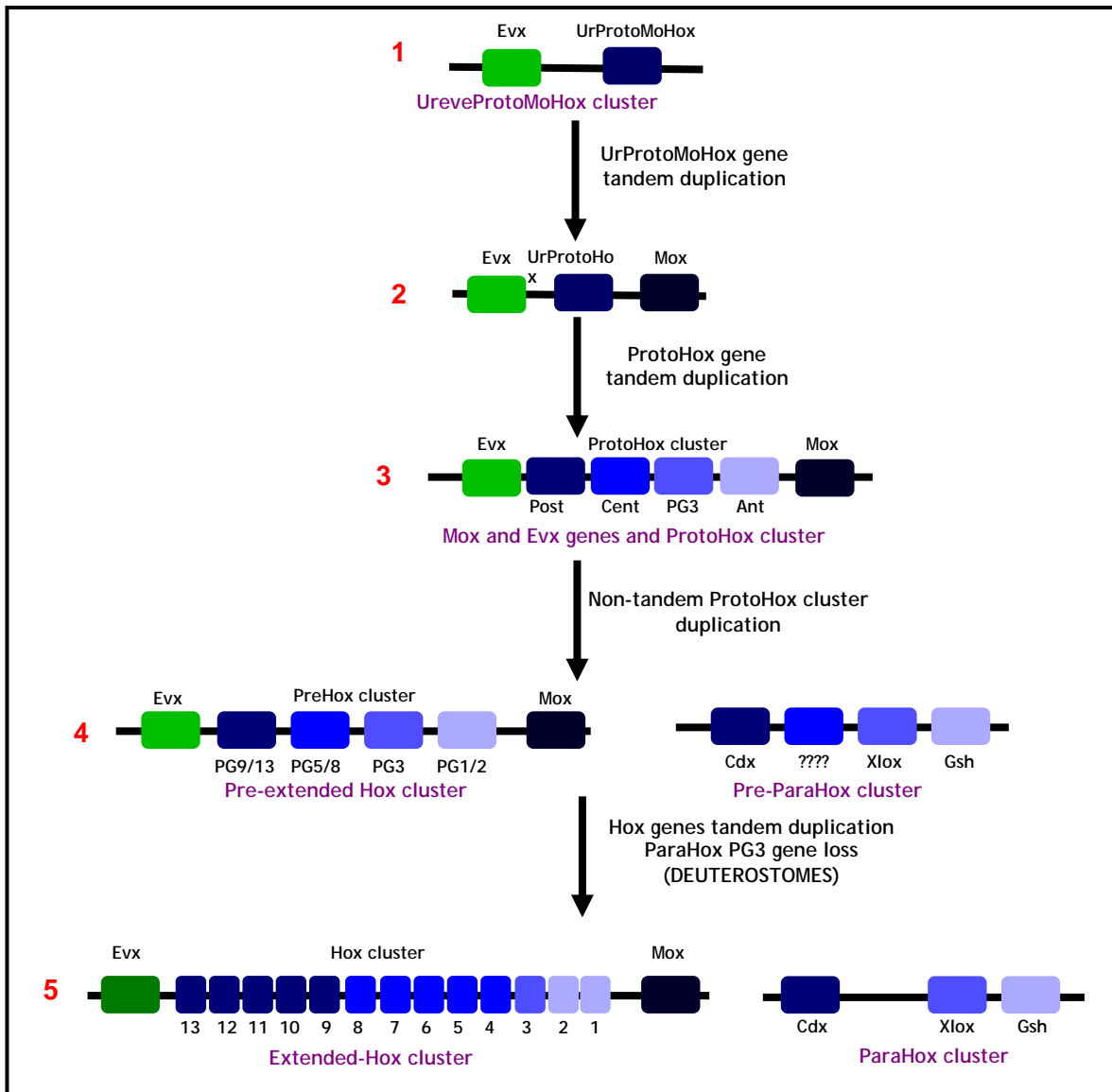


Figure 6. Second proposed scenario for the origin and evolution of extended Hox and ParaHox group genes based on the consensus tree shown in Fig. 3B.

(Brooke *et al.*, 1998) with the *AmphiMox* genomic sequence (chapter I, Results). Although no hybridisation was detected, we cannot rule out this linkage, since we covered only the region containing the ParaHox genes.

2. AMPHIOXUS SOMITOGENESIS AND VERTEBRATE NOVELTIES

The main developmental process that has been studied in this Thesis is somitogenesis. Somites are the most obvious sign of segmentation in vertebrate and cephalochordate embryos (Saga and Takeda, 2001). As they are a synapomorphy of cephalochordates and vertebrates, the use of amphioxus to study vertebrate innovations during somitogenesis is unavoidable. There are similarities but also differences between the vertebrate and amphioxus processes of somitogenesis. In particular, the comparative study of these differences may eventually reveal some vertebrate novelties that originated from the further differentiation of somites into somitic derivatives, such as the vertebrae. This process is radically different between amphioxus and vertebrates. After segmentation from the presomitic mesoderm (PSM), vertebrate somites undergo complex transformations, like epithelial to mesenchyme transition and delamination of cells from the dermomyotome (Pourquié, 2001). In contrast, amphioxus is a predominantly epithelial animal (Whittaker, 1997). Early somites arise as evaginations that bud off from the dorso-lateral wall of the archenteron, resulting in a single-layered epithelium surrounding a cavity, the myocoel. The somites of amphioxus are subdivided into distinct regions, although there is no histological evidence for the sclerotome and no cells migrate away from it (Shimeld and Holland, 2000). Besides, the large medial compartment of each somite is the myotome, which retains its epithelial organisation throughout development. All myotomal cells differentiate in place and become the striated muscle cells that constitute the segmental muscle blocks along the body length (Holland *et al.*, 1995a). Strikingly, certain developmental genes involved in the further differentiation programs that vertebrate but not cephalochordate somites undergo after segmentation from the PSM, are expressed in the vertebrate but not in the amphioxus somites. For instance, *AmphiMox* was not expressed in the amphioxus segmented somites (chapter I, Results), whereas vertebrate *Mox* genes are expressed during somite differentiation (Candia *et al.*, 1992). Genes other than *Mox* behave likewise, since they are expressed in the formed vertebrate somites but not in amphioxus somites [e.g. *Amphisnail* (Langeland *et al.*, 1998) and *AmphiPax1/9* (Holland *et al.*, 1995b)]. This may reflect the simpler structure of cephalochordate somites compared to the higher complexity of vertebrate somites and their derivatives. Hence, the co-option or the maintenance of expression within the somites of such genes after duplication in the vertebrate stem lineage may have been instrumental in the acquisition of some vertebrate innovations such as the vertebrae.

To gain more insight into the process of somitogenesis, we isolated and studied the expression of the *B. floridae* hairy family. The amniote hairy genes have an oscillating expression pattern in the vertebrate PSM; their expression cycles with a periodicity that corresponds to the formation of one somite (Jouve *et al.*, 2000). In the chicken, the onset of the dynamic expression of hairy genes correlates with the ingression of the paraxial mesoderm territory from the epiblast into the primitive streak. Hence, the number of oscillations experienced by somitic cells is correlated with their position along the anterior-posterior axis (Jouve *et al.*, 2002). However, no hairy family gene in non-amniotes behaves similarly to its homologue in amniotes as far as cycling expression in the PSM is concerned (Leve *et al.*, 2001; Pasini *et al.*, 2001; Davis *et al.*, 2001). Thus, the absence of hairy gene cycling activity in lower vertebrates may represent a secondary loss or else an amniote novelty. The lack of cycling behaviour in the four Amphihairy

genes analysed by *in situ* whole mount hybridisation suggested that this cycling behaviour is an amniote novelty. What are the differences in somitogenesis between amniotes and lower vertebrates and cephalochordates? The coelomic cavities of the most anterior somites of amphioxus are formed by enterocoely through the outpocketing of the dorso-lateral wall of the archenteron. This early paraxial mesoderm is formed during gastrulation and used throughout early somitogenesis. The number of somite pairs formed in this phase is 8 in *B. floridae* (Holland *et al.*, 1997). More caudal somites are formed by schizocoely (more similarly to vertebrates), by a splitting process of solid blocks from the proliferating tail bud. Early amphioxus somitogenesis is reminiscent of the formation of the trunk anterior somites of fish and amphibians, from the sequential segmentation of a pre-existing territory that involutes during gastrulation, and not from the division of stem cells in the node/primitive streak of amniotes (Pourquié, 2001). A novel function acquired by higher vertebrate hairy genes may have been instrumental in the more complex process of segmentation that requires stem cell division. Nevertheless, the somites originated from the tail bud in both amphioxus and amniotes do require stem cell division, and amphioxus hairy genes may still cycle during their formation. However, the lack of intervening PSM between the tail bud and the originating somites makes this hypothesis highly improbable or untraceable (see below).

None of the genes with cyclic expression in higher vertebrates behaves like that in *Xenopus*. Note that somitogenesis is also inherently asymmetrical in the frog, since the right side of the embryo temporally advanced in segmentation (Davis *et al.*, 2001). Thus, there may be no way to co-ordinate an asymmetrical cycling pattern, and so these organisms (amphioxus and the frog) use other strategies, or else somitogenesis in both amphioxus and *Xenopus* may not require such an accurate control as in amniotes.

There are some cycling genes other than hairy family members in the zebrafish, namely the Notch ligand *DeltaC* and the E(spl) family members *her1* and *her7* (Jiang *et al.*, 2000; Holey *et al.*, 2000; cited in Leve *et al.*, 2001). An E(spl) family gene, *Hes7*, also cycles in the mouse PSM (Bessho *et al.*, 2002). Therefore, other genes related to the Notch signalling pathway or other genes of the E(spl) family may cycle in the amphioxus PSM. Nevertheless, the extent of the amphioxus PSM makes the cycling of any gene highly improbable. Amphioxus somitogenesis can be divided in two phases. In early somitogenesis, the somites originate by the budding off of the dorso-lateral walls of the archenteron and coeloms are formed by enterocoely (Fig. 7A). The first eight pairs of muscular somites are generated during this early phase from the paraxial mesoderm created during gastrulation (Holland *et al.*, 1997). In contrast, during the second phase, somites arise directly from the proliferative tail bud by schizocoely (Schubert *et al.*, 2001) (Fig. 7B). In the early phase, the PSM is represented in the dorso-lateral walls of the

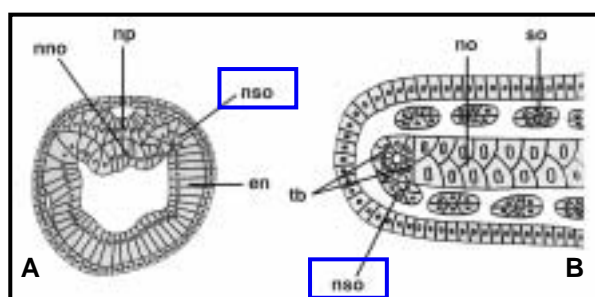


Figure 7. Amphioxus somitogenesis. During the early phase somites arise by the budding off from the dorso-lateral walls of the inner endodermal layer (A), whereas in the late phase they originate directly from the proliferating tail bud (B). Abbreviations used are: en, endoderm; nno, nascent notochord; no, notochord; nso, nascent somite; np, neural plate; tb, tail bud; so, somite. Nascent somites are in blue boxes.

archenteron along eight pairs of somites, and there is no intervening PSM between the tail bud and the nascent somites during the late phase. Thus, if any gene has a cycling behaviour, it would be visible only during the early phase of amphioxus somitogenesis. Nevertheless, the first eight muscular somites are formed very quickly, within a few hours, and so there is not much time for cycling gene transcription regulation. Moreover, the first four somites bud off almost simultaneously from the archenteron (Conklin, 1932).

The study of the *Mox* and hairy genes in *B. floridae* also shed some light into the process of amphioxus somitogenesis itself. Specific aspects of lancelet somitogenesis have been widely discussed among classical morphologists. An example is the inherent asymmetric nature of amphioxus somitogenesis: Conklin (1932) and Hatschek (1892) agreed that somites are roughly symmetrical in position until the 7- or 8-somite stage, but Cerfontaine (1906) claimed that this asymmetry is present from the first pair, with a slight delay in the development of the right side with respect to that of the left side. *AmphiMox* gene expression data led us to conclude that amphioxus somites are asymmetrical from somite number five onwards (Fig. 8A). Left side somites appear first and are thus slightly rostrally located with respect to their right side counterparts (chapter I, Results). The study of the amphioxus hairy genes revealed that cephalochordate somitogenesis was asymmetrical to a higher extent. The asymmetric expression of *AmphihairyB* as stripes prefiguring the presumptive somites in late gastrulae, showed that this asymmetry is present from the very beginning, even before the somites are morphologically discernible (Fig. 8B). Hence, amphioxus left-side somites are slightly more advanced in development than their right-side counterparts, which is noticeable in their more rostral location.

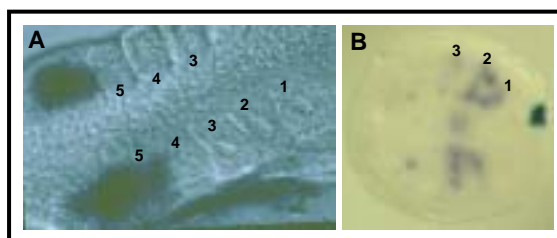


Figure 8. Asymmetric amphioxus somitogenesis, as revealed by *AmphiMox* and *AmphihairyB* gene expression in neurula and gastrula stages, respectively. Anterior is toward the right and numbers refer to the somite or presumptive somite number. **(A)** *AmphiMox* mRNA is still present in the 5th right-side somite, whereas it is not detected in its left-side counterpart. **(B)** *AmphihairyB* expression is seen asymmetrical (the third stripe is observed on the left but not on the right side) even before any somite is morphologically visible.

3. CAUTIONS WITH THE INTERPRETATION OF AMPHIOXUS GENE DATA

Amphioxus has been widely used to understand how vertebrates evolved from an invertebrate ancestor, in particular to elucidate how they acquired some of the morphological innovations they present nowadays. Briefly, by comparison of the expression patterns of a specific gene(s) in vertebrates *versus* the (*a priori*) single orthologous in amphioxus, it is assumed that:

- The amphioxus and vertebrate body regions that express these genes represent homologous structures. This structure was thus already present in their common ancestor, and the particular ancestral gene was also expressed there.
- The expression domains, which are unique to amphioxus, are indicative of cephalochordate specific characters.
- The expression domains that are unique to vertebrates may be the result of their co-option in "vertebrate-invented" functions, which has been instrumental in the emergence of new morphological characters.

For instance, the results presented here have shed some light into the appearance of the midbrain/hindbrain boundary region, with organising properties in vertebrates (chapter III, Results), and the emergence of vertebrae (chapter I, Results). Nevertheless, although this approximation seems reasonable, we should bear in mind the following concepts or facts:

- I. the concept of ancestral species
- II. the evolution of the amphioxus genome
- III. the concept of homology

3.1. The concept of ancestral species

Although it is tempting to regard an extant species as the "lineal ancestor" of a particular group, such a case is actually impossible. Hence, although cephalochordates are thought to resemble the ancestor of vertebrates, any extant species is the result of the mixture of ancestral and specialised characters, in this case amphioxus oddities. The degree of approximation to an inferred ancestral form can be interpreted only if primitive features are distinguished from the products of several evolutionary trends, including the loss of original characters, new specialisations and the effects of convergence to other candidate relatives (Presley *et al.*, 1996).

3.2. The evolution of the amphioxus genome

Several approaches support that amphioxus is both morphologically and genomically similar to the common ancestor of cephalochordates and vertebrates (Conway Morris, 1982; Holland *et al.*, 1994). The current view of the vertebrate genome as the result of major events of gene duplication and co-option is partly supported by the finding that the genome of the cephalochordate amphioxus, the closest living invertebrate relative to vertebrates, contains single members of most of the multigenic families of vertebrates (Holland and Garcia-Fernández, 1996). The amphioxus genome is thus regarded as the prototypical and pre-duplicative

vertebrate-like genome, and the expression pattern of an amphioxus gene is believed to reflect the ancestral chordate function of such family. However, cephalochordates and vertebrates have independently evolved for more than 500 million years (Cañestro *et al.*, 2002). Hence, it is plausible that the norm of "singularity" of the amphioxus genome that fits most of its genes does not fit others. In this Thesis, four of five families have expanded (from 2 to 6) in the *B. floridae* genome (chapters II, III, IV and V, Results). Nevertheless, the location of amphioxus genes as a sister group of the vertebrate gene family has to be maintained, i.e. all the amphioxus genes should be grouped in a phylogenetic tree as trans-orthologues of vertebrate genes, without a 1:1 orthology relationship (Fig. 9).

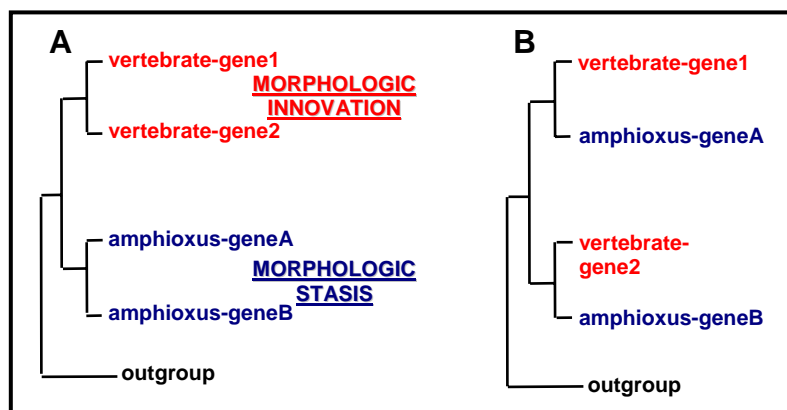


Figure 9. Hypothetic relationships between amphioxus and vertebrate genes. A) Trans-orthology relationship. Note that whereas gene duplication has permitted the acquisition of novelties in vertebrates, it has not had such an effect in cephalochordates (here denoted as morphologic stasis). B) Erroneous tree showing a 1:1 orthology relationship between an amphioxus and a vertebrate gene.

These cases of duplication request for caution when deducing the evolutionary history of a gene family in chordates on the basis of the singularity of the amphioxus genome. We would like to highlight the extreme case, that of amphioxus hairy genes. Our results suggest that during or after duplication, the *cis* regulatory regions of the multiple hairy genes degenerated by differential loss of specific *cis* elements that were not, however, maintained in any of the other duplicate genes, thus ensuring their maintenance as duplicate genes in the genome. They could not degenerate to pseudogenes, as none of the copies retained all regulatory regions, and they were forced to complement for the hairy pleiotropic roles. Similar examples of the DDC model (for Duplication/Degeneration/Complementation Model) have been reported for zebrafish and maize genes after polyploidization (Force *et al.*, 1999), but the finding was unexpected for the otherwise prototypical amphioxus genome. Thus, when studying the pan chordate gene function of a given gene family using cephalochordates as a model system, if additional members of the family are lost, only partial patterns of expression can be found, and the evolutionary scenarios deduced may be misleading. Only the full sequencing of the amphioxus genome or exhaustive searches by various approaches ensure that the evolutionary insights gained from cephalochordate genes are reliable and not blurred by lost genes or fractions of expression patterns. Although amphioxus may resemble the ancestor of vertebrates, but it is not their ancestor, but only its closest living relative, a privileged position that does not involve the freezing of its genome.

3.3. The concept of homology

A central concept in Comparative Biology is that of homology. The characters of distinct organisms are homologous when they share a common origin or, in other words, when they correspond to the same character of their common ancestor. The application of Molecular Biology techniques has broadened the scope of the definition of homology: homologous genes, homologous processes, etc. We should thus differentiate the types of homology, since there is not always a simple correlation between homologies at different levels. For instance, homologous genes can be involved in the development of non-homologous structures, or *vice versa* (reviewed in Abouheif *et al.*, 1997). Although the comparison of gene expression patterns has been widely used as an indicator of structural homology (the structures in which orthologous genes are expressed represent homologous structures in these reports), the above reasoning suggests that this does not always hold true. Therefore, molecular homology (in terms of gene expression) should be regarded merely as an additional indicator of the homology between two structures (Holland and Holland, 1999). Nonetheless, gene expression data are safer indicators of homology when the overall body plans of the animals under comparison are relatively similar, which is the case of cephalochordates and vertebrates.

The prototypical vertebrate-like body plan of amphioxus and its preduplicative vertebrate-like genome have fashioned the use of amphioxus as a reference for establishing the ancestral role of a given gene (gene family) within chordates and gene expression data are considered valuable. In this regard, amphioxus is not only a useful model system, but sometimes also the only possible alternative. For instance, one of the developmental processes that has been more deeply studied in this Thesis is somitogenesis. As reasoned in section 2 (this discussion), this process can be compared only between vertebrates and invertebrates, using amphioxus as a model organism, as it is the closest living invertebrate relative to vertebrates and the only system (apart from vertebrates themselves) in which this process can be analysed, as somites represent a sinapomorphy of vertebrates and cephalochordates (Saga and Takeda, 2001).

4. LIMITATIONS IN THE USE OF AMPHIOXUS AS A MODEL SYSTEM

The first and most serious limitation on the use of amphioxus as a model system is the short breeding season of natural amphioxus populations. In theory, embryos can be obtained during a three month period (between June and September; Stokes and Holland, 1996) but in practice, this period is reduced to one month (August). Hence, experiments *in vivo*, most of which, like transgenesis and generation of knock out organisms by RNA interference (necessary for the study of gene functions in animals with no genetic tradition), are complex and time-consuming, can be performed only during a short period of time. We carried out RNA interference with the *AmphiMox* gene, but further analysis of the embryos revealed that they had no phenotype. We had probably used too low a concentration of *AmphiMox* double-stranded RNA. Moreover, the short period of time, the low amount of embryos and the poor availability of equipment in the laboratory of Tampa did not allow us to perform a huge battery of trials. Other members of our laboratory have also attempted to generate transgenic embryos by either electroporation or transfection. They have encountered the same limiting factors. Moreover, standard gene reporters such as the green fluorescent protein (GFP) and β -galactosidase cannot be used in amphioxus for a transgenic analysis, because they have endogenous corporal fluorescence and β -galactosidase activity (Cañestro *et al.*, 2001; Benito and Garcia-Fernández, personal communication). This does not imply that a transgenic approach cannot be performed in lancelets, but that other reporters (not the commonly used, like the above mentioned) or strategies are needed. Such strategies, less common and so less standardised, require a higher number of trials and are thus more time-consuming.

Since these functional approaches are not yet implementable, amphioxus and vertebrate genes can be compared only according to changes or similarities between their expression patterns. The presence of mRNA in a specific region is not always indicative of the activity of the protein it encodes, e.g. long-range signalling molecules do not operate within the cell that secretes them; accordingly, *in situ* hybridisation detects a region other than that in which the gene function resides. Nevertheless, we studied only genes encoding for transcription factors which are widely accepted to play a role within the cell that has produced them. However, the non-autonomous behaviour of transcription factors, especially of homeodomain-containing transcription factors, has also been reported elsewhere (reviewed in Prochiantz, 2000). This problem can be solved by studying protein distribution rather than mRNA distribution. However, there are still limitations regarding the correlation between protein distribution and gene function. The protein encoded by the gene under study can be present in some regions that lack any overt phenotype in the null mutant, e. g. the *Drosophila ems* gene, whose homologue has been studied in this Thesis, is expressed in a specific domain in all the fly parasegments during embryogenesis. Nonetheless, the abnormal phenotype is restricted to the posterior spiracles (in the eighth abdominal segment) and to the cephalic ones (Dalton *et al.*, 1989).

Another problem is derived from the use of the species selected for this Thesis, *B. floridae*. It is collected in Tampa Bay, Florida, which is its natural habitat. Hence, another obstacle comes to mind, at least for the European labs working on this species. This limitation also strengthens the first one (the length of the amenable breeding season), as although in principle we could stay longer in Tampa (as the breeding season is longer than our three-week

stay), a longer stay would be too expensive. The use of a European species such as *Branchiostoma lanceolatum* may solve this problem. However, all the genes that have been studied in this Thesis have been isolated from the Florida lancelet, and if further *in vivo* work is to be performed on *lanceolatum*, we should use the genes of this species, as the genomes of the two species may differ. In fact, for some cases in which specific genes have been isolated from more than one species, they may be duplicated in one but not in all the species under analysis (chapter IV, Results). Furthermore, the use of *B. floridae* double-stranded RNA to silence *B. lanceolatum* RNAs is not feasible, since an acceptable level of silencing is reached only when there is less than 10% mismatch between the initiator and the target RNA (Baulcombe, 2002).

Another drawback in the use of lancelets is the high level of polymorphism (both somatic and germinal) in amphioxus natural populations (Cañestro, 2001), which hampers the use of certain molecular biology techniques. For example, the difficulty in obtaining a clear genomic Southern banding pattern is well known among the amphioxus community. In addition, it is difficult to ascertain whether different cDNAs represent distinct genes or highly polymorphic alleles of the same gene. This limitation can be solved at the "genomic level" by the use of libraries made from DNA obtained from a single individual, or by the digestion of genomic DNA from single individuals for Southern blots. However, there is no way to solve this problem at the "cDNA level", owing to the low amount of RNA present in single individuals (it is not feasible to construct a cDNA library from RNA extracted from a single embryo).

The last restraint is the lack of compaction of the amphioxus genome, which has three times as much DNA as an ascidian genome (Ohno, 1970), although both lineages escaped the genome wide duplications that took place early at vertebrate evolution, and may thus have a similar number of genes. Hence, we have to assume that the cephalochordate genome has more "junk" DNA (as repetitive elements) than the urochordate genome. Accordingly, the presence of such unwanted DNA hinders the use of some techniques, e. g. chromosome walking, which has been used along this Thesis.

4.1. Amphioxus maintenance and reproduction in the lab

The maintenance of amphioxus populations in the lab may minimize technical and economical limitations. However, all trials have proved unsuccessful so far. Adult *B. floridae* and *B. lanceolatum* have been brought alive from Tampa and from the North of Catalonia, respectively, to the laboratory at Barcelona. They have been maintained in closed circuit aquaria with filtered (or sterilised) seawater and fed with algae and commercial food for alevin fishes. However, after a few weeks most died owing to fungal infections, and those that survived never developed gonads. To our knowledge, the life cycle of cephalochordates has not been reproduced or controlled in lab conditions so far. Nevertheless, if there was a collection site in Europe (and a laboratory nearby), the European amphioxus community would be able to afford the three-month breeding season, and thus collect specimens and perform extensive *in vivo* experiments.

In their natural habitat, amphioxus adults emerge from the sand where they are burrowed down and swim near the bottom to discharge the gametes via the atriopore fortnightly during the breeding season. This phenomenon occurs always soon after sunset, because spawning is a dark-triggered phenomenon in amphioxus (Wickstead, 1975). In the lab, they can be forced to spawn under various stimuli. In Tampa, ripe *B. floridae* spawn after the application of a 2-4 s electric shock (50V, 10 ms pulses, 9 pulses/s) (Holland and Holland, 1989). Japanese and Chinese scientists have maintained for a few months ripe adult amphioxus in continuously aerated 100-l tanks with sand in the bottom and current seawater (roughly 1000 individuals of each sex per tank) at room temperature under natural (or artificial) light cycles and fed with collected plankton. Several individuals naturally spawn mature gametes when the room temperature exceeds 25°C (which corresponds to the breeding season average temperature) (Wu *et al.*, 1994; Watanabe *et al.*, 1999). Thomas Stack (personal communication) from the University of Tübingen, Germany, also obtained *B. lanceolatum* embryos *in vitro*. Ripe animals were collected from the North sea, directly transferred to the laboratory, kept in small dishes with boiled and aerated seawater at 10°C and fed with algae. To trigger natural spawning, a few males and females were transferred to a room at 18°C early in the morning, and spawning took place in the evening. Although a period longer than the standard maximum of the three months of breeding season would be desirable, the period during which spawning can be triggered is unknown.

Some natural populations of *B. lanceolatum* inhabit the shallow soft-bottoms of the north-western Mediterranean coast (mainly Blanes and Banyuls, North Catalonia/South France). Therefore, this project seems reasonable for a Barcelona-based team. The first aim would be to study the ecology of the population. To this end, we should collect specimens daily during the breeding season, and determine the period during which they are ripe and can thus be spawned in the lab, and which conditions are more suitable (temperature, darkness/light cycles and electroshock, among others). Data on the habitat of European lancelets in Blanes have been reported elsewhere (Sardá *et al.*, 1999). Two facts should be taken into account:

- The amphioxus population density in Blanes peaks during spring, decreases sharply throughout summer and remains low in autumn and winter. Unfortunately, these findings refer to the adult biomass, and ripe adult specimens, which are required to trigger spawning, have never been found.
- The collection of adult amphioxus is more problematic here, as the *B. lanceolatum* populations of Blanes are usually found at an average depth of 15 m, which makes the use of boats and dredges necessary.

4.2. Complementary approaches and future prospects

The advent of the post-genomic era strengthens the need for functional information. In the past few years, some genomes have been completely sequenced, and more are currently being sequenced. Owing to their privileged position in the phylogenetic tree as the closest living invertebrate relative to vertebrates (Wada and Saroh, 1994), cephalochordates are candidate organisms for genome projects. Moreover, the sequencing of their genome would without doubt shed light on the red-hot debate about vertebrate origins by two rounds of polyploidization. This

project (the sequencing of the amphioxus genome) has been reported to be necessary in the human genome sequence article in Nature (International Human Genome Sequencing Consortium, 2001). Therefore, it is not surprising that lancelets occupy the second position in the ranking of species whose genome should be sequenced as soon as possible by a NIH initiative. Hence, the cloning of a particular gene (or gene family) will not have its current value, and some high throughput labs will "win the race" of the expression work and thus the comparative approach.

To complement this comparative approach with some functional data, working with heterologous systems may be attractive, although the best solution would obviously be to maintain and reproduce amphioxus in laboratory conditions, and thus direct the functional approach to amphioxus themselves. Ascidians are sometimes used as an alternative, because functional work is possible on their embryos. Nevertheless, they are not good candidates to study the developmental process that has been examined in this Thesis, i. e. somitogenesis, as urochordates do not have somites. However, the isolation and study of the function of Mox and hairy genes in ascidians may allow us to gain more insight into the origin of somitogenesis, by comparison of the amphioxus genes and their ascidian homologues, but not to determine the increase in complexity of vertebrate somitic derivatives.

The use of vertebrate embryos as heterologous systems to study amphioxus genes seems more exciting for continuing the work presented here. This approach would tell us how vertebrates increased in complexity along evolution (i. e. through changes at the regulatory or at the coding sequence levels). As far as the *AmphiMox* gene is concerned, such a strategy would be helpful. The analysis of transgenic mice carrying a reporter gene under the control of the *AmphiMox* regulatory region will tell us how vertebrate Mox genes were co-opted for the differentiation of somitic derivatives after vertebrate genome duplications. If the reporter gene is expressed in segmented somites and somitic derivatives, changes at the level of upstream Mox activating factors have occurred and caused the novel participation of Mox class genes in the differentiation of somites. In contrast, if the *AmphiMox cis* regulatory region fails to drive the reporter expression within the mice segmented somites, changes in this regulatory region must have occurred to permit the acquisition of vertebrate Mox genes novel expression sites. These experiments have been carried out to study the regulatory regions of the anterior-most amphioxus Hox genes (Manzanares *et al.*, 2000). However, they have other drawbacks, e.g. they are very slow (owing to the mice vital cycle compared with, for example, that of the above mentioned urochordates). Therefore, the use of other vertebrate models may also provide useful data. The recent techniques developed to drive the expression of transgenes in particular regions of the chick embryo by electroporation (reviewed in Itasaki *et al.*, 1999), together with the easier manipulation of chick compared with mice embryos, make this system very attractive. In my opinion, the most exciting issue to study hereafter is the nature of the Amphihairy genes promoters and how amniote hairy genes acquired their cycling behaviour. The apparent subdivision of a very complex hairy role by the hairy duplicate copies in the amphioxus genome deserves a deeper analysis. If one assumes that the subfunctionalization has originated by differential losses of *cis* regulatory sequences, one must be struck by the high degree of complexity in the regulatory regions of the amphioxus hairy genes. It would be thrilling, for example, if only the combination of the four Amphihairy genes that may be involved in

embryogenesis rescued the severe defects observed in mice *HES1* null embryos. To test this hypothesis, the use of a vertebrate embryo as a heterologous system is also unavoidable.

Nevertheless, I really hope that functional approaches to the use of amphioxus embryos will soon be developed. The nature of their "non-redundant" genome (apart from some [perhaps more than expected?] exceptions) makes them very valuable to study a gene function within chordates, which is assumed to be very similar to the function of that gene in the ancestor of vertebrates, owing to the astonishing resemblance of extant amphioxus to some "pre-vertebrate" fossils. In this regard, the maintenance and reproduction of lancelets in laboratory conditions should be a priority.

Conclusions

1. RESPECTE ELS GENS DE L'AMFIOX

- Hi ha un únic gen de la classe Mox al genoma de *B. floridae*. Aquest gen -*AmphiMox*- representa el pro-ortòleg dels dos gens de la classe Mox de vertebrats.
- El *cluster* Hox de *B. floridae* conté un mínim de catorze gens, sis dels quals són del tipus Hox posterior, i no estan lligats a cap gen de la classe Evx a una distància de 90 Kb.
- El genoma de *B. floridae* conté dos gens de la classe Evx que estan lligats. *AmphiEvxA* és prototípic i igual de semblant als dos gens Evx de vertebrats, mentre que *AmphiEvxB* ha divergit molt.
- Hi ha dos gens de la classe Emx al genoma de *B. floridae*. Tots dos estan igualment relacionats amb ambdós gens Emx de vertebrats.
- Hi ha un mínim de sis gens de la família hairy al genoma de *B. floridae*. Tots sis representen trans-ortòlegs dels gens hairy de vertebrats.

2. RESPECTE HOMOLOGIES, INNOVACIONS DELS VERTEBRATS, DESENVOLUPAMENT I EXPRESSIÓ DE GENS DE L'AMFIOX

- La somitogènesi és un procés asimètric en cefalocordats. Aquesta asimetria és apreciable molecularment fins i tot abans de que els primers somites musculars siguin visibles.
- Els gens Mox, entre d'altres, podrien haver provocat l'increment de complexitat dels somites i derivats somítics en vertebrats.
- El comportament cíclic dels gens de la família hairy és una sinapomorfia d'amniotes.
- Cal l'estudi de més gens per determinar l'existència o no d'una MHB a la zona anterior del tub nerviós de l'amfiox.
- El *tail bud* apareix com a sinapomorfia dels cordats a nivell molecular.
- La *Hatschek's pit* i la *club-shaped gland* són de caràcter endocrí.

3. EVOLUCIÓ DELS GENS DEL GRUP *EXTENDED HOX* DE GENS AMB *HOMEBOX*

- La posició basal dels gens *Mox* i *Evx* respecte els gens Hox i ParaHox en arbres filogenètics suggereix que el seu origen és molt primerenc i anterior al dels gens Hox i ParaHox.

Conclusions

1. ABOUT AMPHIOXUS GENES

- *There is a single Mox gene in the B. floridae genome that is equally related to both vertebrate Mox genes.*

Similarly to other Mox genes in other species, *AmphiMox* is organised in three exons that in amphioxus span a region of about 13 Kb. One intronic position that is conserved in Mox genes splits the homeobox in two between residues 44 and 45 of the homeodomain.

- *The B. floridae Hox gene cluster is composed of a minimum of 14 Hox genes (six of them belonging to the posterior group), which are not linked to any Evx gene at a distance at least of 90 Kb 5' to the so far posterior-most Hox gene, AmphiHox14.*

In some respects, the organisation of the Posterior region of the amphioxus Hox cluster differs from a consensus mammalian cluster. First, it is at least three times longer than the average human cluster, although most of the relative intergenic distances are the same. Second, *AmphiHox11*, *12* and *14* have introns in their homeoboxes, unlike any vertebrate Hox gene. *AmphiHox12* has an intron between codons 37 and 38 of the homeobox, an unusual site for a homeobox intron. The position of the homeobox intron of *AmphiHox11* and *14* is the same as in *Drosophila AbdB*, between codons 44 and 45 of the homeobox, which is a common intron site for the Antennapedia-like group genes.

- *There are two Evx genes that are linked in the B. floridae genome within a region of 40 Kb. AmphiEvxA is equally related to the multiple vertebrate Evx genes, whereas AmphiEvxB is a highly divergent gene.*

Both genes are, similarly to other species, organised in three exons. The second intron is present within the homeobox, between residues 46 and 47 of the homeodomain, which is highly conserved among Evx-class genes.

- *There are two Emx genes that arose by independent duplication in the B. floridae genome, which are equally related to both vertebrate Emx genes.*
- *There are at least six canonical hairy genes in the B. floridae genome that are equally related to both vertebrate hairy groups.*

Moreover, there are at least two hairy-related sequences in the Florida lancelet genome that most probably represent pseudogenes. Although their genomic organisation is unknown, there is at least one conserved intronic position (intron 2) between them and the vertebrate hairy genes for which their exon/intron organisation has been analysed. Four of the six canonical Amphihairy genes (*AmphihairyA* to *D*) have undergone subfunctionalisation, which guarantees their maintenance in the amphioxus genome as duplicate genes, whereas the other Amphihairy genes (*AmphihairyE* and *F*) may undergo nonfunctionalisation.

2. ABOUT GENE EXPRESSION, HOMOLOGIES, VERTEBRATE NOVELTIES AND AMPHIOXUS DEVELOPMENT

- *Amphioxus somitogenesis is clearly asymmetrical even before the first muscular somites are appreciable.*

This has been demonstrated in terms of *AmphiMox* expression from the fifth enterocoelic somite to the last schyzocoelic somite. The expression of *AmphihairyB*, *C* and *D* shows that this asymmetry is present even from the very beginning of the somitogenesis process.

- ***Mox, among other genes, may have been instrumental in the increase in complexity of vertebrate somites and their derivatives.***

AmphiMox expression data support that the segmented vertebrate somites are much more complex than amphioxus somites.

- ***The cycling behaviour of hairy genes seems to be an amniote novelty.***

No hairy family representative cycles within the lower vertebrate presomitic mesoderm (PSM) nor does so any Amphihairy gene within the amphioxus PSM.

- ***A region similar to the vertebrate MHB seems to be absent in the amphioxus central nervous system, but may be present.***

The lack of *AmphiEvxA* expression in the anterior nerve cord supports that amphioxus does not have a tripartite brain. Nevertheless, *AmphihairyB* expression may support the opposite view. Hence, more expression patterns of the genes expressed within this region in vertebrates need to be studied.

- ***The tail bud emerges as a chordate sinapomorphy at the molecular level.***

AmphiEvxA, *AmphiMox*, and *AmphihairyB*, *C* and *D* expression patterns support the presence of a tail bud in cephalochordates. Thus, this region appears to be homologous to the vertebrate tail bud.

- ***The endocrine character of both the Hatschek's pit and the club-shaped gland is strengthened in terms of Amphihairy gene expression.***

3. ABOUT THE EVOLUTION OF THE EXTENDED-HOX GROUP OF HOMEBOX GENES

- ***New scenarios for the origin and evolution of Mox, Evx, Hox and ParaHox genes are proposed.***

The basal position of *Evx* and *Mox* genes suggests two scenarios for the origin and evolution of the extended Hox and ParaHox cluster genes. In the first, the duplication of the *ProtoMoeve* gene gives rise to *Evx* and *Mox* genes. In the second, the duplication of the *UrProtoMoHox* gene yields *Mox* and *UrProtoHox* genes.

A

- Abouheif, E., Akam, M., Dickinson, W.J., Holland, P.W.H., Meyer, A., Patel, N.H., Raff, R.A., Roth, V.L., and Wray, G.A. (1997). Homology and developmental genes. *Trends Genet.* **13**, 432-433.
- Aguinaldo, A.M., Turbeville, J.M., Linford, L.S., Rivera, M.C., Garey, J.R., Raff, R.A., and Lake, J.A. (1997). Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**, 489-493.
- Ahringer, J. (1996). Posterior patterning by the *Caenorhabditis elegans* even-skipped homolog *vab-7*. *Genes Dev.* **10**, 1120-1130.
- Akam, M. (1989). Hox and HOM: homologous gene clusters in insects and vertebrates. *Cell* **57**, 347-349.
- Akam, M. (1994). Insect development. Is pairing the rule? *Nature* **367**, 410-411.
- Allen JD, Lints T, Jenkins NA, Copeland NG, Strasser A, Harvey RP, Adams JM. (1991). Novel murine homeo box gene on chromosome 1 expressed in specific hematopoietic lineages and during embryogenesis. *Genes Dev.* **5**, 509-520.
- Amores, A., Force, A., Yan, Y.L., Joly, L., Amemiya, C., Fritz, A., Ho, R.K., Langeland, J., Prince, V., Wang, Y.L., Westerfield, M., Ekker, M., and Postlethwait, J.H. (1998). Zebrafish hox clusters and vertebrate genome evolution. *Science* **282**, 1711-1714.
- Atkin, N.B., and Ohno, S. (1967). DNA values of four primitive chordates. *Chromosoma* **23**, 10-13.
- Aulehla, A., and Johnson, R.L. (1999). Dynamic expression of *lunatic fringe* suggests a link between notch signaling and an autonomous cellular oscillator driving somite segmentation. *Dev. Biol.* **207**, 49-61.

B

- Bailey, W.J., Kim, J., Wagner, G.P., and Ruddle, F.H. (1997). Phylogenetic reconstruction of vertebrate Hox cluster duplications. *Mol. Biol. Evol.* **14**, 843-853.
- Balavoine G. (1996). Origine et diversification des complexes de gènes Hox: inventaire de ces gènes chez la planaire *Polycelis nigra*. PhD Thesis. Université Paris-Sud.
- Bastian, H., and Gruss, P. (1990). A murine *even-skipped* homologue, *Evx-1*, is expressed during early embryogenesis and neurogenesis in a biphasic manner. *EMBO J.* **9**, 1839-1852.
- Baulcombe D. (2002). RNA silencing. *Curr. Biol.* **12**, 82-84.
- Beeman, R. W. (1987). A homoeotic gene cluster in the red flour beetle. *Nature* **327**, 247-249.
- Bessho, Y., Sakata, R., Komatsu, S., Shiota, K., Yamada, S., and Kageyama, R. (2001). Dynamic expression and essential functions of *Hes7* in somite segmentation. *Genes Dev.* **15**, 2642-2647.
- Boncinelli, E., Somma, R., Acampora, D., Pannese, M., D'Esposito, M., Faiella, A., and Simeone, A. (1988). Organization of human homeobox genes. *Hum. Reprod.* **3**, 880-886.
- Borday, V., Thäeron, C., Avaron, F., Brulfert, A., Casane, D., Laurenti, P., and Geraudie, J. (2001). *evx1* transcription in bony fin rays segment boundaries leads to a reiterated pattern during zebrafish fin development and regeneration. *Dev. Dyn.* **220**, 91-98.
- Briata, P., Ilengo, C., Van DeWerken, R., and Corte, G. (1997). Mapping of a potent transcriptional repression region of the human homeodomain protein EVX1. *FEBS Lett.* **402**, 131-135.
- Brooke, N.M., Garcia-Fernández, J., and Holland, P.W.H. (1998). The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature* **392**, 920-922.
- Brown, S.J., Parrish, J.K., Beeman, R., and Denell, R.E. (1997). Molecular characterization and embryonic expression of the *even-skipped* ortholog of *Tribolium castaneum*. *Mech. Dev.* **61**, 165-173.
- Bruce, A.E., Oates, A.C., Prince, V.E., and Ho, R.K. (2001). Additional Hox clusters in the zebrafish: divergent expression patterns belie equivalent activities of duplicate Hoxb5 genes. *Evol. Dev.* **3**, 127-144.
- Brusca, R.C., and Brusca, G.J. (1990). Other Deuterostomes: Chaetognatha, Hemichordata, Chordata. In "Invertebrates". (A. D. Sinauer, Ed.). pp.841-889. Sunderland, Massachusetts.
- Bürglin, T.R. (1994). In "Guidebook to the Homeobox genes" (D. Duboule, Ed.). pp.25-71. Oxford University Press, Oxford.
- Bürglin, T. R. (1995) The Evolution of Homeobox genes. In "Biodiversity and Evolution" (R. Arai, M. Kato, and Y. Doi, Eds.). pp. 291-336. The National Science Museum Foundation, Tokyo.

Burke, A.C., Nelson, C.E., Morgan, B.A., and Tabin, C. (1995). *Hox* genes and the evolution of vertebrate axial morphology. *Development* **121**, 333-346.

C

Cameron, C.B., Garey, J.R., and Swalla, B.J. (2000). Evolution of the chordate body plan: New insights from phylogenetic analyses of deuterostome phyla. *Proc. Natl. Acad. Sci. USA*. **97**, 4469-4474.

Candia, A.F., Hu, J., Crosby, J., Lalley, P.A., Noden, D., Nadeau, J.H., and Wright, C.V.E. (1992). *Mox-1* and *Mox-2* define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse embryos. *Development* **116**, 1123-1136.

Candia, A.F., and Wright, C.V.E. (1995). The expression pattern of *Xenopus Mox-2* implies a role in initial mesodermal differentiation. *Mech. Dev.* **52**, 27-36.

Candia, A.F., and Wright, C.V.E. (1996). Differential localization of *Mox-1* and *Mox-2* proteins indicates distinct roles during development. *Int. J. Dev. Biol.* **40**, 1179-1184.

Cañestro, C., Albalat, R., Escrivà, H., and González-Duarte, R. (2001). Endogenous beta-galactosidase activity in amphioxus: a useful histochemical marker for the digestive system. *Dev. Genes Evol.* **211**, 154-156.

Cañestro, C. (2001). Caracterització molecular i evolutiva de la família de les Alcohol Deshidrogenases de cadena mitjana als cordats. PhD Thesis. Universitat de Barcelona.

Cañestro, C., Albalat, R., Hjelmqvist, L., Godoy, L., Jornvall, H., and 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.

Carroll, S.B., Laughon, A., and Thalley, B.S. (1988). Expression, function, and regulation of the *hairy* segmentation protein in the *Drosophila* embryo. *Genes Dev.* **2**, 883-890.

Carroll, S.B. (1995). Homeotic genes and the evolution of arthropods and chordates. *Nature* **376**, 479-85.

Cau, E., Gradwohl, G., Casarosa, S., Kageyama, R., and Guillemot, F. (2000). *Hes* genes regulate sequential stages of neurogenesis in the olfactory epithelium. *Development* **127**, 2323-2332.

Cerfontaine, P. (1906). Recherches sur le développement de l'Amphioxus. *Arch. de Biol.*, T. 22.

Chen, J.Y., Dzik, J., Edgecombe, G.D., Ramsköld, L., and Zhou, G.Q. (1995). A possible Early Cambrian chordate. *Nature* **377**, 720-722.

Chen, G., Fernandez, J., Mische, S., and Courey, A.J. (1999). A functional interaction between the histone deacetylase Rpd3 and the corepressor groucho in *Drosophila* development. *Genes Dev.* **13**, 2218-2230.

Chiang, C., Patel, N.H., Young, K.E., and Beachy P.A. (1994). The novel homeodomain gene *buttonless* specifies differentiation and axonal guidance functions of *Drosophila* dorsal median cells. *Development* **120**, 3581-3593.

Cohn, M.J., and Tickle, C. 1999. Developmental basis of limblessness and axial patterning in snakes. *Nature* **399**, 474-479.

Conklin, E.G. (1932). The embryology of amphioxus. *J. Morphol.* **54**, 69-151.

Conway Morris, S. (1982). In "Atlas of the Burgess Shale" (S. Conway Morris, Ed.) 26 Paleontological Association, London.

Cooke, J., and Zeeman, E.C. (1976). A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *J. Theor. Biol.* **58**, 455-476.

Cooke, J., Nowak, M.A., Boerlijst, M., and Maynard-Smith, J. (1997). Evolutionary origins and maintenance of redundant gene expression during metazoan development. *Trends Genet.* **13**, 360-364.

Coulier, F., Popovici, C., Villet, R., Birnbaum, D. (2000). MetaHox gene clusters. *J. Exp. Zool.* **288**, 345-351.

D

Dalton, D., Chadwick, R. and , McGinnis, W. (1989). Expression and embryonic function of *empty spiracles*: a *Drosophila* homeobox gene with two patterning functions on the anterior-posterior axis of the embryo. *Genes Dev.* **3**, 1940-1956.

- Dalton, R. (2000). Biologists flock to 'evo-devo' in a quest to read the recipes of life. *Nature* **403**, 125.
- David, J.R. (2001). Evolution and development: some insights from evolutionary theory. *An. Acad. Bras. Cienc.* **73**, 385-95.
- Davis, R.L., Turner, D.L., Evans, L.M., and Kirschner, M.W. (2001). Molecular targets of vertebrate segmentation: two mechanisms control segmental expression of *Xenopus hairy2* during somite formation. *Dev. Cell.* **1**, 553-565.
- Dawson, S.R., Turner, D.L., Weintraub, H., and Parkhurst, S.M. (1995). Specificity of the hairy/Enhancer of split Basic Helix-Loop-Helix (bHLH) Proteins Maps outside the bHLH domain and Suggests Two Separable Modes of Transcriptional Repression. *Mol. Cel. Biol.* **15**, 6923-6931.
- Degnan, B.M., Degnan, S.M., Fenteny G., and Morse, D.E. (1997). A Mox homeobox gene in the gastropod mollusc *Haliotis rufescens* is differentially expressed during larval morphogenesis and metamorphosis. *FEBS Lett.* **411**, 119-122.
- Del Barco Barrantes, I., Elia, A.J., Wunsch, K., De Angelis, M.H., Mak, T.W., Rossant, J., Conlon, R.A., Gossler, A., and de la Pompa, J.L. (1999). Interaction between Notch signalling and Lunatic fringe during somite boundary formation in the mouse. *Curr. Biol.* **9**, 470-480.
- Dermatzakis, E.T., and Clark, A.G. (2001). Differential selection after duplication in mammalian developmental genes. *Mol. Biol. Evol.* **18**, 557-562.
- de Rosa, R., Grenier, J.K., Andreeva, T., Cook, C.E., Adoutte, A., Akam, M., Carroll, S.B., and Balavoine, G. (1999). Hox genes in brachiopods and priapulids and protostome evolution. *Nature* **399**, 772-776
- Deschamps, J., van den Akker, E., Forlani, S., De Graaff, W., Oosterveen, T., Roelen, B., and Roelfsema, J. (1999). Initiation, establishment and maintenance of Hox gene expression patterns in the mouse. *Int. J. Dev. Biol.* **43**, 635-650.
- D'Esposito, M., Morelli, F., Acampora, D., Migliaccio, E., Simeone, A., and Boncinelli, E. (1991). *EVX2*, a human homeobox gene homologous to the *even-skipped* segmentation gene, is localized at the 5' end of HOX4 locus on chromosome 2. *Genomics* **10**, 43-50.
- Di Gregorio, A., Spagnuolo, A., Ristatore, F., Pischetola, M., Aniello, F., Branno, M., Cariello, L., and Di Lauro, R. (1995). Cloning of ascidian homeobox genes provides evidence for a primordial chordate cluster. *Gene* **156**, 253-257.
- Doe, C.Q., Smouse, D., and Goodman, C.S. (1988). Control of neuronal fate by the *Drosophila* segmentation gene *even-skipped*. *Nature* **333**, 376-378.
- Dollé, P., Fraulob, V., and Duboule, D. (1994). Developmental expression of the mouse *Evx-2* gene: relationship with the evolution of the HOM/Hox complex. *Development (Suppl.)*, 143-153.
- Duboule, D., and Dollé, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of *Drosophila* homeotic genes. *EMBO J.* **8**, 1497-1505.
- Duboule, D. (1994). Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development (Suppl.)*, 135-42.
- Duboule, D., and Morata, G. (1994). Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet.* **10**, 358-364.
- Duboule, D. (1995). Vertebrate Hox genes and proliferation: an alternative pathway to homeosis? *Curr. Opin. Genet. Dev.* **5**, 525-528.
- Duboule, D. (1998). Vertebrate hox gene regulation: clustering and/or colinearity? *Curr. Opin. Genet. Dev.* **8**, 514-518.
- Duboule, D., and Wilkins, A.S. (1998). The evolution of 'bricolage'. *Trends Genet.* **14**, 54-59.
- Dubrulle, J., McGrew, M.J., and Pourquié, O. (2002). FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* **106**, 219-232.
- Dush, M.K., and Martin, G.R. (1992). Analysis of mouse *Evx* genes: *Evx-1* displays graded expression in the primitive streak. *Dev. Biol.* **151**, 273-287.

E

Ellenberger, T., Fass, D., Arnaud, M., and Harrison, S.C. (1994). Crystal structure of transcription factor E47: E-box recognition by a basic region helix-loop-helix dimer. *Genes Dev.* **8**, 970-980.

Escrivà, H., Safi, R., Hänni, C., Langlois, M.C., Saumitou-Laprade, P., Stehelin, D., Capron, A., Pierce, R., and Laudet, V. (1997). Ligand binding was acquired during evolution of nuclear receptors. *Proc. Natl. Acad. Sci. USA.* **94**, 6803-6808.

F

Faiella, A., D'Esposito, M., Rambaldi, M., Acampora, D., Balsafiore, S., Stornaiuolo, A., Mallamaci, A., Migliaccio, E., Gulisano, M., and Simeone, A. (1991). Isolation and mapping of *EVX1*, a human homeobox gene homologous to *even-skipped*, localized at the 5' end of HOX1 locus on chromosome 7. *Nucleic Acids Res.* **19**, 6541-6545.

Ferrier, D.E.K., and Holland P.W.H. (2001). Ancient origin of the Hox gene cluster. *Nat. Rev. Genet.* **2**, 33-38.

Finkelstein, R., and Perrimon, N. (1990). The *orthodenticle* gene is regulated by *bicoid* and *torso* and specifies *Drosophila* head development. *Nature* **346**, 485-458.

Finkelstein, R., Smouse, D., Capaci, T.M., Spradling, A.C., and Perrimon, N. (1990). The *orthodenticle* gene encodes a novel homeo domain protein involved in the development of the *Drosophila* nervous system and ocellar visual structures. *Genes Dev.* **4**, 1516-1527.

Fisher, A., and Caudy, M. (1998). The function of hairy-related bHLH repressor proteins in cell fate decisions. *BioEssays* **20**, 298-306.

Force, A., Lynch, M., Pickett, F.B., Amores, A, Yan, Y., and Postlethwait, J. (1999). Preservation of Duplicate Genes by Complementary, Degenerative Mutations. *Genetics* **151**, 1531-1545.

Frasch, M., Hoey, T., Rushlow, C., Doyle, H., and Levine, M. (1987). Characterization and localization of the even-skipped protein of *Drosophila*. *EMBO J.* **6**, 749-759.

Frigerio, G., Burri, M., Bopp, D., Baumgartner, S., and Noll, M. (1986). Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. *Cell* **47**, 735-746.

Futreal, P.A., Cochran, C., Rosenthal, J., Miki, Y., Swenson, J., Hobbs, M., Bennett, L.M., Haugen-Strano, A., Marks, J., Barrett, J.C., Tavtigian, S.V., Shattuck-Eidens, D., Kamb, A., Skolnick, M., and Wiseman, R.W. (1994). Isolation of a divergent homeobox gene, MOX1, from the BRCA1 region on 17q21 by solution hybrid capture. *Hum. Mol. Genet.* **3**, 1359-1364.

G

Galant, R., and Carroll, S.B. (2002). Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* **415**, 910-913.

Gans, C. (1996). Study of lancelets: the first 200 years. *Isr. J. Zool.* **42**, S3-11.

García-Bellido, A. (1994). *Génética del desarrollo y de la evolución*. pp. 97-110. Ed. Arbor, Madrid.

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

Gauchat, D., Mazet, F., Berney, C., Schummer, M., Kreger, S., Pawlowski, J., and Galliot, B. (2000). Evolution of Antp-class genes and differential expression of Hydra Hox/paraHox genes in anterior patterning. *Proc. Natl. Acad. Sci. USA.* **97**, 4493-4498.

Gee, H. Vertebrate morphology. Return of the amphioxus. *Nature* **370**, 504-505.

Gee, H. (1996). *Before the backbone: views on the origin of the vertebrates*. Chapman and Hall, London.

Gehring, W.J. (1985). The molecular basis of development. *Sci. Am.* **253**, 153-162.

Gehring, W.J. (1992). The homeobox in perspective. *Trends Biochem. Sci.* **17**, 277-280.

Gellon, G., and McGinnis, W. (1998). Shaping animal body plans in development and evolution by modulation of Hox expression patterns. *BioEssays* **20**, 116-125.

- Gibson, T.J., and Spring, J. (1998). Genetic redundancy in vertebrates: polyploidy and persistence of genes encoding multidomain proteins. *Trends Genet.* **14**, 46-49.
- Gionti, M., Ristatore, F., Di Gregorio, A., Aniello, F., Branno, M., and Di Lauro, R. (1998). *Cihox5*, a new *Ciona intestinalis* Hox-related gene, is involved in regionalization of the spinal cord. *Dev. Genes Evol.* **207**, 515-523.
- Goodman, C.S., and Coughlin, B.C. (2000). Introduction. The evolution of evo-devo biology. *Proc. Natl. Acad. Sci. USA.* **97**, 4424-4425.
- Gorski, D.H., LePage, D.F., Patel, C.V., Copeland, N.G., Jenkins, N.A. and Walsh, K. (1993). Molecular cloning of a diverged homeobox gene that is rapidly down-regulated during the G0/G1 transition in vascular smooth muscle cells. *Mol. Cell. Biol.* **13**, 3722-3733.
- Goto, T., Macdonald, P., and Maniatis, T. (1989). Early and late periodic patterns of *even skipped* expression are controlled by distinct regulatory elements that respond to different spatial cues. *Cell* **57**, 413-422.
- Graham, A., Papalopulu, N., and Krumlauf, R. (1989). The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. *Cell* **57**, 367-378.
- Grigoriou, M., Kastrinaki, M.C., Modi, W.S., Theodorakis, K., Mankoo, B., Pachnis, V., and Karagogeos, D. (1995). Isolation of the human *MOX2* homeobox gene and localization to chromosome 7p22.1-p21.3. *Genomics* **26**, 550-555.
- Grossniklaus, U., Cadigan, K.M., and Gehring, W.J. (1994). Three maternal coordinate systems cooperate in the patterning of the *Drosophila* head. *Development* **120**, 3155-3171.

H

- Han, K., and Manley, J.L. (1993). Transcriptional repression by the *Drosophila* even-skipped protein: definition of a minimal repression domain. *Genes Dev.* **7**, 491-503.
- Harris, C.L. (1996). Introduction to Chordates. In "Concepts in Zoology". (V. Cohen, Ed.) pp. 724-744. Harper Collins Publishers Inc., New York.
- Hart, C.P., Awgulewitsch, A., Fainsod, A., McGinnis, W., and Ruddle, F.H. (1985). Homeo box gene complex on mouse chromosome 11: molecular cloning, expression in embryogenesis, and homology to a human homeo box locus. *Cell* **43**, 9-18.
- Hatschek, B. (1893). The amphioxus and its development. Swan Sonnenschein and Co., London.
- Hérault, Y., Hraba-Renevey, S., van der Hoeven, F., and Duboule, D. (1996). Function of the *Evx-2* gene in the morphogenesis of vertebrate limbs. *EMBO J* **15**, 6727-6738.
- Holland, L.Z., Pace, D.A., Blink, M.L., Kene, M., and Holland, N.D. (1995a). Sequence and expression of amphioxus alkali myosin light chain (*AmphiMLC-alc*) throughout development: implications for vertebrate myogenesis. *Dev. Biol.* **171**, 665-676.
- Holland, L.Z., Kene, M., Williams, N.A., and Holland, N.D. (1997). Sequence and embryonic expression of the amphioxus *engrailed* gene (*AmphiEn*): the metameric pattern of transcription resembles that of its segment-polarity homolog in *Drosophila*. *Development* **124**, 1723-1732.
- Holland, N.D., and Holland, L.Z. (1993). Embryos and larvae of invertebrate deuterostomes. In "Essential Developmental Biology: a Practical Approach" (C. D. Stern and P. W. H. Holland, Eds.) pp. 21-32. IRL Press, Oxford.
- Holland, N.D., Holland, L.Z., and Kozmik, Z. (1995b). An amphioxus Pax gene, *AmphiPax-1*, expressed in embryonic endoderm, but not in mesoderm: implications for the evolution of class I paired box genes. *Mol. Mar. Biol. Biotechnol.* **4**, 206-214.
- Holland, N.D., and Holland, L.Z. (1999). Amphioxus and the utility of molecular genetic data for hypothesizing body part homologies between distantly related animals. *Amer. Zool.* **39**, 630-640.
- Holland, P.W.H., Garcia-Fernández, J., Williams, N.A., and Sidow, A. (1994). Gene duplications and the origins of vertebrate development. *Development (Suppl.)*, 125-133.
- Holland, P.W.H., and Garcia-Fernández, J. (1996). Hox genes and chordate evolution. *Dev. Biol.* **173**, 382-395.

Holland, P.W.H. (1998). Major transitions in animal evolution: a developmental genetic perspective. *Amer. Zool.* **38**, 829-842.

Holland, P.W.H. (1999). The future of evolutionary developmental biology. *Nature* **402**, C41-44.

Holley, S.A., Geisler, R., Nusslein-Volhard, C. (2001) Control of *her1* expression during zebrafish somitogenesis by a delta-dependent oscillator and an independent wave-front activity. *Genes Dev.* **14**, 1678-1690.

Hugues, A.L., da Silva, J., Friedman, R. (2001). Ancient Genome duplications did not structure the human *Hox*-bearing Chromosomes. *Genome Res.* **11**, 771-780.

I

Ingham, P.W., Howard, K.R., and Ish-Horowicz, D. (1985). Transcription pattern of the *Drosophila* segmentation gene *hairy*. *Nature* **318**, 439-445.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860-920.

Ishibashi, M., Moriyoshi, K., Sasai, Y., Shiota, K., Nakanishi, S., and Kageyama, R. (1994). Persistent expression of helix-loop-helix factor *HES-1* prevents mammalian neural differentiation in the central nervous system. *EMBO J.* **13**, 1799-1805.

Ishibashi, M., Ang, S.L., Shiota, K., Nakanishi, S., Kageyama, R., and Guillemot, F. (1995). Targeted disruption of mammalian *hairy* and Enhancer of split homolog-1 (*HES-1*) leads to up-regulation of neural helix-loop-helix factors, premature neurogenesis, and severe neural tube defects. *Genes Dev.* **9**, 3136-3148.

Itasaki, N., Bel-Vialar, S., and Krumlauf, R. (1999). 'Shocking' developments in chick embryology: electroporation and in ovo gene expression. *Nat. Cell Biol.* **1**, 203-207.

Izpisua-Belmonte, J.C., Falkenstein, H., Dollé, P., Renucci, A., and Duboule, D. (1991). Murine genes related to the *Drosophila AbdB* homeotic genes are sequentially expressed during development of the posterior part of the body. *EMBO J.* **10**, 2279-2289.

J

Jarriault, S., Brou, C., Logeat, F., Schroeter, E.H., Kopan, R., and Israel, A. (1995). Signalling downstream of activated mammalian Notch. *Nature* **377**, 355-358.

Jensen, J., Pedersen, E.E., Galante, P., Hald, J., Heller, R.S., Ishibashi, M., Kageyama, R., Guillemot, F., Serup, P., and Madsen, O.D. (2000). Control of endodermal endocrine development by *Hes-1*. *Nat. Genet.* **24**, 36-44.

Jiang, Y.J., Aerne, B.L., Smithers, L., Haddon, C., Ish-Horowicz, D., and Lewis, J. (2000). Notch signalling and the synchronization of the somite segmentation clock. *Nature* **408**, 475-479.

Jiménez, G., Paroush, Z., and Ish-Horowicz, D. (1997). Groucho acts as a corepressor for a subset of negative regulators, including *Hairy* and *Engrailed*. *Genes Dev.* **11**, 3072-3082.

Joly, J.S., Joly, C., Schulte-Merker, S., Boulekbache, H., and Condamine, H. (1993). The ventral and posterior expression of the zebrafish homeobox gene *eve1* is perturbed in dorsalized and mutant embryos. *Development* **119**, 1261-1275.

Jouve, C., Palmeirim, I., Henrique, D., Beckers, J., Gossler, A., Ish-Horowicz, D., and Pourquié, O. (2000). Notch signalling is required for cyclic expression of the hairy-like gene *HES1* in the presomitic mesoderm. *Development* **127**, 1421-1429.

Jouve, C., Iimura, T., Pourquié, O. (2002). Onset of the segmentation clock in the chick embryo: evidence for oscillations in the somite precursors in the primitive streak. *Development* **129**, 1107-1117.

Jürgens, G., Wieschaus E., Nüsslein-Volhard, C., and Klüding, H. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. II. Zygotic loci on the third chromosome. *Wilhelm Roux's Arch.* **193**, 283-295.

K

- Katsuyama Y, Wada S, Yasugi S, and Saiga H. (1995). Expression of the labial group Hox gene *HrHox-1* and its alteration induced by retinoic acid in development of the ascidian *Halocynthia roretzi*. *Development* **121**, 3197-3205.
- Kenyon, C., and Wang, B. (1991). A cluster of Antennapedia-class homeobox genes in a nonsegmented animal. *Science* **253**, 516-517.
- Kim, C.B., Amemiya, C., Bailey, W., Kawasaki, K., Mezey, J., Miller, W., Minoshima, S., Shimizu, N., Wagner, G., Ruddle, F. (2000). Hox cluster genomics in the horn shark, *Heterodontus francisci*. *Proc. Natl. Acad. Sci. USA* **97**, 1655-1660.
- Kimura, M., and King, J.L. (1979). Fixation of a deleterious allele at one of two "duplicate" loci by mutation pressure and random drift. *Proc. Natl. Acad. Sci. USA* **76**, 2858-2861.
- Kondo, T., and Duboule, D. (1999). Breaking colinearity in the mouse HoxD complex. *Cell* **97**, 407-417.
- Kourakis, M.J., and Martindale, M.Q. Combined-method phylogenetic analysis of Hox and ParaHox genes of the metazoa. *J. Exp. Zool.* **288**, 175-191.
- Koyanagi, M., Ono, K., Suga, H., Iwabe, N., and Miyata, T. (1998). Phospholipase C cDNAs from sponge and hydra: antiquity of genes involved in the inositol phospholipid signaling pathway. *FEBS Lett.* **439**, 66-70.
- Kuhn, K., Streit, B., and Schierwater, B. (1999). Isolation of Hox genes from the scyphozoan *Cassiopeia xamachana*: implications for the early evolution of Hox genes. *J. Exp. Zool.* **285**, 63-75.

L

- Langeland, J.A., Tomsa, J.M., Jackman, W.R, and Kimmel, C.B. (1998). An amphioxus *snail* gene: Expression in paraxial mesoderm and neural plate suggests a conserved role in patterning the chordate embryo. *Dev. Genes Evol.* **208**, 569-577.
- Lardelli, M., and Ish-Horowicz, D. (1993). *Drosophila hairy* pair-rule gene regulates embryonic patterning outside its apparent stripe domains. *Development* **118**, 255-266.
- Ledent, V., and Vervoort, M. (2001). The Basic Helix-Loop-Helix Protein Family: Comparative Genomics and Phylogenetic Analysis. *Genome Res.* **11**, 754-770.
- Leimeister, C., Dale, K., Fischer, A., Klamt, B., Hrabe de Angelis, M., Radtke, F., McGrew, M.J., Pourquié, O., and Gessler, M. (2000). Oscillating expression of *c-Hey2* in the presomitic mesoderm suggests that the segmentation clock may use combinatorial signaling through multiple interacting bHLH factors. *Dev. Biol.* **227**, 91-103.
- Leve, C., Gajewski, M., Rohr, K.B., and Tautz, D. (2001). Homologues of *c-hairy1* (*her9*) and *lunatic fringe* in zebrafish are expressed in the developing central nervous system, but not in the presomitic mesoderm. *Dev. Genes Evol.* **211**, 493-500.
- Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565-570.
- Li, X., and Noll, M. (1994). Evolution of distinct developmental functions of three *Drosophila* genes by acquisition of different *cis* regulatory regions. *Nature* **367**, 83-87.
- Locascio, A., Aniello, F., Amoroso, A., Manzanares, M., Krumlauf, R., and Branno, M. (1999). Patterning the ascidian nervous system: structure, expression and transgenic analysis of the *CiHox3* gene. *Development* **126**, 4737-4748.
- Lokki, J., and Saura, A. (1980). Polyploidy in Insect evolution. (W. H. Lewis, Ed.). pp. 270-320. Plenum Press, New York.
- Long, S., and Byrne, M. (2001). Evolution of the echinoderm Hox gene cluster. *Evol. Dev.* **3**, 302-311.
- Lukianov, K. A., Tarabykin, V. S., Potapov, V. K., Bekman, E. P., and Lukianov, S. A. (1994). The cloning of fragments of homeobox genes expressed during planarian regeneration *Ontogenez* **25**, 28-32.
- Lundin, L.G. (1993). Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. *Genomics* **16**, 1-19.
- Lundin, L.G. (1999). Gene duplications in early metazoan evolution. *Semin. Cell Dev. Biol.* **10**, 523-530.
- Lynch, M., and Walsh, J.B. (1998). Genetics and Analysis of Quantitative Traits. Sinauer Associates, Sunderland, MA.

Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151-1155.

Lynch, M., and Force, A. (2000). The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**, 459-473.

M

Macdonald, P.M., Ingham, P., and Struhl, G. (1986). Isolation, structure, and expression of *even-skipped*: a second Pair-Rule gene of *Drosophila* containing a Homeo Box. *Cell* **47**, 721-734.

Macías, A., and Morata, G. (1996). Functional hierarchy and phenotypic suppression among *Drosophila* homeotic genes: the *labial* and *empty spiracles* genes. *EMBO J.* **15**, 334-343.

Mankoo, B.S., Collins, N.S., Ashby, P., Grigorieva, E., Pevny, L. H., Candia, A., Wright, C.V.E., Rigby, P.W. J., and Pachnis, V. (1999). *Mox2* is a component of the genetic hierarchy controlling limb muscle development. *Nature* **400**, 69-73.

Mann, R.S. (1997). Why are Hox genes clustered? *BioEssays* **19**, 661-664.

Manzanares, M., Wada, H., Itasaki, N., Trainor, P.A., Krumlauf, R., and Holland, P.W.H. (2000). Conservation and elaboration of *Hox* gene regulation during evolution of vertebrate head. *Nature* **408**, 854-857.

Martindale, M.Q., and Henry, J.Q. (1999). Intracellular fate mapping in a basal metazoan, the ctenophore *Mnemiopsis leidyi*, reveals the origins of mesoderm and the existence of indeterminate cell lineages. *Dev. Biol.* **214**, 243-257.

Martinez, P., Rast, J.P., Arenas-Mena, C., and Davidson, E.H. Organization of an echinoderm Hox gene cluster. *Proc. Natl. Acad. Sci. USA* **96**, 1469-1474.

Massari, M.E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* **20**, 429-440.

Matsui, T., Hirai, M., Hirano, M., and Kurosawa, Y. (1993). The *HOX* complex neighbored by the *EVX* gene, as well as two other homeoboxcontaining genes, the *GBX*-class and the *EN*-class, are located on the same chromosomes 2 and 7 in humans. *FEBS Lett.* **336**, 107-110.

McGinnis, W., Hart, C. P., Gehring W. J., and Ruddle, F. H. (1984). Molecular Cloning and Chromosome Mapping of a Mouse DNA Sequence Homologous to Homeotic Genes of *Drosophila*. *Cell* **38**, 675-680

McGinnis W, Krumlauf R. (1992). Homeobox genes and axial patterning. *Cell.* **68**, 283-302.

McGinnis, W. (1994). A century of homeosis, a decade of homeoboxes. *Genetics* **137**, 607-611.

McGrew, M.J., Dale, J.K., Fraboulet, S., and Pourquié, O. (1998). The *lunatic fringe* gene is a target of the molecular clock linked to somite segmentation in avian embryos. *Curr. Biol.* **8**, 979-982.

Miller, D.J., and Miles, A. (1993). Homeobox genes and the zootype. *Nature* **365**, 215-216.

Mito, T., and Endo, K. (2000). PCR survey of Hox genes in the crinoid and ophiuroid: evidence for anterior conservation and posterior expansion in the echinoderm Hox gene cluster. *Mol. Phylogenet. Evol.* **14**, 375-388.

Mokady, O., Dick, M.H., Lackschewitz, D., Schierwater, B., and Buss, L.W. (1998). Over one-half billion years of head conservation? Expression of an *ems* class gene in *Hydractinia symbiolongicarpus* (Cnidaria: Hydrozoa). *Proc. Natl. Acad. Sci. USA* **95**, 3673-3678.

Moloney DJ, Panin VM, Johnston SH, Chen J, Shao L, Wilson R, Wang Y, Stanley P, Irvine KD, Haltiwanger RS, Vogt TF. (2000). Fringe is a glycosyltransferase that modifies Notch. *Nature* **406**, 369-375.

Moore, A.W., Barbel, S., Jan, L.Y., and Jan, Y.N. (2000). A genomewide survey of basic helix-loop-helix factors in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **97**, 10436-10441.

Morita, T., Nitta, H., Kiyama, Y., Mori, H., and Mishina, M. (1995). Differential expression of two zebrafish *emx* homeoprotein mRNAs in the developing brain. *Neurosci. Lett.* **198**, 131-134.

Murre, C., McCaw, P.S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **56**, 777-783.

Myojin, M., Ueki, T., Sugahara, F., Murakami, Y., Shigetani, Y., Aizawa, S., Hirano, S., and Kuratani, S. (2001). Isolation of *Dix* and *Emx* gene cognates in an agnathan species, *Lampetra japonica*, and their

expression patterns during embryonic and larval development: conserved and diversified regulatory patterns of homeobox genes in vertebrate head evolution. *J. Exp. Zool.* **291**, 68-84.

N

Nadeau, J.H., and Sankoff, D. (1997). Comparable rates of gene loss and functional divergence after genome duplications early in vertebrate evolution. *Genetics* **147**, 1259-1266.

Naito, M., Ishiguro, H., Fujisawa, T., Kurosawa, Y. (1993). Presence of eight distinct homeobox-containing genes in cnidarians. *FEBS Lett.* **333**, 271-274.

Naruse, K., Fukamachi, S., Mitani, H., Kondo, M., Matsuoka, T., Kondo, S., Hanamura, N., Morita, Y., Hasegawa, K., Nishigaki, R., Shimada, A., Wada, H., Kusakabe, T., Suzuki, N., Kinoshita, M., Kanamori, A., Terado, T., Kimura, H., Nonaka, M., and Shima, A. A detailed linkage map of medaka, *Oryzias latipes*: comparative genomics and genome evolution. *Genetics* **154**, 1773-1784.

Nelsen, O.E. (1953). Comparative embryology of vertebrates. Mc-Graw Hill, New York.

Neyt, C., Jagla, K., Thisse, B., Haines, L., and Currie P. D. (2000). Evolutionary origins of vertebrate appendicular muscle. *Nature* **408**, 82-86.

O

Oda, I., and Saiga, H. (2001). *Hremx*, the ascidian homologue of *ems/emx*, is expressed in the anterior and posterior-lateral epidermis but not in the central nervous system during embryogenesis. *Dev. Genes Evol.* **211**, 291-298.

Ohno, S. 1970. Evolution by gene duplication. George Allen and Unwin, London.

Ohtsuka, T., Ishibashi, M., Gradwohl, G., Nakanishi, S., Guillemot, F., and Kageyama, R. (1999). Hes1 and Hes5 as notch effectors in mammalian neuronal differentiation. *EMBO J.* **18**, 2196-2207.

Ono, K., Suga, H., Iwabe, N., Kuma, K., and Miyata, T. (1999). Multiple protein tyrosine phosphatases in sponges and explosive gene duplication in the early evolution of animals before the parazoan-eumetazoan split. *J. Mol. Evol.* **48**, 654-662.

P

Palmeirim, I., Henrique, D., Ish-Horowicz, D., and Pourquié, O. (1997). Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell.* **91**, 639-648.

Pankratz, M.J. and Jäckle, H. (1993). Blastoderm Segmentation. In "The Development of *Drosophila melanogaster*". (M. Bate and A. Martinez Arias, Eds.). pp. 467-516. Cold Spring Harbor Laboratory Press, New York.

Pannese, M., Lupo, G., Kablar, B., Boncinelli, E., Barsacchi, G., and Vignali, R. (1998). The *Xenopus* *Emx* genes identify presumptive dorsal telencephalon and are induced by head organizer signals. *Mech. Dev.* **73**, 73-83.

Pasini, A., Henrique, D., and Wilkinson, D.G. (2001). The zebrafish Hairy/Enhancer-of-split-related gene *heró* is segmentally expressed during the early development of hindbrain and somites. *Mech. Dev.* **100**, 317-321.

Patel, N.H., Ball, E.E., and Goodman, C.S. (1992). Changing role of *even-skipped* during the evolution of insect pattern formation. *Nature* **357**, 339-342.

Patel, N.H., Condrón, B.G., and Zinn, K. (1994). Pair-Rule expression patterns of *even-skipped* are found in both short- and long-germ beetles. *Nature* **367**, 429-434.

Pébusque, M.J., Coulier, F., Birnbaum, D., and Pontarotti, P. (1998). Ancient large-scale genome duplications: phylogenetic and linkage analyses shed light on chordate genome evolution. *Mol. Biol. Evol.* **15**, 1145-1159.

Pendleton, J.W., Nagai, B.K., Murtha, M.T., and Ruddle, F.H. (1993). Expansion of the Hox gene family and the evolution of chordates. *Proc. Natl. Acad. Sci. USA* **90**, 6300-6304.

Pollard, S.L., and Holland, P.W.H. (2000). Evidence for 14 homeobox gene clusters in human genome ancestry. *Curr. Biol.* **10**, 1059-1062.

- Popodi, E., Kissinger, J.C., Andrews, M.E., and Raff, R.A. (1996). Sea urchin Hox genes: insights into the ancestral Hox cluster. *Mol. Biol. Evol.* **13**, 1078-1086.
- Postlethwait, J.H., Yan, Y.L., Gates, M.A., Horne, S., Amores, A., Brownlie, A., Donovan, A., Egan, E.S., Force, A., Gong, Z., Goutel, C., Fritz, A., Kelsh, R., Knapik, E., Liao, E., Paw, B., Ransom, D., Singer, A., Thomson, M., Abduljabbar, T.S., Yelick, P., Beier, D., Joly, J.S., Larhammar, D., Rosa, F., et al. (1998). Vertebrate genome evolution and the zebrafish gene map. *Nat. Genet.* **18**, 345-9.
- Pourquié, O. (2001). Vertebrate somitogenesis. *Annu. Rev. Cell Dev. Biol.* **17**, 311-350.
- Presley, R., Horder, T.J., Slípka, J. (1996). Lancelet development as evidence of ancestral chordate structure. *Isr. J. Zool.* **42**, S-97-116.
- Prochiantz, A. (2000). Messenger proteins: homeoproteins, TAT and others. *Curr. Opin. Cell Biol.* **12**, 400-406.

Q

- Qiu, M., Anderson, S., Chen, S., Meneses, J.J., Hevner, R., Kuwana, E., Pedersen, R.A., and Rubenstein, J.L. (1996). Mutation of the *Emx-1* homeobox gene disrupts the corpus callosum. *Dev. Biol.* **178**, 174-178.

R

- Raff, R.A. (2000). Evo-devo: the evolution of a new discipline. *Nat. Rev. Genet.* **1**, 74-79.
- Rallis, C., Stamatakis, D., Pontikakis, S., Mankoo, B.S., and Karagogeos, D. (2001). Isolation of the avian homologue of the homeobox gene *Mox2* and analysis of its expression pattern in developing somites and limbs. *Mech. Dev.* **104**, 121-124.
- Ronshaugen, M., McGinnis, N., and McGinnis, W. (2002). Hox protein mutation and macroevolution of the insect body plan. *Nature* **415**, 914-917.
- Ruiz i Altaba, A., and Melton, D.A. (1989). Bimodal and graded expression of the *Xenopus* homeobox gene *Xhox3* during embryonic development. *Development* **106**, 173-183.

S

- Saga, Y. and Takeda, H. (2001). The making of the somite: Molecular events in vertebrate segmentation. *Nat. Rev. Genet.* **2**, 835-845.
- Sardá, R., Pinedo, S., and Martin, D. (1999). Seasonal dynamics of macroinfaunal key species inhabiting shallow soft-bottoms in the Bay of Blanes (NW Mediterranean). *Acta Oecologica* **20**, 315-326.
- Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R., and Nakanishi, S. (1992). Two mammalian helix-loop-helix factors structurally related to *Drosophila hairy* and Enhancer of split. *Genes Dev.* **12B**, 2620-2634.
- Schubert, F.R., Nieselt-Struwe, K., and Gruss, P. (1993). The Antennapedia-type homeobox genes have evolved from three precursors separated early in metazoan evolution. *Proc. Natl. Acad. Sci. USA* **90**, 143-147.
- Schubert, M., Holland, L.Z., Stokes, M.D., and Holland, N.D. (2001). Three amphioxus *Wnt* genes (*AmphiWnt3*, *AmphiWnt5*, and *AmphiWnt6*) associated with the Tail Bud: the evolution of somitogenesis in chordates. *Dev. Biol.* **240**, 262-273.
- Schultz, R.J. (1980). Role of Polyploidy in the Evolution of Fishes. (W. H. Lewis, Ed.). pp. 313-340. Plenum Press, New York.
- Scott, M.P., Weiner, A.J. (1984). Structural relationships among genes that control development: sequence homology between the *Antennapedia*, *Ultrabithorax*, and *fushi tarazu* loci of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **81**, 4115-9
- Shankland, M. (1994). Leech segmentation: a molecular perspective. *BioEssays* **16**, 801-808.
- Sharman, A.C., and Holland, P.W.H. (1998). Estimation of Hox gene cluster number in lampreys. *Int. J. Dev. Biol.* **42**, 617-620.
- Sharman, A.C. (1999). Some new terms for duplicated genes. *Semin. Cell Dev. Biol.* **10**, 561-563.

- Shimeld, S.M. (1999). Gene function, gene networks and the fate of duplicated genes. *Semin. Cell Dev. Biol.* **10**, 549-553.
- Shimeld, S.M., and Holland, P.W.H. (2000). Vertebrate innovations. *Proc. Natl. Acad. Sci. USA* **97**, 4449-4452.
- Sidow, A. (1996). Gen(om)e duplications in the evolution of early vertebrates. *Curr. Opin. Gen. Dev.* **6**, 715-722.
- Simeone, A., Acampora, D., Arcioni, L., Andrews, P.W., Boncinelli, E., and Mavilio, F. (1990). Sequential activation of HOX2 homeobox genes by retinoic acid in human embryonal carcinoma cells. *Nature* **346**, 763-766.
- Simeone, A., Gulisano, M., Acampora, D., Stornaiuolo, A., Rambaldi, M., and Boncinelli, E. (1992a). Two vertebrate homeobox genes related to the *Drosophila* empty spiracles gene are expressed in the embryonic cerebral cortex. *EMBO J.* **11**, 2541-2550.
- Simeone, A., Acampora, D., Gulisano, M., Stornaiuolo, A., and Boncinelli, E. (1992b). Nested expression domains of four homeobox genes in developing rostral brain. *Nature* **358**, 687-690.
- Skrabaneck, L., and Wolfe, K.H. (1998). Eukaryote genome duplication -where's the evidence? *Curr. Opin. Genet. Dev.* **8**, 694-700.
- Slack, J.M.W., Holland, P.W.H., and Graham, C.F. (1993). The zootype and the phylotypic stage. *Nature* **361**, 490-492.
- Sommer, R.J., and Tautz, D. (1993). Involvement of an orthologue of the *Drosophila* pair-rule gene *hairy* in segment formation of the short germ-band embryo of *Tribolium*. *Nature* **361**, 448-450.
- Sordino, P., Duboule, D., and Kondo, T. (1996). Zebrafish *Hoxa* and *Evx-2* genes: cloning, developmental expression and implications for functional evolution of posterior *Hox* genes. *Mech. Dev.* **59**, 165-175.
- Spring J. (1997). Vertebrate evolution by interspecific hybridisation -are we polyploid? *FEBS Lett.* **400**, 2-8.
- Spyropoulos, D.D., and Capecchi, M.R. (1994). Targeted disruption of the *even-skipped* gene, *evx1*, causes early postimplantation lethality of the mouse conceptus. *Genes Dev.* **8**, 1949-1961.
- Stamatakis, D., Kastrinaki, M-C., Mankoo, B.S., Pachnis, V., and Karagogeos, D. (2001). Homeodomain proteins Mox1 and Mox2 associate with Pax1 and Pax3 transcription factors. *FEBS Lett.* **499**, 274-278.
- Stellwag, E.J. (1999). Hox gene duplication in fish. *Semin. Cell Dev. Biol.* **10**, 531-40.
- Stokes, M.D., and Holland, N.D. (1996) Life-history of the Florida lancelet, *Branchiostoma floridae*: some factors affecting population dynamics in Tampa Bay. *Isr. J. Zool.* **42**, S67-86.
- Suga, H., Kuma, K., Iwabe, N., Nikoh, N., Ono, K., Koyanagi, M. Oshiyama, D., and Miyata, T. (1997). Intermittent divergent of the protein tyrosine kinase family during animal evolution. *FEBS Lett.* **412**, 540-546.
- Suga, H., Koyanagi, M., Joshiyama, D., Ono, K., Iwabe, N., Kuma, K., and Miyata, T. (1999). Extensive gene duplication in the early evolution of animals before the parazoan-eumetazoan split demonstrated by G proteins and protein tyrosine kinases from sponge and Hydra. *J. Mol. Evol.* **48**, 646-653.

T

- Taylor, J.S., Van de Peer, Y., Braasch, I., and Meyer, A. (2001). Comparative genomics provides evidence for an ancient genome duplication event in fish. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **356**, 1661-1679.
- Thaëron, C., Avaron, F., Casane, D., Borday, V., Thisse, B., Thisse, C., Boulekbache, H., and Laurenti, P. (2000). Zebrafish *evx1* is dynamically expressed during embryogenesis in subsets of interneurons, posterior gut and urogenital system. *Mech. Dev.* **99**, 167-172.
- The *C. elegans* Sequencing Consortium. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012-2018.

U

- Umbhauer, M., Boucaut, J.C., and Shi, D.L. (2001). Repression of *XMyoD* expression and myogenesis by *Xhairy-1* in *Xenopus* early embryo. *Mech. Dev.* **109**, 61-68.

V

- van Auken, K., Weaver, D.C., Edgar, L.G., and Wood, W.B. *Caenorhabditis elegans* embryonic axial patterning requires two recently discovered posterior-group Hox genes. *Proc. Natl. Acad. Sci. USA* **97**, 4499-4503.
- van der Hoeven, F., Zakany, J., and Duboule, D. (1996). Gene transpositions in the HoxD complex reveal a hierarchy of regulatory controls. *Cell* **85**, 1025-1035.
- von Dassow, G., and Munro, E. (1999). Modularity in animal development and evolution: elements of a conceptual framework for EvoDevo. *J. Exp. Zool.* **285**, 307-325.

W

- Wada, H., and Satoh, N. (1994). Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. *Proc. Natl. Acad. Sci. USA* **91**, 1801-1804.
- Wada, H., Garcia-Fernández, J., and Holland, P.W.H. (1999). Colinear and segmental expression of amphioxus Hox genes. *Dev. Biol.* **213**, 131-141.
- Walldorf, U., and Gehring, W.J. (1992). *Empty spiracles*, a gap gene containing a homeobox involved in *Drosophila* head development. *EMBO J.* **11**, 2247-2259.
- Walsh, J.B. (1995). How do often duplicated genes evolve new functions? *Genetics* **139**, 421-428.
- Warren, R.W., Nagy, L., Selegue, J., Gates, J., and Carroll, S. (1994). Evolution of homeotic gene regulation and function in flies and butterflies. *Nature* **372**, 458-461.
- Watanabe, T., Yoshida, M., and Shirai, H. (1999). Effect of Light on the Time of Spawning in the Amphioxus, *Branchiostoma belcheri tsingtauense*. *Contributions from the Ushimado Marine Laboratory* **37**, 1-7.
- Watterson, G.A. (1983). On the time for gene silencing at duplicate loci. *Genetics* **105**, 745-766.
- Weatherbee, S.D., Halder, G., Kim, J., Hudson, A., and Carroll, S. (1998). Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the *Drosophila* haltere. *Genes Dev.* **12**, 1474-1482.
- Weatherbee, S.D., Nijhout, H.F., Grunert, L.W., Halder, G., Galant, R., Selegue, J., and Carroll, S. (1999). Ultrabithorax function in butterfly wings and the evolution of insect wing patterns. *Curr. Biol.* **9**, 109-115.
- Whittaker, J. R. (1997). Cephalochordates, the Lancelets. In "Embryology: constructing the organism" (S. F. Gilbert and A. M. Raunio, Eds.) pp. 365-381. Sinauer Associates Inc. Publishers, Sunderland.
- Willey, A. (1894). *Amphioxus and the Ancestry of Vertebrates*. Macmillan, New York.
- Williams, N.A., and Holland, P.W.H. (2000). An amphioxus Emx homeobox gene reveals duplication during vertebrate evolution. *Mol. Biol. Evol.* **17**, 1520-1528.
- Wrischnik, L.A., and Kenyon, C.J. (1997). The role of *lin-22*, a *hairy/Enhancer of split* homolog, in patterning the peripheral nervous system of *C. elegans*. *Development* **124**, 2875-2888.
- Wolfe, K.H., and Shields, D.C. (1997). Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708-713.
- Wolfe, K. H. (2001). Yesterday's polyploids and the mystery of diploidization. *Nat. Rev. Genet.* **5**, 333-341.
- Wu, X.H., Zhang, S.C., Wang, Y.Y., Zhang, B.L., Qu, Y.M., and Jiang, X.J. (1994). Laboratory observation on spawning, fecundity and larval development of amphioxus (*Branchiostoma belcheri tsingtauense*). *Chinese J. Oceanol. Limnol.* **12**, 289-294.

Y

- Yoshida, M., Suda, Y., Matsuo, I., Miyamoto, N., Takeda, N., Kuratani, S., and Aizawa, S. (1997). *Emx1* and *Emx2* functions in development of dorsal telencephalon. *Development* **124**, 101-111.
- Younger-Shepherd, S., Vaessin, H., Bier, E., Jan, L.Y., and Jan, Y.N. (1992). *deadpan*, an essential pan-neural gene encoding an HLH protein, acts as a denominator in *Drosophila* sex determination. *Cell* **70**, 911-922.