The amphioxus Hairy family: differential fate after duplication

Carolina Minguillón¹,*, Senda Jiménez-Delgado¹, Georgia Panopoulou² and Jordi García-Fernàndez¹†

¹Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Avenida Diagonal 645, E-08028 Barcelona, Spain
²Evolution and Development Group, Depart. Prof. Hans Lehrach, Max-Planck-Institut fur Molekulare Genetik, 14195 Berlin (Dahlem), Germany

*Present address: Division of Developmental Biology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK
†Author for correspondence (e-mail: jgarcia@bio.ub.es)

Accepted 21 August 2003

Development 130, 5903-5914
© 2003 The Company of Biologists Ltd
doi:10.1242/dev.00811

Summary
Vertebrate Hairy genes are highly pleiotropic and have been implicated in numerous functions, such as somitogenesis, neurogenesis and endocrine tissue development. In order to gain insight into the timing of acquisition of these roles by the Hairy subfamily, we have cloned and studied the expression pattern of the Hairy gene(s) in amphioxus. The cephalochordate amphioxus is widely believed to be the living invertebrate more closely related to vertebrates, the genome of which has not undergone the massive gene duplications that took place during vertebrate evolution. Surprisingly, we have isolated eight Hairy genes from the ‘pre-duplicative’ amphioxus genome. In situ hybridisation on amphioxus embryos showed that Hairy genes had experienced a process of subfunctionalisation that is predicted in the DDC model (for duplication-degeneration-complementation). Only the summation of four out of the eight Amphi-Hairy genes expression resembles the expression pattern of vertebrate Hairy genes, i.e. in the central nervous system, presomitic mesoderm, somites, notochord and gut. In addition, Amphi-Hairy genes expression suggest that amphioxus early somites are molecularly prefigured in an anteroposterior sequence in the dorsolateral wall of the archenteron, and the presence of a midbrain/hindbrain boundary. The expansion of the amphioxus Hairy subfamily request for caution when deducing the evolutionary history of a gene family in chordates based in the singularity of the amphioxus genome. Amphioxus may resemble the ancestor of the vertebrates, but it is not the ancestor, only its closest living relative, a privileged position that should not assume the freezing of its genome.

Supplemental data available online

Key words: Hairy, bHLH, amphioxus, DDC model, Subfunctionalisation, Duplication

Introduction
Genes in the Hairy subfamily encode class E of basic helix-loop-helix (bHLH) transcription factors and act as transcriptional repressors regulating differentiation and embryonic patterning in many organisms (Ledent and Vervoort, 2001). In protostomes, Hairy members have been isolated from *Drosophila melanogaster* and the beetle *Tribolium*. There are three members of the Hairy family in the *D. melanogaster* genome ([*hairy (h)*], [*deadpan (dpn)*] and [*similar to deadpan (side)*]), that have arisen by independent duplications in the fly lineage (Moore et al., 2000), whereas a single Hairy homologue has been so far isolated from the beetle (Sommer and Tautz, 1993). In *Drosophila*, *h* acts during segmentation as a primary pair-rule gene (Carroll et al., 1988), and also during the development of the fly peripheral nervous system as a pre-pattern gene (Fisher and Caudy, 1998). *dpn* function is required during somatic sex determination as an autosomal factor, and also during the development of the peripheral nervous system where it acts as a precursor gene (Fisher and Caudy, 1998). In contrast to *h* and *dpn*, little is known about *side* function during *Drosophila* development. It is expressed in specific subsets of cells in the central nervous system (Moore et al., 2000). In *Tribolium*, the *hairy* gene also functions during segmentation, but is not involved in nervous system development (Sommer and Tautz, 1993).

In vertebrates, there are two Hairy genes in chicken [*Hairyl* and *Hairy2* (Palmeirim et al., 1997; Jouve et al., 2000)], in zebrafish [*her6* and *her9* (Pasini et al., 2001; Leve et al., 2001)] and in the frog [*X-Hairy1* and *X-Hairy2* (Davis et al., 2001)]. However, in mice, only a single Hairy gene has been so far identified [*Hes1* (Sasai et al., 1992)]. Other HES and HER genes (for Hairy-enhancer of split related) belong to the Enhancer of split [E(spl)] subfamily and not to the Hairy subfamily of the bHLH-O family of transcription factors (Davis and Turner, 2001).

For mouse *Hes1* and both chicken Hairy genes, a striking dynamic expression pattern in the presomitic mesoderm (PSM), where they cycle with a temporal periodicity that corresponds to the formation of one somite, has been observed (Palmeirim et al., 1997; Jouve et al., 2000). Besides this cycling pattern in PSM, their expression is also detected in several endoderm-derived tissues, the notochord, and the central nervous system (CNS) (Sasai et al., 1992). Accordingly, mouse mutant for the *Hes1* gene exhibit severe defects in neural and endocrine development. Briefly, these defects are thought to be due to the premature differentiation of postmitotic neurons or endocrine cells, respectively (Ishibashi et al., 1995; Jensen et
In neither the zebrafish nor *Xenopus* is there a dynamic expression pattern of Hairy genes in the PSM comparable with that in amniotes. During somitogenesis, the zebrafish Hairy gene *her6* is expressed in the posterior part of each segmented somite and in stripes in the anterior PSM. Within the CNS, *her6* is expressed first in the prospective forebrain, and later in hindbrain segmentation with a very dynamic, segmentally restricted pattern. Low levels of *her6* expression are also present in the notochord (Pasini et al., 2001). The zebrafish Hairy gene *her9* is also expressed during CNS development but in contrast to its parologue, neither in segmented somites nor in the PSM. Within the CNS, *her9* is predominantly expressed in the fore- and midbrain, and transiently in the hindbrain, leaving a non-expressing gap at the midbrain-hindbrain boundary (MHB) while it is also expressed in the midline mesoderm. However, other bHLH genes closely related to the Hairy subfamily, i.e. E(spl) subfamily members, exhibit a cyclic expression comparable with that of amniote Hairy genes (e.g. *her1* (Holley et al., 2000)). This similar behaviour has led some authors to propose that these genes are behaving as (or represent) functional Hairy genes, although they do not represent orthologous genes (Gajewski and Voolstra, 2002). In the frog, the two Hairy homologues are expressed in the CNS, somites and PSM. Both genes have identical expression patterns as a band prefiguring a new somite formation in the anterior PSM and are also transcribed weakly in segmented somites. Within the neuroectoderm, they exhibit a non-overlapping expression pattern (Davis et al., 2001).

A characteristic of vertebrate Hairy genes is therefore to be very pleiotropic, in contrast to the protostome Hairy subfamily members. To understand when this multiplicity of functions was acquired during evolution, we studied the Hairy subfamily in the cephalochordate amphioxus. Amphioxus is the closest living invertebrate relative to vertebrates and has not undergone the massive gene duplications events (up to polyploidization) that took place early during vertebrate evolution (see Spring, 2002). It is believed that amphioxus represents a direct descendant of the cephalochordate/vertebrate ancestor. Hence, amphioxus has been widely used as a model system to study the ancestral function of a gene family at vertebrate origins, represented by a single gene in the chordate ancestor that may be very similar to that of modern amphioxus. Surprisingly, we have isolated eight canonical Hairy family members from the ‘pre-duplicative’ amphioxus genome that we have called canonical Hairy family members from the ‘pre-duplicative’ modern amphioxus. Surprisingly, we have isolated eight genes closely related to the Hairy subfamily, i.e. E(spl) homologues, which we named *Amphi-hairyA*, *Amphi-hairyB*, *Amphi-hairyC* and *Amphi-hairyD*. The *h1* and *h2* primers were used to amplify by PCR most of the bHLH domain of *Amphi-hairyA* to *Amphi-hairyD*, and all four products were used in conjunction to perform a further screening over the same cDNA library under less stringent conditions (50°C). We thus isolated an additional Hairy homologue that we denominated *Amphi-hairyE*. *Amphi-hairyA* to *Amphi-hairyD* plus an additional *Amphi-hairyF* genes were also identified through an EST project of gastrula and neurula stages (Panopoulou et al., 2003).

To ascertain whether we cloned all amphioxus Hairy genes, we amplified the putative second intron of Amphi-Hairy genes by PCR with the degenerate primers *h1* (exon 2) and SPLIT R (exon 3) on multiple extractions of genomic DNA from single individuals. As no bands were visible, after primer cleaning, we performed a nested PCR reaction with the degenerate primers SPLIT F (exon 2) and *h2* (exon 3) (see Table S1 at http://dev.biologists.org/supplemental/) under the same conditions was performed. A band of the expected size was excised and subcloned. Sequencing showed its similarity to the Hairy subfamily genes.

The cloned fragment was used to screen a cDNA library from 6-20 hours postfertilization amphibious embryos (Langeland et al., 1998). Approximately 4x10⁶ pfu were screened at moderate stringency [55°C, Church’s buffer (Shifman and Stein, 1995)] with the PCR product. This led to the identification of four Hairy homologues, which we named *Amphi-hairyA*, *Amphi-hairyB*, *Amphi-hairyC* and *Amphi-hairyD*. The *h1* and *h2* primers were used to amplify by PCR most of the bHLH domain of *Amphi-hairyA* to *Amphi-hairyD*, and all four products were used in conjunction to perform a further screening over the same cDNA library under less stringent conditions (50°C). We thus isolated an additional Hairy homologue that we denominated *Amphi-hairyE*. *Amphi-hairyA* to *Amphi-hairyD* plus an additional *Amphi-hairyF* genes were also identified through an EST project of gastrula and neurula stages (Panopoulou et al., 2003).

In addition, we screened a PAC library gently provided by Chris Amemiya (Yale University, USA) with a mixture of Amphi-Hairy clones. The positive clone 51K5 was used as a template in PCR reactions (2 minutes at 94°C and then 35 cycles of 15 seconds at 94°C, 10 seconds at 60°C and 20 seconds at 72°C) with Amphi-Hairy specific oligonucleotides (hairyAF, hairyAR, hairyBF, hairyBR, RT-CF, RT-CR, hairyDF, hairyDR, RT-EF, RT-ER, RT-FF, RT-RR, RT-GF, RT-GR, RT-HF and RT-HR; see Table S1 at http://dev.biologists.org/supplemental/). *Amphi-hairyA* to *Amphi-hairyF* were contained in the PAC clone, whereas *Amphi-hairyG* and *Amphi-hairyAmphi-hairyH* were not.
Phylogenetic analysis

The putative protein sequences of the Amphi-Hairy genes were aligned with their homologues from other organisms using ClustalX (Thompson et al., 1994). The complete proteins were used to construct a Neighbour-joining tree. Topology robustness was assessed by 1000 bootstrap resampling of the data.

Obtaining embryos and in situ hybridisation

Ripe adults of the Florida lancelet, Branchiostoma floridae, were collected from Old Tampa Bay (Florida, USA) during the summer breeding season. The males and females were spawned electrically in the laboratory, and selected developmental stages were raised as described in Holland and Holland (Holland and Holland, 1993).

In situ hybridisation was performed according to previously published methods (Holland et al., 1996). When available, the 3’ coding region of each Amphi-Hairy gene (region that included neither the conserved bHLH domain nor the orange domain), which was amplified using primers h5 and h3rep (see Table S1 at http://dev.biologists.org/supplemental/), was used as a template for the DIG-labelled antisense probe.

After photographed as wholemounts, selected embryos were contrasted in 1% Poinceau S, 1% acetic acid, dehydrated through an ethanol series and embedded in Spurr’s resin. Serial 3 μm sections were obtained with a glass knife, mounted in DePeX and photographed under Nomarski optics.

RT-PCR

We performed RT-PCR experiments with the Amphi-Hairy genes for which we did not detect any expression in the whole-mount hybridisation, i.e. Amphi-hairyE to Amphi-hairyH. Specific primers (RT-EF, RT-ER, RT-FF, RT-FR, RT-GF, RT-GR, RT-HF and RT-HR; see Table S1 at http://dev.biologists.org/supplemental/) were designed and multiplex separate reactions were performed for each gene in conjunction with the use of Amphi-hairyC as an internal positive control. The primers used for Amphi-hairyC (RT-CF and RT-CR; see Table S1 at http://dev.biologists.org/supplemental/) were designed in such a way that the region amplified was larger than the region amplified for Amphi-hairyE to Amphi-hairyH.

cDNA from 12-, 15-, 18- and 21-hour embryos and adults was obtained by standard methods (J. R. Bayascas, PhD Thesis, University of Barcelona, 1997). Samples were used as a template for the PCR. PCR conditions were 1 minute at 94°C, and 30 cycles of 20 seconds at 94°C, 20 seconds at 55°C, 20 seconds at 72°C. The full-length cDNA clones of Amphi-hairyC, Amphi-hairyE and Amphi-hairyF, and genomic clones for Amphi-hairyG and Amphi-hairyH were used as templates for positive controls. After electrophoresis, gels were blotted and hybridised with Amphi-hairyC plus Amphi-hairyE-, Amphi-hairyF-, Amphi-hairyG- or Amphi-hairyH-specific probes.

Results

Isolation and characterisation of the amphioxus Hairy family

Using PCR, screenings at high and moderate stringency, and EST sequencing we isolated eight amphioxus Hairy genes that we called Amphi-hairyA to Amphi-hairyH. An alignment of Amphi-HairyA to Amphi-HairyG shows that the putative proteins encoded share higher similarity in the bHLH DNA binding and dimerisation domain, the orange domain, and the C-terminal WRPW domain of interaction with the co-repressor protein Groucho (Jiménez et al., 1997), plus further amino acid stretches between these domains (Fig. 1A). The bHLH domain of Amphi-HairyH, which is truncated at the most C-terminal end also shows high similarity to the other Amphi-Hairy proteins (Fig. 1A).

An alignment with other Hairy proteins from vertebrates and protostomes shows that the similarity is mainly observed within these three functional domains (see Fig. S1A at http://dev.biologists.org/supplemental/). Based on these alignments we concluded that all Amphi-hairyA to Amphi-hairyH are members of the Hairy subfamily of the bHLH-O family of transcription factors. Amphi-Hairy proteins bHLH domain is more closely related to that of vertebrate Hairy proteins (74-97%) than to those of other invertebrate Hairy proteins (67-88%) (see Fig. S1B at http://dev.biologists.org/supplemental/).

Phylogenetic analysis

To gain more insight into the relationships between Hairy genes, we conducted a molecular phylogenetic analysis by the neighbour-joining method on the complete Hairy proteins, using mouse and human HES2 sequences [E(spl) subfamily of bHLH-O] as outgroups (Fig. 1B). Vertebrate Hairy proteins clustered together (bootstrap value 100%) into two well-supported groups (values 100 and 78%), in agreement with other phylogenetic studies (Gajewski and Voolstra, 2002). All the Amphi-Hairy proteins form a monophyletic group (73%) that branches immediately outside these groups (as their sister group; 80%) suggesting that all they have originated by independent duplication in the cephalochordate lineage, in harmony with the hypothesis that vertebrate genes have originated by duplication after the cephalochordate-vertebrate divergence.

In a recent study (Gajewski and Voolstra, 2002), the existence of 5 Hairy-type genes in the Fugu rubripes genome, named FrHer6.1, FrHer6.2, FrHer9, FrHer10.1 and FrHer10.2 is highlighted. The authors conclude that more Hairy genes should exist in the zebrafish genome and that at least one of those types (the Her10 type) in other vertebrate genomes. To further analyse the phylogenetic relationships among the HES/HER subfamilies of bHLH transcription factors, we conducted similar phylogenetic trees with the whole set of human (Ledent et al., 2002), mouse, Drosophila, zebrafish and Fugu proteins. In our analyses, all vertebrate and invertebrate Hairy genes group together (71% bootstrap value), whereas the Her10 Fugu class groups with the enhancer of split representative HES2 (92%). Thus, only the fish Her6 and 9 classes represent truly Hairy-subfamily homologues, whereas the rest group with enhancer-of-split related sequences (see Fig. S2 at http://dev.biologists.org/supplemental/). We thus argue that there are only two classes of Hairy genes in vertebrates. However, in agreement with Gajewski and Voolstra data, it is still plausible that further Hairy homologues remain to be discovered in the zebrafish genome, as it is accepted that extensive gene duplications up to tetraploidisation occurred in the fish lineage (Amores et al., 1998; Robinson-Rechavi et al., 2001). This fact is represented here by the Fugu specific duplication of the Her6 gene (FrHer6.1 and FrHer6.2) (Gajewski and Voolstra, 2002). Furthermore, the misplacement of the chicken Hairy genes at the base of the rest of tetapode genes was noted previously (Gajewski and Voolstra, 2002) and may be due to a slight higher divergence rate of the chicken genes.

Amphi-hairyA to F are clustered in the lancelet genome

In order to ascertain whether amphioxus Hairy specific gene
duplications occurred by tandem or transgene duplication, we isolated and analysed a positive clone (51K5) from an amphioxus PAC library. We performed PCR on the PAC DNA with specific Amphi-Hairy primers to determine their presence in the genomic region isolated. We thus deduced that six of the eight Hairy genes (Amphi-hairyA to Amphi-hairyF) are linked in the amphioxus genome and contained in the same PAC (Fig. 2). However, we were unable to amplify Amphi-hairyG or Amphi-hairyH with the same strategy, concluding that they are absent in the genomic region contained in the PAC isolated.

**Amphi-hairyA, Amphi-hairyB, Amphi-hairyC and Amphi-hairyD genes expression in gastrula stages**

The expression of four different Amphi-Hairy genes (Amphi-hairyA, Amphi-hairyB, Amphi-hairyC and Amphi-hairyD) was studied by whole-mount in situ hybridisation. In brief, all these four genes have a specific expression pattern whose summation resembles that of the single (for the mouse) or the two Hairy genes that exist in other vertebrates.

No signal is detected at the blastula or early gastrula stages for any of the Amphi-Hairy genes. During mid-gastrula stage, Amphi-hairyA is expressed in two domains: in the anterior endoderm (Fig. 3A) and just outside the dorsal lip of the blastopore in the presumptive neural plate (arrow in Fig. 3A).

Expression of Amphi-hairyB, Amphi-hairyC and Amphi-hairyD is first detected at the very late gastrula stage in both the neural plate and the presumptive somitic mesoderm. In the presumptive somitic mesoderm all three genes are expressed in a striped pattern prefiguring the definitive somites. Specimens with 2, 3 and 4 bands are shown for Amphi-hairyB (Fig. 3B,C,E, respectively).

The expression pattern in the presumptive somitic mesoderm is very similar in Amphi-hairyB, Amphi-hairyC and Amphi-hairyD. For example, a three-stripe pattern is shown for Amphi-hairyC (Fig. 3F) and a four-stripe pattern is shown for Amphi-hairyD (Fig. 3G,H).

Amphi-hairyB, Amphi-hairyC and Amphi-hairyD are also expressed in the presumptive neural plate of the late gastrula. For Amphi-hairyB, expression is at about the level of the first and second somites (Fig. 3C, arrow) and is strongest laterally (Fig. 3D, arrows). For Amphi-hairyC, the relatively weak signal in the neural plate is in two regions: an anterior one between the
Fig. 2. Linkage of Amphi-hairyA to Amphi-hairyF genes in the amphioxus genome. DNA from PAC 51K5 was subjected to PCR reactions with Amphi-Hairy specific oligonucleotides. Amphi-Hairy A to Amphi-Hairy F (lines A to F) gave positive bands of the expected sizes (Amphi-hairyA, 300 bp; Amphi-hairyB, 156 bp; Amphi-hairyC, 311 bp; Amphi-hairyD, 249 bp; Amphi-hairyE, 235 bp; Amphi-hairyF, 207 bp), indicating that they are contained in the same genomic region.

Amphi-hairyA, B, C and D genes expression in neurula stages

During neurula stage, the expression of the four Amphi-Hairy genes becomes more gene-specific. Amphi-hairyA gene behaves according to the pattern observed for earliest stages: it is expressed in the endoderm and the posterior-most part of the neural tube. In early neurula stages the signal is conspicuous along most of the gut and also in the posterior-most third of the dorsal nerve cord (Fig. 4A, arrow). The nascent Hatschek’s left diverticulum also expresses the gene (arrowhead in Fig. 4A). In later neurulae, the signal remains in the posterior third of the neural tube (arrow in Fig. 4B), whereas the signal in the gut is now restricted to specific regions. It is confined to its ventral posterior-most part, a middle part, and in the anterior-most part, the signal is restricted dorsally and in the Hatschek’s left anterior diverticulum (arrowhead in Fig. 4B) and the anterior wall of the gut.

Amphi-hairyB also continues to be expressed in the same tissues as it was in gastrula stages. It is strongly expressed in the neural plate and also in the posterior part of the formed somites and in the most posterior paraxial mesoderm. In an oblique dorsolateral view of a very young neurula, the signal is detected in stripes at the posterior-most part of the segmented somites (the arrow in Fig. 4C marks the posterior border of the last formed somite) and very highly in the anterior-most PSM. It is also detected all along the neural plate (Fig. 4C). A similar pattern was observed for older neurulae within the somites and anterior PSM, but the signal in the neural plate is no longer detected over its entire length and appears more region-specific (Fig. 4D). In transverse sections (Fig. 4E-G from anterior to posterior), the signal in the neural plate (asterisk in Fig. 4E), in the neural plate and the somites (asterisk and arrows in Fig. 4F) or in the somites (arrows in Fig. 4G) is better observed. Later on, the signal in the somites becomes extinguished, and is only detected in the most posterior paraxial mesoderm (arrows in Fig. 4H,I). In the anterior part of the neural tube, there is a gap between the anterior-most expression domain and the next domain (lines in Fig. 4H,I).

During neurula stage, Amphi-hairyC (Fig. 4J) and Amphi-hairyD (Fig. 4K) have slightly different expression patterns. Although both where expressed in the neural plate (arrowheads in Fig. 4K,S, respectively), they were complementarily

---

Fig. 3. Amphi-hairyA, Amphi-hairyB, Amphi-hairyC and D genes expression in amphioxus gastrulae by whole-mount in situ hybridisation. All embryos in dorsal views (A-C,E-G) have anterior towards the left. The lateral view (H) has anterior towards the left and dorsal upwards. In the transverse optical dissection (D), dorsal is upwards looking from a posterior position. The presumptive somitic bands are numbered and the arrows indicate expression domains in the neural plate (with the exception of D). (A) Amphi-hairyA expression within the anterior inner endodermal cell layer and the posterior neural plate. (B) Dorsal view of an Amphi-hairyB two- to three-stripe gastrula. (C) Dorsal view of an Amphi-hairyB three-stripe embryo. (D) Optical transverse dissection of the specimen in C showing the highest Amphi-hairyB expression levels at the edges of the neural plate (asterisks). The expression in the dorsolateral walls of the archenteron is shown by arrows. (E) Dorsal view of an Amphi-hairyB 4-stripe pattern embryo. (F) Dorsal view of a three-stripe pattern embryo stained for the Amphi-hairyC probe. (G) Dorsal view of a four-stripe embryo stained for the Amphi-hairyD labelled probe. (H) Lateral view of the gastrula in G showing the Amphi-hairyD gene expression as pre-somitic stripes.
expressed in the dorsal portion of the anterior gut. *Amphi-hairyC* was conspicuously expressed as two patches at the immediate lateral to the midline anterior endoderm (arrows in Fig. 4K), whereas *Amphi-hairyD* mRNA was present in the dorsal-most part (arrow in Fig. 4S) and in two lateral domains of the anterior endoderm (double arrow in Fig. 4S). In a medial section, both genes were also similarly transcribed in the neural plate (arrowhead in Fig. 4L). *Amphi-hairyC* was also highly...
Amphi-hairyA, Amphi-hairyB, Amphi-hairyC and Amphi-hairyD genes expression in amphioxus larvae by whole-mount in situ hybridisation. Lateral views (A,D,F,J) have anterior towards the left and dorsal upwards. In dorsal and ventral views (I,K,N), anterior is towards the right. In transverse sections (B,C,E,G,H,L,M), dorsal is upwards looking from a posterior position. (A) Amphi-hairyA expression in the endoderm and the posterior dorsal nerve cord of an amphioxus larva. Scattered positive cells are in the middle gut are indicated by a double arrow and expression in the left anterior gut diverticulum is indicated by an arrow. (B) Cross-section through level i in A showing the expression of Amphi-hairyA in the preoral pit (arrow). (C) Cross-section through level ii in A showing the expression of Amphi-hairyA in the gut (arrow) and the posterior neural tube (asterisk). (D) Amphi-hairyB expression in the anterior half of the dorsal nerve cord. The gap in the anterior neural tube domain is shown by a double arrow. (E) Cross-section through level i in D showing the expression of Amphi-hairyB in the neural tube (asterisk). (F) Lateral view of an amphioxus larva showing Amphi-hairyC expression in the posterior tail bud, and in the anterior endoderm (arrow indicates the pharyngeal endoderm, and asterisk indicates the preoral pit). (G) Cross-section through level i in F showing the expression of Amphi-hairyC in pharyngeal endoderm (arrow). (H) Cross-section through level ii in F showing the expression of Amphi-hairyC in the posterior mesoderm. (I) Dorsal view of the specimen in F showing the Amphi-hairyC expression within the tail bud. (J) Lateral view of a larva showing Amphi-hairyD expression in the anterior club-shaped gland (asterisk) and the presumptive first and second gill slits (arrows), and in the posterior mesoderm. (K) Ventral view of the specimen in J. (L) Cross-section through level i in J showing the expression of Amphi-hairyD in the club-shaped gland (arrow). (M) Cross-section through level ii in J showing the expression of Amphi-hairyD in the posterior mesoderm. (N) Dorsal view of the specimen in J showing the Amphi-hairyD expression within the tail bud.

Interestingly, all four genes are similarly expressed within the segmented somites (arrows in Fig. 4L). In a posterior transverse section, another difference came to light. Although both genes where highly expressed in the forming somites, Amphi-hairyC was the only one expressed in the neural plate (arrowhead in Fig. 4M and asterisk in Fig. 4T). Moreover, whereas Amphi-hairyC is weakly expressed in the forming notochord (asterisk in Fig. 4M), Amphi-hairyD is expressed at high levels in this territory (arrow in Fig. 4T). During the late neurula stage, both genes are expressed in dorsal structures (Fig. 4N for Amphi-hairyC and Fig. 4V for Amphi-hairyD) with several differences. Amphi-hairyC signal is mainly seen in the neural tube (arrowhead in Fig. 4P), the gut, the segmented somites (arrows in Fig. 4P) and only at low levels within the dorsal notochord (asterisk in Fig. 4P). Amphi-hairyD signal is also detected in the neural tube (arrowhead in Fig. 4W) and in the gut (Fig. 4W, arrow). However, Amphi-hairyD signal is not detected within the segmented somites, but is conspicuous in the ventral notochord (asterisk in Fig. 4W). In a posterior section through level ii in Fig. 4N and 4V, the signal is detected in the forming somites for both genes (Fig. 4Q,X, respectively). However, Amphi-hairyD signal is stronger detected in the forming notochord (asterisk in Fig. 4X) than Amphi-hairyC signal (Fig. 4Q).

Interestingly, all four genes are similarly expressed within the neural tube of late neurulae as they where expressed in the neural plate of gastrula stages. That is, Amphi-hairyA being the Hairy gene expressed in a posterior-most domain of the neural tube (Fig. 4B), Amphi-hairyB being highly expressed in the anterior-most domain (Fig. 4H), and Amphi-hairyC and Amphi-hairyD similarly expressed in between (Fig. 4N,V).

**Amphi-hairyA, Amphi-hairyB, Amphi-hairyC and Amphi-hairyD genes expression in larvae**

During larval stages, Amphi-hairyA expression is similar to that observed in late neurulae. Its mRNA continues to be restricted to the posterior-most part of the dorsal nerve cord. Within the gut, Amphi-hairyA is strongly expressed in three regions (Fig. 5A): in the most posterior gut it has a two-domain pattern; in a middle region there are scattered positive cells (double arrow in Fig. 5A); and in the anterior gut it is expressed dorsally in the left anterior gut diverticulum that will become the Hatschek’s pit (arrow in Fig. 5A), in the anterior wall of the gut, and in the pharynx endoderm more ventrally. Cross-sections through level i in Fig. 5A show Amphi-hairyA expression in the left anterior gut diverticulum (arrow in Fig. 5B). A cross section through a more posterior level (ii in Fig. 5A) shows the signal within the posterior neural tube (asterisk in Fig. 5C) and the gut (arrow in Fig. 5C).

Amphi-hairyB expression is still restricted to the anterior part of the neural tube during larval stages (Fig. 5D), which is
better observed in a transverse section through level i in Fig. 5D (asterisk in Fig. 5E). There is still a gap between the anterior patch of Amphi-hairyB expression and the rest along the mid-anterior neural tube (double arrow in Fig. 5D).

In contrast to earlier stages, Amphi-hairyC is no longer detected in neural tissues. Its expression is mainly observed in the posterior paraxial mesoderm associated with the tail bud, and in the anterior endoderm, where it is conspicuously expressed in a region ventral to the mouth and the branchial anlage (arrow in Fig. 5F), and in the ventral part of the left anterior gut diverticulum or preoral pit (asterisk in Fig. 5F). A cross-section through level i in Fig. 5F shows the expression in the ventral pharyngeal endoderm (arrow in Fig. 5G), and in the posterior mesoderm through a more posterior level (Fig. 5H). In a dorsal view, the signal in the posterior tail bud is better observed (Fig. 5I).

Last, Amphi-hairyD mRNA is not longer present in the neural tube during larval stages. It is expressed only in the anterior endoderm and in the posterior tail bud, similar to Amphi-hairyC. In the anterior region of the larva, it is expressed in a three-stripe pattern. The first one marks the region where the ventral duct of the club-shaped gland is developing (asterisk in Fig. 5J), and the two posterior ones that are in the branchial anlage, may prefigure the first two gill slits (arrows in Fig. 5J). This pattern is clearly seen from a ventral view (Fig. 5K). A section through level i in Fig. 5J makes further visible Amphi-hairyD expression within the club-shaped gland (arrow in Fig. 5L), and in the posterior mesoderm through a more posterior level (Fig. 5M). It is also noticeable from a dorsal view, the signal within the entire tail bud, although a bit higher in its anterior part, the chordoneural hinge of the tail bud (Fig. 5N).

Amphi-hairyE to H expression

Although we tried extensively in situ whole-mount hybridisations with Amphi-hairyE and Amphi-hairyF probes, we were unable to detect any expression on amphioxus embryos. Similarly, we failed to detect these genes by RT-PCR using specific primers for each of the two genes, in conjunction with specific primers for Amphi-hairyC (Fig. 6A for Amphi-hairyE and Amphi-hairyC, and Fig. 6B for Amphi-hairy-F and Amphi-hairyC). As only single clones for Amphi-hairyE and Amphi-hairyF were isolated from gastrula and neurula embryonic libraries (see Table S2 at http://dev.biologists.org/supplemental/), we concluded that either they are expressed during embryogenesis at levels which are very low to be detected with whole-mount in situ hybridisation or RT-PCR or they are present due to transcription leakage. We also performed RT-PCR experiments on adult mRNA and were unable to detect their expression (data not shown).

Amphi-hairyG and Amphi-hairyH were never found in the cDNA screenings or EST projects, and were isolated from genomic libraries. Therefore, we carried out similar RT-PCR experiments as with Amphi-hairyE and Amphi-hairyF. Again, and using Amphi-hairyC as internal control, neither Amphi-hairyY (Fig. 6C) nor Amphi-hairyR (Fig. 6D) expression was detectable during embryogenesis. RT-PCR on adult mRNA also gave negative results (not shown). Thus, we conclude that Amphi-hairyG and Amphi-hairyH if at all, are expressed at undetectable levels in the stages analysed.

Discussion

The amphioxus Hairy family

We report here the characterisation of the unexpectedly large Hairy subfamily in amphioxus (Branchiostoma floridae). It is composed of a minimum of eight genes, four of which are expressed at different territories during B. floridae embryogenesis. Although other cases of gene duplication have been reported in amphioxus (Minguillon et al., 2002), to our knowledge, this is the first report of such an extreme case of duplication. Moreover, this is also the first case in which a differential fate after duplication can be asserted for the duplicate copies. In other cases of duplicate genes in amphioxus, only two copies have been reported and there are two reports of three copies [calmodulin (Karabinos and Bhattacharya, 2000) and PTPR4 tyrosine phosphatase (Ono-Koyanagi et al., 2000)], and either the expression of both copies was too similar to distinguish as in the case of the brachyury genes (Holland et al., 1995), or only quantitative differences accounted for their expression patterns as in the case of HNF3 (Shimeld, 1997).

All eight amphioxus Hairy genes are closely related to both vertebrate and protostome Hairy genes (Fig. 1B; see Fig. S1 at http://dev.biologists.org/supplemental/). The existence of multiple copies of Hairy genes in the amphioxus genome is not in conflict with the ‘pre-duplicative’ state of its genome. All the copies put together would constitute the pro-orthologue of vertebrate Hairy genes (Fig. 1B) and each Amphi-Hairy gene is a trans-orthologue of each vertebrate gene, which means that they have arose by independent duplication in the B. floridae genome. This is enhanced by the finding that at least six of the Hairy genes are closely linked in the genome, suggesting their origin by tandem duplication. The fact that the cephalochordate
lineage predated the massive gene/genome duplications at the origins of vertebrates (Spring, 2002) does not imply that its genome is not evolving, meaning that it can undergo specific gene duplications, gene losses, etc. Amphioxus may resemble the ancestor of the vertebrates, but it is not the ancestor, only its closest living relative, a privileged position that did not include the freezing of its genome.

**Differential fate after duplication of the amphioxus Hairy genes: the DDC model**

Several mathematical models have been developed to explain the future of paralogue genes after duplication from a single ancestral gene. These models predict that duplicate genes initially have fully overlapping, redundant functions, such that one copy may shield the second from natural selection, if gene dose is not crucial. Because deleterious mutations occur more frequently than beneficial ones (Lynch and Walsh, 1998), the classical model predicts that one of the duplicate loci should most commonly deteriorate into a pseudogene (Watterson, 1983). The classical model also considers a rarer alternative: maintenance of duplicate copies, owing to the fixation of a rare beneficial mutation in one copy that endows it with a novel function, while the other maintains the original role (Ohno, 1970). The DDC model (for duplication-degeneration-complementation) was proposed by Force et al. (Force et al., 1999) as a possible explanation for the higher maintenance of duplicate genes in duplicated genomes observed than that expected under the classical mathematical models. According to this model, duplicate copies of an ancient single gene preserve their maintenance in the genome by subfunctionalisation, i.e. by differential degeneration of regulatory regions among the duplicate genes. Hence, paralogue genes decouple a pleiotropic role that was carried out by their single ancestor gene and become thus all copies necessary for carrying out all the functions that the ancestral pre-duplicate gene carried out.

Amphi-hairyA, Amphi-hairyB, Amphi-hairyC and Amphi-hairyD genes have expression patterns that generally do not overlap, although some of the differences are subtle. Briefly, Amphi-hairyA is mainly expressed within the endoderm and to a lesser extent at the posterior neural tube, Amphi-hairyB in the neural tube, posterior paraxial mesoderm and somites, and Amphi-hairyC and Amphi-hairyD are more widely expressed in all three embryonic layers. There are several cases of combinatorial patterns among the Amphi-Hairy genes all along amphioxus development. More noticeable are those within the neural plate first and the neural tube latter, and those in some anterior endoderm-derived structures in amphioxus larvae. The single mouse gene Hes1 is expressed all along the neural tube, but strikingly, only the summation of expression for all four Hairy genes of amphioxus covers the entire amphioxus neural plate and neural tube. Amphi-hairyB is expressed at the very front of the animal, Amphi-hairyA at the posterior, and Amphi-hairyC and Amphi-hairyD genes in between. It is also interesting that a combination of two Hairy genes (Amphi-hairyA plus Amphi-hairyB) is required to include all the cells of Hatschek’s left diverticulum, a suspected homologue of the vertebrate adenohypophysis. By contrast, zebrafish her9 is expressed in the whole pituitary. Hence, if we assume that this subfunctionalisation has originated by differential losses of cis regulatory sequences, one must be struck by the high degree of complexity in the regulatory regions of Hairy genes. Only for these two examples, we have to think in at least three distinct regulatory elements driving the expression of an ancestral pre-duplicative Hairy gene within the neural tube, and at least two driving the expression in the left anterior diverticulum, that have been differentially lost by the duplicate copies during or after the process of duplication. Fig. 7A summarises Amphi-hairyA to Amphi-hairyD expression data. This points out for a complex organisation of Hairy regulatory sequences. Based on the partitioned expression of Amphi-Hairy genes, distinct regulatory elements must be invoked to account for particular places of expression, which have been individualised in amphioxus duplicates. A plausible schematic representation of territorial enhancers deduced from Amphi-hairyA-Amphi-hairyD expression data is shown in Fig. 7B.

The Hairy genes in *B. floridae* seem to have undergone different processes of fate determination after duplication from a single gene in its lineage. As shown above, Amphi-hairyA to Amphi-hairyD genes have undergone subfunctionalisation, a
process ensuring their maintenance in the genome as duplicate genes. Conversely, Amphi-hairyE to Amphi-hairyH may be in the process of nonfunctionalisation, as we have been unable to detect their expression, or being expressed at very low levels and/or in particular instants that escaped the analyses.

Insights into amphioxus early somitogenesis from Amphi-Hairy genes expression

All three amphioxus Hairy genes that are expressed in the presumptive somitic mesoderm (Amphi-hairyB, Amphi-hairyC and Amphi-hairyD) shed light onto the long-standing debate regarding the formation of the first four pair of amphioxus muscular somites. These somites have the peculiarity that they bud off virtually simultaneously from the dorsolateral walls of the archenteron. Hatschek (Hatschek, 1893) claimed to have observed embryos with a single pair of somites, whereas Conklin (Conklin, 1932) never detected them and claimed that always more than one somite pair was present in all the embryos he analyzed. Regardless of their simultaneous or sequential anteroposterior appearance, they are molecularly prefigured one by one in the dorsolateral wall of the archenteron, as we have detected gastrulae with a two- to three-stripe, three-stripe and four-stripe pattern (Fig. 3). Moreover, the maturation of those four first muscular somites also appears as a sequential anteroposterior process, as the intensity of Amphi-hairyB, Amphi-hairyC and Amphi-hairyD expression is stronger in posterior (and thus younger) pre-somites than in the anterior (older) ones (Fig. 3).

Similarities between amphioxus and vertebrate Hairy genes

Within the CNS, the mouse gene Hes1 is expressed in undifferentiated neuronal precursor cells in the ventricular zone, and its transcription decreases as neurogenesis proceeds until expression is no longer detected in mature neurons or glial cells (Sasai et al., 1992). Accordingly, mutant mice for the Hes1 gene exhibit severe neural defects. Moreover, in their brain there is an upregulation of some neural bHLH factors and postmitotic neurons appear prematurely. It thus appears that Hes1, like the Drosophila hairy gene, acts as a negative regulator of neurogenesis, and that its downregulation is required for precursor cells to enter the differentiation processes (Ishibashi et al., 1995). It seems reasonable to think that amphioxus Hairy genes are carrying out a similar function within the amphioxus nerve cord. In the mouse, Hes1 is expressed all along the neural tube, which corresponds to the summation of the expression patterns of all four Amphi-Hairy genes. In larval stages, only the anterior- and the posterior-most parts of the neural tube are positive for a Hairy gene (Amphi-hairyB, Fig. 5D, and Amphi-hairyA, Fig. 5A, respectively). It is tempting to speculate that these differences may account for differential maturation rates along the neural tube. In the zebrafish, the Hairy orthologue her9 is expressed in the mid- and hindbrain but not in the midbrain-hindbrain boundary (MHB) (Leve et al., 2001). It is striking therefore to note a gap also in the anterior amphioxus neural tube (Amphi-hairyB expression in Figs 4, 5). The existence of a tripartite brain in cephalochordates is still debatable (Ferrier et al., 2001) and the gap of expression in this region suggests the presence of a MHB.

In endodermal derivatives, Hes1 is expressed in, and required for, the proper development of the endocrine islet cells of the mouse pancreas as well as other dispersed endocrine cells along the entire gut. It was suggested that Hes1 functions as a general negative regulator of endodermal endocrine differentiation, similar to its action within neural precursors (Jensen et al., 2000). Similarly, the amphioxus Hairy genes are expressed in the developing gut and certain derivatives in the larva (Amphi-hairyA, Amphi-hairyC and Amphi-hairyD). Amphioxus does not have a discrete pancreas but has several types of endocrine cells incorporated into the gut epithelium, some of which are possibly homologous of the pancreas-islet cells of mammals (Holland et al., 1997b). Interestingly, patches of Amphi-hairyA expression in the gut may represent regions with presumptive endocrine cell types. All Hairy genes but Amphi-hairyB are also conspicuously expressed in combinatorial patterns in some endoderm-derived glands. Briefly, in the left gut diverticulum, Amphi-hairyA is predominantly expressed in the dorsoposterior region and Amphi-hairyC in its ventral region, whereas Amphi-hairyD is conspicuously expressed in the club-shaped gland. Hatschek’s left gut diverticulum contributes to the Hatschek’s pit in the adult, a structure thought to be homologous to the vertebrate adenohypophysis (Whittaker, 1997), an organ where the zebrabash Hairy gene her9 is expressed (Leve et al., 2001).

In summary, the addition of the expression patterns of Amphi-hairyA to Amphi-hairyD genes closely resembles the expression of the single mouse Hes1 gene or the multiple Hairy genes in other vertebrate species. They are expressed in the CNS (all of them in complementary patterns), in the posterior paraxial mesoderm (Amphi-hairyB, Amphi-hairyC and Amphi-hairyD), in the posterior compartment of the segmented somites (Amphi-hairyB and Amphi-hairyC), in the gut (Amphi-hairyA always and Amphi-hairyC and Amphi-hairyD in neurola and larval stages), and in the notochord (Amphi-hairyD). Hence, the common ancestor of cephalochordates and vertebrates already possessed a single Hairy gene of a very pleiotropic nature, in contrast to protostome Hairy genes. This pre-duplicative gene at the origins of vertebrates already functioned in similar territories as it does in extant vertebrates. In the amphioxus lineage, the highly pleiotropic gene suffered duplication and distinct regulatory regions were most probably maintained in particular copies.

Do the Amphi-Hairy genes cycle within the amphioxus posterior paraxial mesoderm?

There is not a single technique available to demonstrate whether one (or more) amphioxus Hairy genes are cycling within the PSM as they do in amniotes. Amphioxus somitogenesis is divided in two different phases. During an earlier phase, the somites originate by the budding off of the dorsolateral walls of the archenteron (which could be considered a sort of PSM) forming their coeloms by enterochoely. The first eight pairs of muscular somites are formed during this early phase from paraxial mesoderm formed during gastrulation (Holland et al., 1997a). By contrast, during the second phase somites arise directly from the proliferative tail bud by a schyzocoelic process (Schubert et al., 2001) without the intervention of any visible PSM between the tail bud and the nascent somites. Hence, if any Hairy gene has a cycling behaviour, it would only be visible during the early phase of amphioxus somitogenesis or in the tail bud itself. Only Amphi-hairyB, Amphi-hairyC and Amphi-
Amphi-Hairy genes expression within the paraxial mesoderm development may had happened. Besides, we have detected was not the case, although again masking of cycling due to fast blastopore, the equivalent structure to the primitive streak. This to detect their ‘on and off’ expression within or around the accordingly to this cyclic behaviour, we should have been able (Jouve et al., 2002). If any Hairy gene is being regulated correlated with their position along the anteroposterior axis (her6) and (Amphi-hairyD) as stripes that prefigure the first muscular somites. Moreover, we saw gastrulae with two to three, three or four stripes (Fig. 3), which may indicate that they are formed one by one, in the anterior PSM, similar to Hairy genes in lower vertebrates such as cycling in the zebrafish "herb" gene or both "Xenopus" Hairy genes. The expression of these genes is seen as one to three stripes at the anterior PSM that prefigure the regions where new somites will be added (Pasini et al., 2001; Davis et al., 2001). Hence, we suggest that the cycling behaviour of the Hairy subfamily may be used to study this signalling pathway, such as the Notch ligand DeltaC, cycle in the zebrafish PSM (Jiang et al., 2000). Thus, it is possible that other genes related to the Notch signalling pathway or other genes from the E(spl) family cycle in the amphibian PSM.

We are indebted to Jim Langeland for cDNA library and Chris Amemiya for the PAC library, to the Serveis Científico-Tècnics of the Universitat de Barcelona for the sequencing, and to Elisa Serra and Elisa Bundke for technical help. We specially thank Nic Holland for unvaluable help with in situ hybridisation analyses of highly synchronous embryos.

In the chicken, the onset of the dynamic expression of Hairy genes correlates with ingression of the paraxial mesoderm territory from the epiblast into the primitive streak. Hence, the number of oscillators experienced by somitic cells is correlated with their position along the anteroposterior axis (Jouve et al., 2002). If any Hairy gene is being regulated according to this cyclic behaviour, we should have been able to detect their ‘on and off’ expression within or around the blastopore, the equivalent structure to the primitive streak. This was not the case, although again masking of cycling due to fast development may had happened. Besides, we have detected Amphi-Hairy genes expression within the paraxial mesoderm (Amphi-hairyB, Amphi-hairyC and Amphi-hairyD) as stripes that prefigure the first muscular somites. Moreover, we saw gastrulae with two to three, three or four stripes (Fig. 3), which may indicate that they are formed one by one, in the anterior PSM, similar to Hairy genes in lower vertebrates such as the zebrafish "herb" gene or both "Xenopus" Hairy genes. The expression of these genes is seen as one to three stripes at the anterior PSM that prefigure the regions where new somites will be added (Pasini et al., 2001; Davis et al., 2001). Hence, we suggest that the cycling behaviour of the Hairy subfamily may be an amniote novelty. However, as discussed above, a cycling behavior of amphibious Hairy genes cannot be totally discarded. Alternatively, other genes belonging to the E(spl) subfamily may be cycling in amphibious, as "her1" and "her7" in zebrafish (Holley et al., 2000; Leve et al., 2001). In addition, other non-bHLH genes related to the Notch signalling pathway, such as the Notch ligand DeltaC, cycle in the zebrafish PSM (Jiang et al., 2000). Thus, it is possible that other genes related to the Notch signalling pathway or other genes from the E(spl) family cycle in the amphibious PSM.


