Developmental Biology 327 (2009) 252–262

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/developmentalbiology

Evolution of Developmental Control Mechanisms

Differential regulation of ParaHox genes by retinoic acid in the invertebrate chordate amphioxus (*Branchiostoma floridae*)

Peter W. Osborne [a,1], Gérard Benoit [b], Vincent Laudet [b], Michael Schubert [b,2], David E.K. Ferrier [a,*,1,2]

[a] *Zoology Department, Oxford University, South Parks Road, Oxford, OX1 3PS, UK*
[b] *Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, CNRS, INRA, Université Claude Bernard Lyon 1, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07, France*

ARTICLE INFO

Article history:
Received for publication 2 October 2008
Revised 19 November 2008
Accepted 19 November 2008
Available online 7 December 2008

Keywords:
Amphioxus
Retinoic acid
Gsx
Xlox
Cdx
ParaHox
Gene regulation
Evolution
Endoderm patterning

ABSTRACT

The ParaHox cluster is the evolutionary sister to the Hox cluster. Like the Hox cluster, the ParaHox cluster displays spatial and temporal regulation of the component genes along the anterior/posterior axis in a manner that correlates with the gene positions within the cluster (a feature called collinearity). The ParaHox cluster is however a simpler system to study because it is composed of only three genes. We provide a detailed analysis of the amphioxus ParaHox cluster and, for the first time in a single species, examine the regulation of the cluster in response to a single developmental signalling molecule, retinoic acid (RA). Embryos treated with either RA or RA antagonist display altered ParaHox gene expression: *AmphiGsx* expression shifts in the neural tube, and the endodermal boundary between *AmphiXlox* and *AmphiCdx* shifts its anterior/posterior position. We identified several putative retinoic acid response elements and in vitro assays suggest some may participate in RA regulation of the ParaHox genes. By comparison to vertebrate ParaHox gene regulation we explore the evolutionary implications. This work highlights how insights into the regulation and evolution of more complex vertebrate arrangements can be obtained through studies of a simpler, unduplicated amphioxus gene cluster.

Introduction

Within chordates, a single intact ParaHox cluster of three genes has been conserved in tetrapods and amphioxus, whose lineages separated over 500 million years ago (Brooke et al., 1998; Coulier et al., 2000; Ferrier et al., 2005). The three ParaHox gene families, Gsx (or Gsh or ind), Xlox (or Xlhbox8, IPF1, PDX1, IDX1, STF1 or Lox3) and Cdx (or caudal/cad) are widespread within Bilateria (Ferrier and Holland, 2001a) and at the origin of chordates the ParaHox genes were linked in a tight cluster with Xlox located between Gsx and Cdx (Brooke et al., 1998; Finnerty and Martindale, 1999; Ferrier and Holland, 2001b). Unlike the single cluster of amphioxus however, tetrapods also have several extra ParaHox genes (an extra Gsx and two extra Cdx genes) as a result of two rounds of whole genome duplications followed by gene loss. Teleost fish in contrast, have disrupted the structure of their ParaHox cluster because of an extra round of whole genome duplications and subsequent gene loss, although the overall number and complement of vertebrate ParaHox genes has been retained in the teleost lineage (Mulley et al., 2006; Prohaska and Stadler, 2006; Siegel et al., 2007). The single hagfish ParaHox cluster is also in the process of degeneration, with Xlox mutated to a pseudo-gene (Furlong et al., 2007). Urochordates and echinoderms have not undergone genome duplications but have also disrupted their ParaHox cluster, due to rearrangement of their respective genomes. These invertebrate deuterostome lineages have thus retained all of their ParaHox genes, but not the ParaHox cluster, which contrasts with the mechanism of cluster disruption via gene loss in some vertebrates (Ferrier and Holland, 2002; Arnone et al., 2006). Although widely conserved, the ParaHox cluster, like its paralogue the Hox cluster, is therefore not completely immune to disruption (Ferrier and Minguillón, 2003).

Outside of the chordates, lophotrochozoans have representatives of all three genes (Ferrier and Holland, 2001a) while only representatives of the Cdx and Gsx families have been identified so far in ecdysozoans. In both groups no intact cluster has yet been identified. Overall, the ancestral bilaterian clearly had three ParaHox genes that were most likely arranged in a cluster similar to that maintained by amphioxus. However, it remains a mystery as to why some groups have retained a ParaHox cluster over long periods of time, while other groups have not. Either a selective constraint has kept the ParaHox cluster together, for example through shared *cis*-regulation (Mulley et al., 2006; Duboule, 2007), or alternatively they have simply been retained as an evolutionary relic via slow genomic evolution in some lineages. The single ParaHox cluster of amphioxus

* Corresponding author.
E-mail address: dekf@st-andrews.ac.uk (D.E.K. Ferrier).
[1] Current address: Gatty Marine Institute, University of St. Andrews, St. Andrews, Fife, KY16 8LB, Scotland, UK.
[2] These authors contributed equally to this paper.

0012-1606/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2008.11.027

constitutes a good system to investigate these possible regulatory mechanisms because it will avoid the problem of functional redundancy between duplicates inherent with studying vertebrate ParaHox genes (Schubert et al., 2006a). By identifying amphioxus regulatory mechanisms conserved with other chordates we can also reconstruct the ancestral regulation of the cluster and begin to answer questions on how and why the ParaHox cluster has been maintained in certain lineages.

Retinoic acid (RA) is a derivative of vitamin A (retinol) involved in regulating chordate anterior/posterior (A/P) patterning (Mark et al., 2006; Marlétaz et al., 2006; Maden, 2007). RA function is mediated by retinoic acid receptor (RAR) and retinoid X receptor (RXR) heterodimers binding to specific regulatory DNA elements (called RA response elements or RAREs) and activating transcription in a ligand-dependent manner (Mark et al., 2006). RAREs consist of two direct repeats (DR), with the canonical nucleotide sequence (A/G)G(G/T)TCA, that are usually separated by either 2 or 5 nucleotide spacers (Bastien and Rochette-Egly, 2004). Compared with Hox genes, there are relatively few examples of RA regulating ParaHox genes in chordates and the ParaHox cluster as a whole has not been examined.

Currently there are no reports of vertebrate Gsx genes responding to RA but there are several studies examining the interaction between RA and the other two ParaHox genes. In mice, *Cdx1* is directly regulated by RA and RA treatment causes an early induction and a posterior expansion of *Cdx1* (Houle et al., 2000). The direct regulation of *Cdx1* by RA signalling is partially mediated by an atypical DR5-type RARE located upstream of the *Cdx1* gene (Houle et al., 2000, 2003; Lickert and Kemler, 2002; Pilon et al., 2007). Interestingly, from *Cdx1* RARE null mutants a second RA signalling pathway was identified and found to regulate *Cdx1* (Houle et al., 2003), probably through a DR2-type RARE conserved in the mouse and chicken *Cdx1/CdxA* introns (Gaunt et al., 2003). In contrast to these results in mouse and chicken, one of the two zebrafish *Cdx1* genes is not regulated by RA (at least not in the tissues or stages studied) (Stafford and Prince, 2002). In addition, the other two mouse Cdx paralogues are not induced by RA (Roelen et al., 2002). Instead, excess RA represses *Cdx4* anteriorly (Iulianella et al., 1999) and may subtly reduce *Cdx2* expression (Roelen et al., 2002). Finally, recent work indicates zebrafish Cdx4 may actually prevent RA signal transduction in posterior endoderm (Kinkel et al., 2008). This variation of Cdx responses to RA signalling indicates the regulation of these genes has evolved since the ParaHox cluster duplicated, making an analysis of the unduplicated amphioxus Cdx gene critical for understanding the ancestral regulation of the gene.

In RA-depleted embryos, Xlox expression is abolished in the dorsal pancreatic anlage in zebrafish (Stafford and Prince, 2002) and mice (Martin et al., 2005; Molotkov et al., 2005). Conversely, maternal replacement of RA in deficient mice allows recovery of Xlox expression (Martin et al., 2005; Molotkov et al., 2005). Furthermore, excess RA also causes up-regulation of Xlox expression in mouse cell lines (Tulachan et al., 2003; Micallef et al., 2005). RA has also been demonstrated to induce ectopic Xlox expression in chick/quail chimera experiments (Kumar et al., 2003).

In this report, we aimed to examine the evolution of RA regulation of ParaHox genes in chordates. We undertook a comprehensive examination of all three ParaHox genes from the Florida amphioxus, *Branchiostoma floridae*, and assessed the response to exogenous RA and RA antagonist. The treatment of amphioxus with RA has been well characterised in the past. Embryos treated with excess RA at gastrulation display a stereotypical change in development (Holland and Holland, 1996). The most distinctive changes are the loss of mouth and gill slits and a massive reduction of the pharyngeal endoderm. This is despite a lack of neural crest in amphioxus which was long believed to be the main mediator for the effects of RA in the vertebrate branchial arches (Escriva et al., 2002). In a similar manner to vertebrates, gene expression is also dramatically altered in amphioxus treated with RA: for instance Hox genes are shifted anteriorly by RA in the nerve cord (Holland and Holland, 1996; Schubert et al. 2006b). In contrast, the development of the notochord and musculature of amphioxus is not clearly affected by the addition of RA. Moreover, while *AmphiRXR* is weakly expressed throughout the amphioxus embryo (Escriva et al., 2002), *AmphiRAR* is expressed throughout the mesendoderm at gastrula stages and by neurula stages is widely expressed in the posterior neural plate and most mesendodermal derivatives (Escriva et al., 2002). During subsequent development, RAR is downregulated anteriorly and posteriorly and is restricted to more central regions of the nerve cord, somites and endoderm (Escriva et al., 2002). Thus, RA is likely to be acting in locations relevant to the expression of the amphioxus ParaHox genes.

We provide a full description of the expression of amphioxus ParaHox genes in embryogenesis and early larval development, revealing a novel expression domain of *AmphiGsx* that has not previously been documented. For the first time in a single species we show the response of all three ParaHox genes to a single signalling molecule. All three genes are regulated by RA, with RA causing an A/P shift in some expression domain boundaries. In addition, *AmphiCdx* and *AmphiXlox* share a RA-sensitive boundary of expression in the posterior endoderm. We have also begun to establish whether this RA regulation is via a direct mechanism by identifying putative RAREs in the amphioxus ParaHox cluster.

Materials and methods

Embryology, in situ hybridisation, and microscopy

Embryos of the Florida amphioxus (*B. floridae*) were treated at the very late blastula stage with 1×10^{-6} M RA or RA antagonist (BMS009) diluted in DMSO or with DMSO alone (as described previously in Escriva et al., 2002; Schubert et al. 2006b). In situ hybridisations were performed essentially as described elsewhere (Holland et al., 1996). Staining was conducted at 4 °C. The *AmphiCdx* clone (pC15.1) was a kind gift from Jordi Garcia-Fernàndez, initially cloned from an amphioxus cDNA library (Langeland et al., 1998). First and second exons of *AmphiGsx* and *AmphiXlox* were amplified from the two amphioxus ParaHox PACs (Genbank accession numbers AC129948 and AC129947) and subsequently fused by PCR. Total sizes of the clones were: *AmphiXlox* 1174 bp and *AmphiGsx* 1018 bp. Primers used to clone *AmphiGsx* and *AmphiXlox* were: BfGsx5′-F GTCGAACGCCTTTGTGAAGT; BfGsx5′-R TGATGCCACTCCAGAGGG-GAAAGGTAGG; BfGsx3′-F TCCCCTCTGGAGTGGCATCAGATGGTC; BfGsx3′-R TACGACAACGCAAAGTAACG; BfXlox5′-F TTCAAACGA-TACCGGACAAAC; BfXlox5′-R AAGGACGCACCACCTGGCCATTGAGAC; BfXlox3′-F GGCCAGGTG-GTGCGTCCTTTGCTGTTG; BfXlox3′-R ATGAAAAACCACCTGCGTTG. Stained amphioxus embryos were visualised on a Zeiss Axioskop2 microscope. The percentage of body length expressing each gene was calculated from measurements of gene expression domains and embryo length using Axiovision 4. One-way ANOVA tests were undertaken on raw (cells numbers expressing *AmphiGsx*) or log transformed data (relative lengths of expression of *AmphiXlox* or *AmphiCdx*) to determine if there were significant differences between treatments at each stage. See Supplementary Tables 2–4 for numbers of embryos used in the analysis.

Identification of RAREs

NHR scan (Sandelin and Wasserman, 2005) identified four DR5 and two DR2 element sequences in the amphioxus ParaHox cluster (Supplementary Table S1). These six putative RAREs were further analysed, as was a putative DR5 element located upstream of *AmphiGsx* that was identified by a manual search (using Gene Palette (Rebeiz and Posakony, 2004)) for the consensus sequence (A/G)G(G/T)TCA[2/5](A/G)G(G/T)TCA (Perlmann et al., 1993).

Electrophoretic mobility shift assays (EMSAs)

EMSAs were carried out as previously described (Vanacker et al., 1999). Complementary primers (Supplementary Table S1) corresponding to the putative RAREs (plus flanking sequences) with 5′ overhanging restriction sites were annealed and labelled with γ^{32}P-ATP. Radiolabelled RAREs were bound to the amphioxus RAR and RXR proteins at 4 °C (in Hepes, 2 mM; NaCl, 50 mM; KCl, 50 mM; $MgSO_4$, 3 mM; glycerol, 10%; DTT, 2 mM; poly dIdC, 0.04 μg/μL) and run on a 6% acrylamide gel, followed by gel drying and signal detection.

Transactivation assays

Reporter construct cloning and cell culture experiments were performed as described (Wang et al., 2003). Complementary primers (Supplementary Table S1) corresponding to the putative RAREs (plus flanking sequence) with 5′ overhanging restriction sites were annealed and cloned into the pGL2 promoter vector (Promega). RARE island 1 (7702 bp) was PCR amplified using the primers RI1F-TTCTTTCGGCGTCGTTATTC and RI1R-CTCGAGAACCGCAAGAGCAACAC which are 1163 bp upstream and 2033 bp downstream of DR5b and DR5c respectively. RARE island 2 (7061 bp) was PCR amplified using the primers RI2F-GGATCCCGCAATAGATACGTCAACA and RI2R-GGATCCCACCAGCTCACCCTAAA which are 363 bp upstream and 1482 bp downstream of DR5e and DR2b respectively. Both PCR fragments were subcloned into the pGL2 promoter vector upstream of a minimal promoter element and the luciferase reporter gene to assay possible activities as enhancer elements. The human embryonic kidney cell line 293 was grown in DMEM+10% foetal calf serum and antibiotics in 2 cm^2 wells to 80–90% confluence. All cells were co-transfected with a DNA solution of vector containing the RARE construct to be tested plus a plasmid encoding β-galactosidase as a baseline control for transfection efficiency between treatments. Transfection was conducted in the presence of lipofectamine and Plus Reagent (Invitrogen) but in the absence of foetal calf serum. In addition, controls were transfected with either *AmphiRAR* (in pSG5), *AmphiRXR* (in pSG5) or empty vector alone, whereas experimental treatments contained both *AmphiRAR* and *AmphiRXR* constructs. Cells were transfected for 5 h and then foetal calf serum was added to all cells, while all-*trans* RA at a final concentration of 5×10^{-7} M was added to half of the wells. After 24 h, levels of β-galactosidase and luciferase activity were measured.

Results

Wild-type expression of amphioxus ParaHox genes

To examine the regulation of amphioxus ParaHox genes, we first undertook a detailed description of the endogenous gene expression. *AmphiCdx* is the first ParaHox gene to be expressed. It initiates during mid-gastrulation in a ring around the blastopore and this expression continues through blastopore closure (Fig. 1L). As neurulation begins, expression is strongest dorsally in the neuroectoderm, but there is also ventral and lateral expression throughout the posterior of the embryo in ectoderm and to a lesser extent in ventral endoderm (Fig. 1M). During early stages of somitogenesis, expression is present in the neuroectoderm, the walls of the posterior archenteron and in the tail bud region in a continuous domain (Fig. 1N). Neuroectoderm expression is strong posteriorly and weaker at more anterior levels, such that *AmphiCdx* expression in the neuroectoderm is graded from posterior to anterior. Moreover, the anterior limit of expression is generally more rostral in the neuroectoderm than in the archenteron. This difference in A/P levels is maintained in late embryos (Figs. 1O, P), although after 1 week of development neural expression is down-regulated anteriorly, hence reversing these A/P level differences (Fig. 1Q). Throughout development there is expression in the tail bud, and the posterior elongation of the embryo results in expansion of expression so that *AmphiCdx* is also expressed throughout the hindgut and posterior neural tube until at least 7 days post-fertilisation, the oldest stage examined (Figs. 1M–Q).

The earliest *AmphiXlox* expression was observed after gastrulation in the ventral posterior archenteron wall at low levels (data not shown). This domain subsequently strengthens and at approximately the same A/P position a weaker domain develops dorsally (Fig. 1E). This dorsal expression subsequently increases in intensity and spreads to include the neuroectoderm and mesendoderm (Fig. 1F). During neurulation, expression is initiated posteriorly in the embryo, linking the dorsal and ventral domains. There is also strong *AmphiXlox* expression in two cells of the neural tube approximately level with the anterior boundary of somite five (where the first pigment spot will form) (Figs. 1G, H). In late neurulae, *AmphiXlox* expression is maintained in the neural tube and the posterior expression in the embryo separates again into two distinct domains. The first is a ventral domain in the posterior gut, while the second is located more posteriorly in the dorsal mesendoderm close to the tail bud (Fig. 1I). In early larvae, neural expression is down-regulated and the mesendodermal domain lost. Larval growth causes an A/P expansion of *AmphiXlox* endodermal expression so that it now marks a specific domain in the posterior midgut/anterior hindgut (Fig. 1J). This distinctive domain is maintained until at least 7 days post-fertilisation with strong expression at its anterior limit, which weakens posteriorly (Fig. 1K).

AmphiGsx is the last gene to be activated, and it is expressed in a more restricted pattern. There are two temporally distinct domains of *AmphiGsx* expression, an 'early' and a 'late' domain. The early domain arises during neurulation and consists of four cells in the neural tube, level with somite five (Figs. 1B, C). This early *AmphiGsx* domain has not previously been identified and overlaps with *AmphiXlox* neural expression, although *AmphiGsx* is expressed in slightly more cells (compare Figs. 1C and H). In late embryos, a transient expression domain is detectable in the most anterior part of the neural tube, the cerebral vesicle (CV) (Fig. 1D). Expression is in the centre of the CV (along the A/P axis) on both left and right sides and remains until early larval stages (Fig. 1D inset). This detailed description of amphioxus ParaHox expression patterns extends and improves the initial description provided by Brooke et al. (1998).

RA changes amphioxus ParaHox expression in the endoderm

To assess whether RA regulates amphioxus ParaHox gene expression we examined the expression of all three ParaHox genes in amphioxus embryos treated with either RA, RA antagonist (BMS009) or DMSO (control). For both *AmphiCdx* and *AmphiXlox*, the earliest effects of RA and BMS009 treatments that we observe are detectable in the late embryo endoderm (Figs. 2, 3). At this stage, a significant response to varying RA signalling levels for both *AmphiCdx* (one-way ANOVA, $p=0.001$) and *AmphiXlox* (one-way ANOVA, $p<0.0001$) is observed in the posterior endoderm. These alterations remain significant during subsequent development (Figs. 2, 3). Both genes respond in a similar but complementary manner to changing RA signalling levels (Figs. 2, 3). The anterior limit of endodermal *AmphiCdx* expression in RA-treated embryos is shifted posteriorly, and consistent with this response, the anterior limit of expression is shifted rostrally in BMS009-treated embryos (Fig. 2). Interestingly, it is the posterior limit of *AmphiXlox* that is changed in response to RA and BMS009 rather than the anterior boundary (Fig. 3). In RA-treated amphioxus embryos, *AmphiXlox* is expanded posteriorly, and BMS009 treatment moves the posterior expression limit rostrally.

From these complementary shifts in expression, it appears RA signalling influences the boundary of expression between the two ParaHox genes. In an attempt to elucidate the details of this shift, we

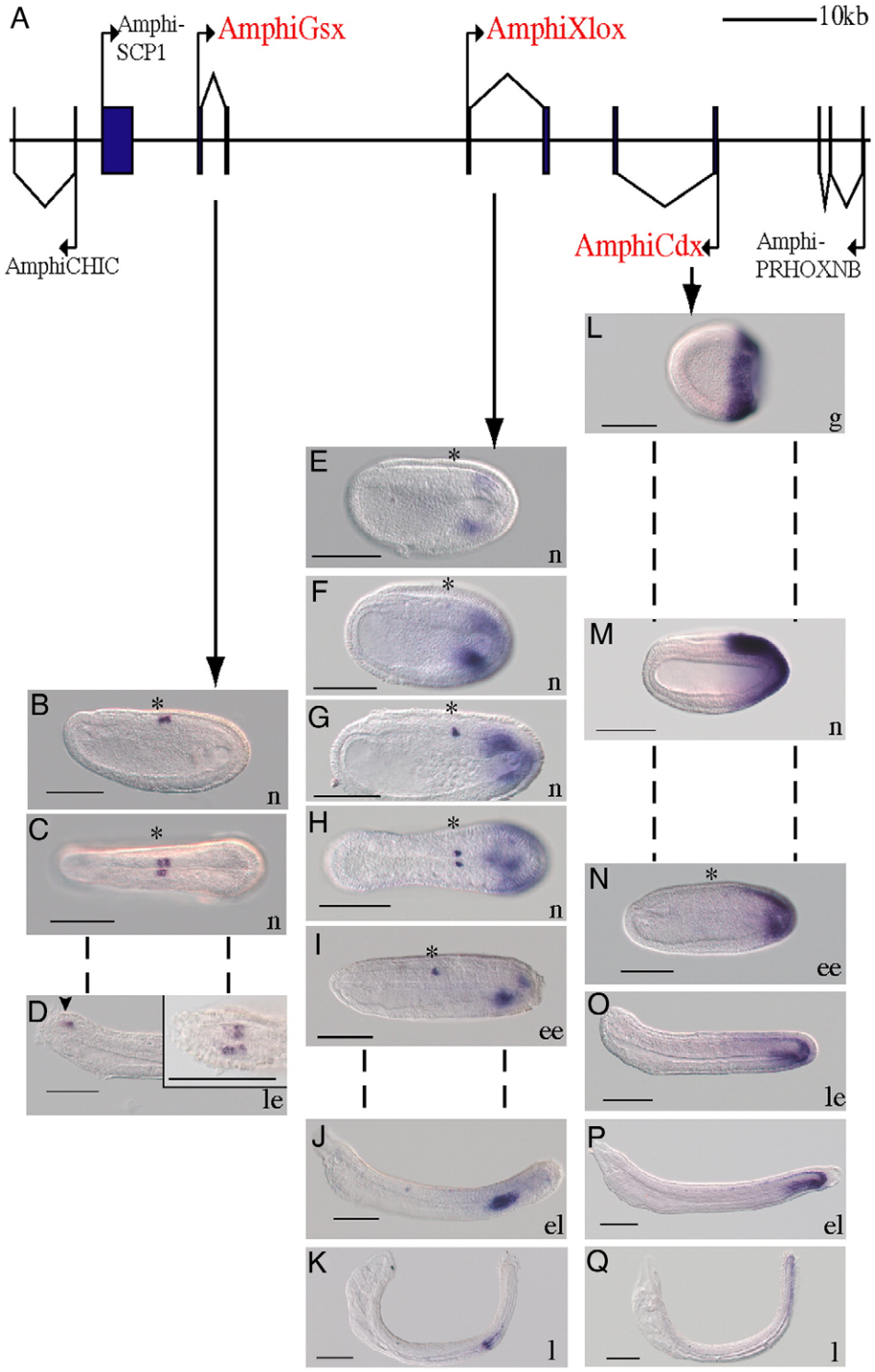

Fig. 1. Wild-type expression of the amphioxus ParaHox genes is related to the genomic organisation (A) of the ParaHox gene cluster. *AmphiGsx* (B–D) is expressed most anteriorly, *AmphiXlox* (E–K) is expressed more centrally and *AmphiCdx* (L–Q) is expressed most posteriorly. *AmphiCdx* is the first gene to be detectable during gastrulation around the closing blastopore (L) and then remains in a continuous domain in the posterior of the animal in the neural tube, hindgut and tail bud (M–Q). *AmphiXlox* expression commences slightly later than *AmphiCdx* in the posterior endoderm in two distinct domains, (dorsally and ventrally) (E–I) but later becomes restricted to a more central region of the gut in the developing larva (J, K). There is also transient expression in two neural tube cells in a region coinciding with the position of the future first pigment spot (G–I). The last ParaHox gene to be activated is *AmphiGsx*, which is initially expressed in the neural tube at the same level as the neural domain of *AmphiXlox* (B, C). Subsequently, *AmphiGsx* expression is down-regulated in this early domain and is activated later in the cerebral vesicle (D). The inset in D is a magnified dorsal view of the embryo at the level of the arrowhead. Embryos are presented as side views with anterior to the left except for dorsal views of the embryos in C and H. The anterior/posterior position of the boundary between somites 4 and 5 is indicated by an asterisk. Lowercase lettering denotes the developmental stage: g, gastrula; n, neurula, ee, early embryo; le, late embryo; el, early larvae; l, larvae, with similarly aged embryos/larvae aligned horizontally. Vertical dashed lines denote a continuation of the same expression pattern shown in the panels. Scale bars represent 100 μm.

examined the expression of both genes simultaneously. In single colour double in situ hybridisation experiments on control embryos (Figs. 4A, B) and on embryos treated with either RA or BMS009 (data not shown), only a single continuous domain of expression can be observed, indicating that there is no gap in expression between the two genes. From two colour double in situ hybridisations the

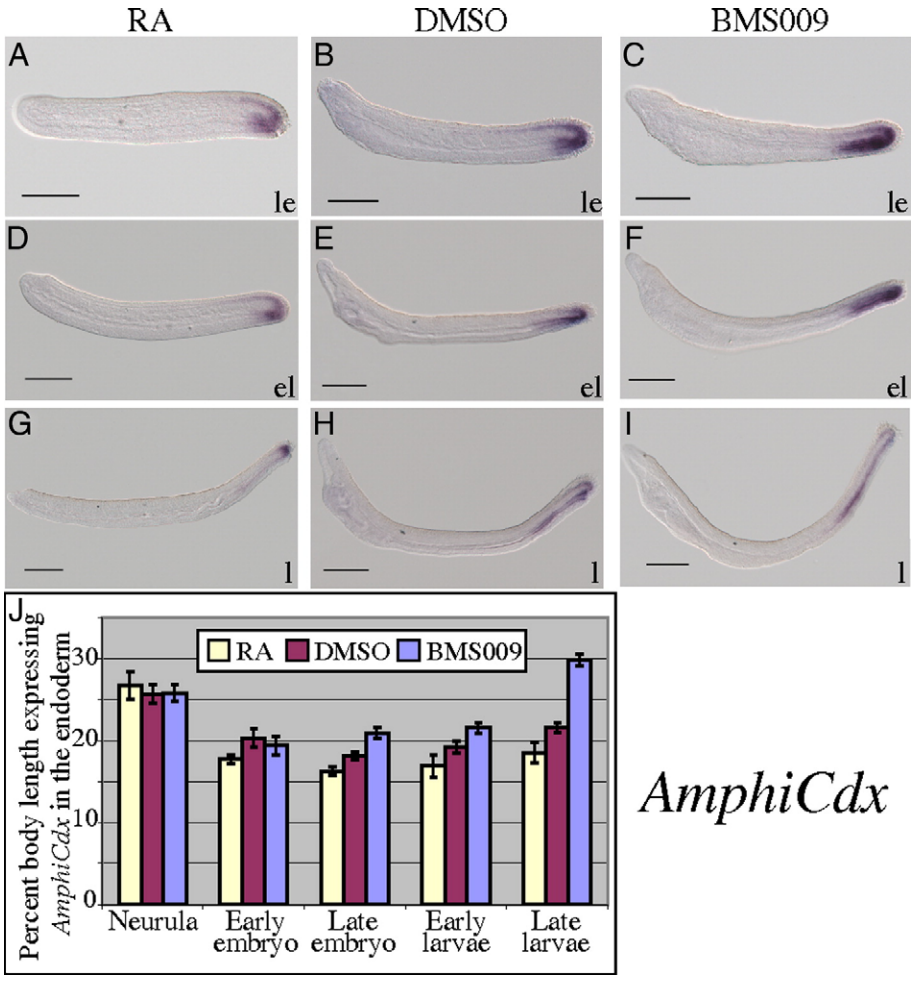

Fig. 2. Retinoic acid (RA) regulates the expression of *AmphiCdx* in the posterior endoderm. RA causes a posterior compression of the endodermal expression (A, D, G) relative to control embryos (B, E, H), whereas treatment with the RA antagonist BMS009 expands the *AmphiCdx* domain anteriorly (C, F, I). Embryos are presented as side views with anterior to the left. Lowercase lettering denotes the developmental stage: le, late embryo; el, early larvae; l, larvae. Scale bars represent 100 μm. Expression was not affected by treatment at early stages of development (J) with one-way ANOVA *p*-values of 0.86 and 0.24 for neurulae and early embryo, respectively. A significant difference exists between treatments at all other stages with *p*-values of less than 0.005. Error bars are standard error of the mean.

expression patterns clearly overlap in late neurulae in the ventral archenteron walls (data not shown). However as the pharynx begins to form, the overlap shrinks so by the late embryo stage little if any overlap in the expression domains in the archenteron can be seen (Figs. 4E, F). This is also obvious in RA- and BMS009-treated embryos (Figs. 4C, D, G, H) implying that subsequent to the initial activation of *AmphiCdx* and *AmphiXlox*, there is a secondary refinement of the limits of their expression domains, so that the two domains become adjacent to each other. It is the A/P position of this boundary during later stages of amphioxus development that is sensitive to RA.

RA alters AmphiGsx expression in the neuroectoderm

Unlike the endodermal modification of *AmphiCdx* and *AmphiXlox*, RA does not regulate expression of these two ParaHox genes in the neuroectoderm. However, RA does have a profound effect on *AmphiGsx* in the neuroectoderm. Expression of *AmphiGsx* is not modified in the cerebral vesicle in response to either RA or BMS009 but RA signalling does modify the A/P position of the early *AmphiGsx* domain (Fig. 5). Treatment with exogenous RA shifts and expands *AmphiGsx* expression anteriorly by increasing the number of *AmphiGsx*-expressing cells. An anterior shift was observed in all RA-treated embryos (Fig. 5A); however some embryos displayed a pattern of individual cells with high levels of expression interspersed by cells with low or no observable expression. In control embryos, the anterior and posterior limits of *AmphiGsx* expression were level with the anterior and posterior limits of somite 5 respectively (Fig. 5B). In RA-treated embryos the posterior limit of expression was shifted rostrally by approximately one somite length. The anterior limit of *AmphiGsx* was affected even more profoundly by RA, shifting rostrally by several somite lengths. An extreme effect was also observed in BMS009-treated embryos where the signal becomes completely undetectable by in situ hybridisation (even when stained for 30 days) (Fig. 5C). To quantify early *AmphiGsx* expression, the total number of *AmphiGsx*-expressing cells induced by RA and repressed by BMS009 were counted in treated embryos and compared to control embryos (Fig. 5G). The total number of cells expressing *AmphiGsx* are significantly different in response to treatment with RA or BMS009 (one-way ANOVA, $p<0.0001$).

Identification and analysis of putative ParaHox RAREs

As expression of all three ParaHox genes is clearly modified by RA treatment, we hypothesised a direct regulation and thus attempted to identify putative RAREs in the ParaHox cluster. Seven putative RAR/RXR binding sites were identified within the amphioxus ParaHox cluster. Six RAREs were localised within two distinct genomic segments, a region (Island 1) containing three putative RAREs between *AmphiGsx* and *AmphiXlox* (DR5a, DR5b and DR5c) and a region (Island 2) of three more putative RAREs between *AmphiXlox*

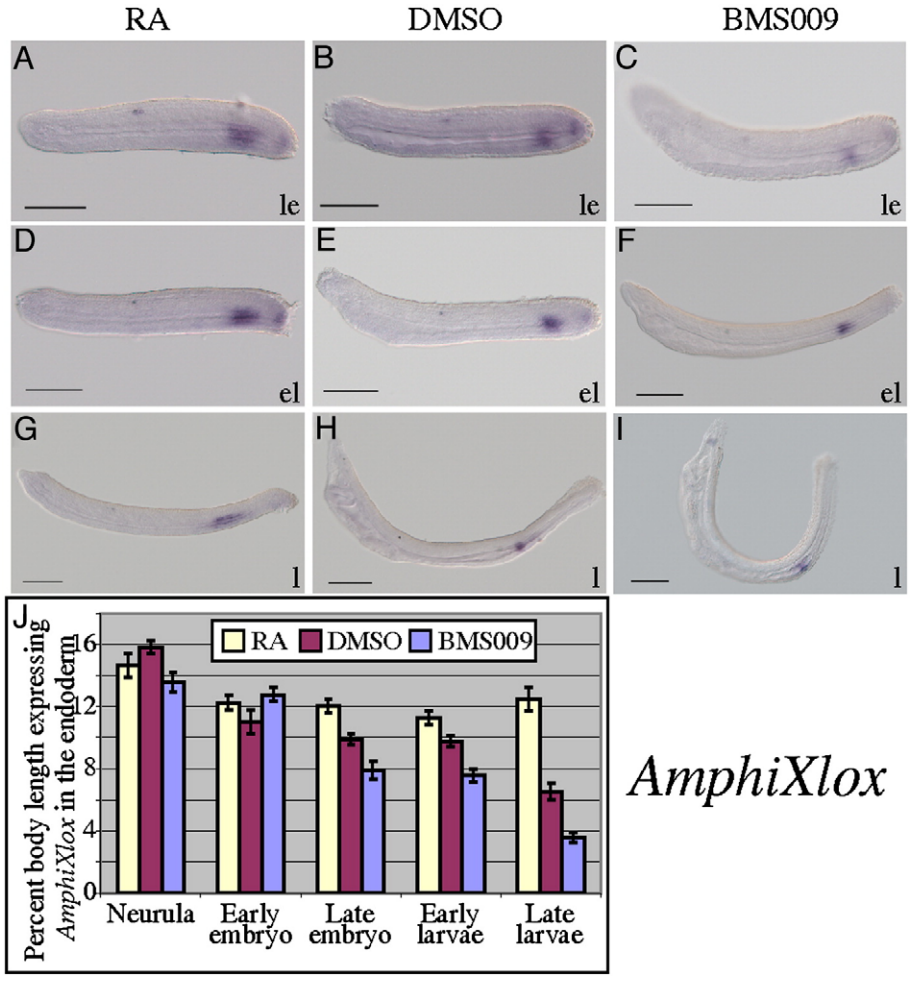

Fig. 3. Retinoic acid (RA) regulates the expression of *AmphiXlox* in the posterior endoderm. RA causes an expansion caudally of the posterior limit of endodermal *AmphiXlox* expression (A, D, G) relative to control embryos (B, E, H), whereas BMS009 treatment reduces the endodermal *AmphiXlox* domain posteriorly (C, F, I). Embryos are presented as side views with anterior to the left. Lowercase lettering denotes the developmental stage: le, late embryo; el, early larvae; l, larvae. Scale bars represent 100 μm. Expression was not affected by treatment at early stages of development (J) with one-way ANOVA *p*-values of 0.04 and 0.27 for mid-neurulae and late neurulae, respectively (the weak significance at the mid-neurula stage is due to the high variability amongst BMS009-treated embryos). A highly significant difference exists between treatments at all other stages with *p*-values of less than 0.0001. Error bars are standard error of the mean.

and *AmphiCdx* (DR5d, DR2a and DR2b). In addition, a single weak match to a RARE consensus sequence was identified upstream of *AmphiGsx* (DR5e) (Fig. 6A). As *AmphiGsx* expression was so dramatically altered by RA we tested this putative RARE as well, despite its lack of significance in the NHR scan. These seven putative ParaHox RAREs (Supplementary Table S1) were tested in vitro using a combination of EMSAs and heterologous cell culture transactivation experiments. In both assays, DR5c was the only single element that was bound weakly by the amphioxus RAR/RXR heterodimer (Fig. 6B, Supplementary Fig. S1). To further test DR5c, we analysed the ability of the amphioxus RAR/RXR heterodimer to activate transcription of a reporter construct containing single or tandem copies of DR5c. Moreover, we also tested the activity of DR5c in its native genomic context, ParaHox RARE Island 1 (Fig. 6B). The ParaHox Island 1 used in the analysis includes DR5b-DR5c (and several additional sequences with only a single nucleotide difference from the consensus DR5 which were not significant under NHR scan criteria). We also assayed the activity of ParaHox RARE Island 2 (Fig. 6A). A single copy of DR5c consistently causes more than 2-fold induction after stimulation with RA (Fig. 6B). However, multiple copies of the DR5c element did not significantly increase the level of induction of the reporter construct. The Island 1 construct (containing DR5c) consistently shows approximately 3-fold induction after treatment with RA. The RARE Island 2 construct is not significantly activated by RA, indicating it is not efficiently regulated by the amphioxus RAR/RXR heterodimer. Taken together, and considering that we are probably working at the limits of detection for such a heterologous assay system, this implies the amphioxus ParaHox DR5c element can be bound by amphioxus RAR/RXR and potentially mediate a direct RA-dependent response.

Discussion

Evolutionary constraints on the chordate ParaHox gene cluster

Intact ParaHox clusters have been identified in amphioxus, mice, humans (Brooke et al., 1998; Ferrier et al., 2005) and *Xenopus* (H.V. Isaacs personal communication) and it is generally assumed chordates have maintained this gene cluster as a result of selective constraints. It has been proposed these constraints are due to shared enhancer elements (Brooke et al., 1998). Alternatively it may be that the amphioxus and tetrapod lineages have simply retained the ParaHox cluster as an evolutionary relic. Indeed, recent demonstrations that both the hagfish and teleost lineages have lost their ParaHox clusters (Mulley et al., 2006; Furlong et al., 2007) along with cluster break-up in the urochordates (Ferrier and Holland, 2002) could argue against selective constraints maintaining chordate ParaHox clusters. A better understanding of the regulation of the ParaHox genes in intact clusters will allow us to understand whether the cluster is maintained by

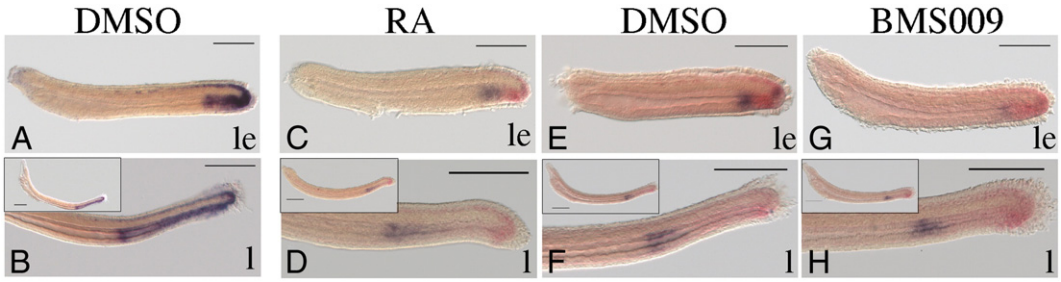

Fig. 4. Retinoic acid (RA) regulates the expression boundary between *AmphiXlox* and *AmphiCdx* in the posterior endoderm. Amphioxus larvae subjected to single colour double in situ hybridisations against both *AmphiCdx* and *AmphiXlox* reveal the expression domains are either overlapping or adjacent to each other, with no gap between the expression domains (A, B). Two colour double in situ hybridisations on early larvae show that the two genes have adjacent domains from the stages shown (E, F). Once the boundary between *AmphiXlox* (blue staining) and *AmphiCdx* (red staining) has formed in the late embryo, this boundary becomes responsive to treatment with RA or BMS009 (C, D, G, H). The images in B, D, F, H are views of the posterior half of the larvae displayed in the insets. Embryos are presented as side views with anterior to the left. Lowercase lettering denotes the developmental stage: le, late embryo; l, larvae. Scale bars represent 100 μm.

chance or through selective constraints, released in animals with unusual features (such as derived development, degenerate morphology or duplicated genomes).

Detailed descriptions of the expression of each ParaHox gene are an absolute prerequisite for studying their regulation. To this end we present a more extensive analysis of amphioxus ParaHox gene expression than previously described (Brooke et al., 1998). We show the importance of such careful examinations by identifying an extra *AmphiGsx* expression domain in the neural tube that overlaps with the *AmphiXlox* neural domain. This newly discovered *AmphiGsx* domain does not break the spatial or temporal collinearity of the amphioxus ParaHox gene expression (Brooke et al., 1998). The genes are still activated temporally (first to last) and spatially (posterior to anterior) in the order *AmphiCdx*, *AmphiXlox* and then *AmphiGsx*. The *AmphiXlox* and *AmphiCdx* genes also overlap during early embryonic development and later share a boundary in the endoderm. These overlapping/abutting domains raise the possibility that pairs of ParaHox genes share upstream transcriptional regulators and perhaps even share enhancers, which if present could provide a selective constraint for cluster retention.

RA regulates endodermal expression of AmphiXlox and AmphiCdx

The *AmphiXlox* domain in the gut endoderm is expanded posteriorly in response to exogenous RA, whereas the *AmphiCdx* domain in the gut is reduced anteriorly. Double in situ hybridisations suggest RA is actually shifting the A/P position of the boundary between the two genes and these results imply a tight mechanistic link between the regulation of *AmphiXlox* and *AmphiCdx*. Previous studies have demonstrated RA signalling in amphioxus is involved in development of the anterior (pharyngeal) and fore/midgut endoderm (Holland and Holland, 1996; Cañestro et al., 2001; Schubert et al., 2005). However, this is the first demonstration that posterior endodermal marker genes also respond to RA in amphioxus, indicating the entire A/P axis of the amphioxus endoderm is patterned through RA-dependent mechanisms.

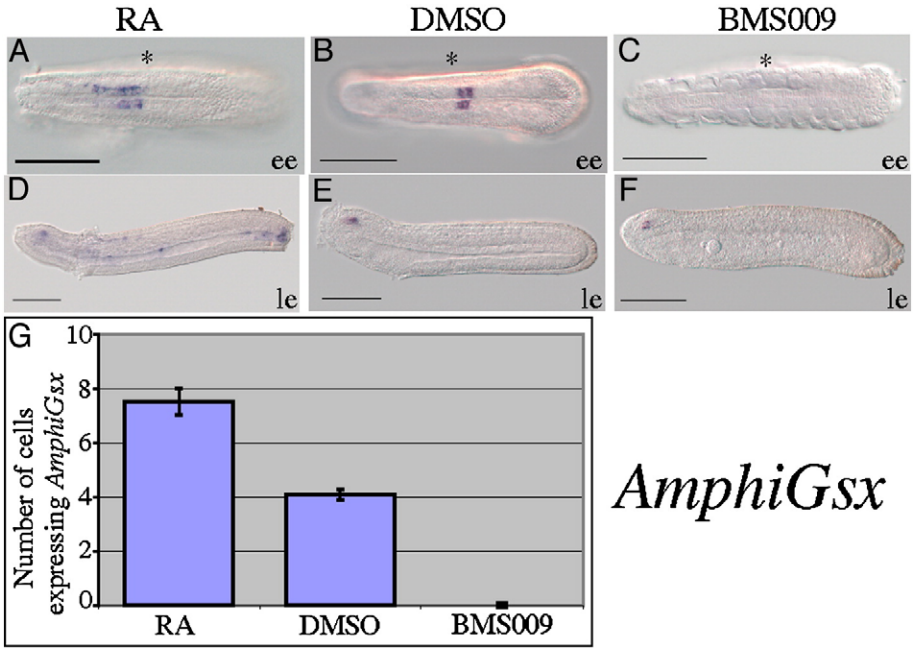

Fig. 5. Retinoic acid (RA) regulates the expression of *AmphiGsx* in the neural tube. The early *AmphiGsx* expression domain corresponds to four cells in the neural tube at the level of somite five (B). This domain is dramatically shifted and expanded anteriorly by treatment with RA (A) and reduced below the level of detection by in situ hybridisation by treatment with BMS009 (C). The embryo in C is at a different focal plane to those in A and B to display the somites and hence the stage of the embryo. The anterior/posterior position of the boundary between somites 4 and 5 is indicated by an asterisk. Treatment does not affect the later cerebral vesicle expression domain (D–F). Embryos are presented with anterior to the left, embryos in A–C are dorsal and in D–F are side views. Lowercase lettering denotes the developmental stage: ee, early embryo; le, late embryo. Scale bars represent 100 μm. Treatment affects the number of cells expressing *AmphiGsx* in the early domain causing a significant increase after RA treatment and a reduction after BMS009 treatment (one-way ANOVA *p*-value of less than 0.001) (G). Error bars are standard error of the mean.

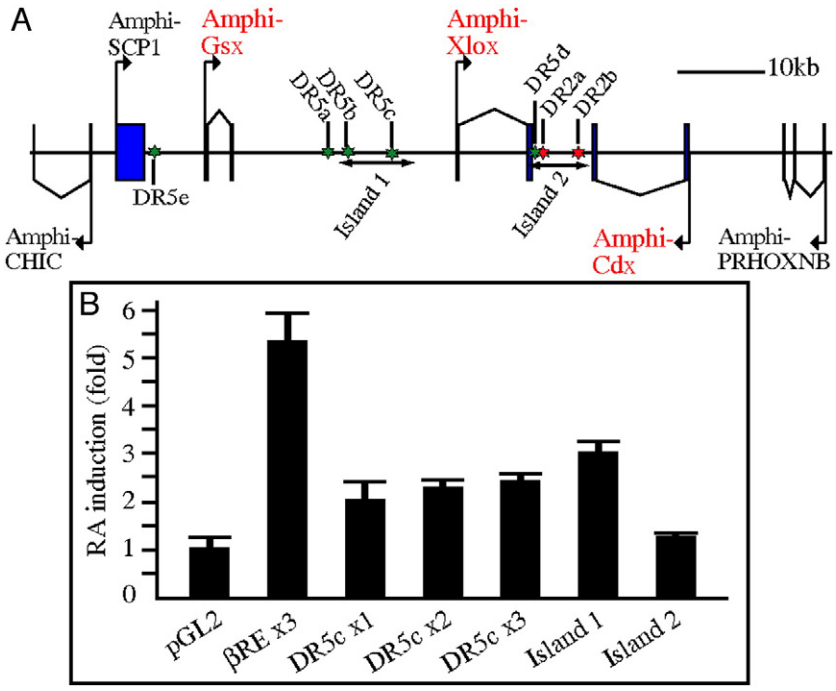

Fig. 6. Identification (A) and analysis (B) of putative retinoic acid response elements (RAREs), DR5a-e and DR2a-b, in the amphioxus ParaHox cluster. These putative RAREs are clustered in two islands, Island 1 upstream of *AmphiXlox* and Island 2 between *AmphiXlox* and *AmphiCdx*. In a heterologous cell culture transactivation assay, DR5c weakly activates transcription of a reporter gene after RA stimulation both as a single copy construct (DR5cx1) and when tested as a double (DR5cx2) or a triple (DR5cx3) copy construct. A construct containing the majority of Island 1 (including DR5b and DR5c) also weakly induces the reporter gene, while a construct of Island 2 (including DR5d, DR2a and DR2b) is less efficient in activating reporter gene expression. Empty pGL2 promoter vector was used as a negative control. A vector containing three copies of the RARE found in the human *RARβ2* regulatory region (βREx3) was used as a positive control for RA-stimulated, amphioxus RAR/RXR-dependent transcriptional activation in the cell culture transactivation assays. Error bars are standard error of the mean.

There are three major mechanisms that could explain the shift of the *AmphiCdx*/*AmphiXlox* boundary in the endoderm: (1) both ParaHox genes could be regulated directly and independently through one or more RAREs, (2) one of the ParaHox genes could be regulated directly through a RARE and the second ParaHox gene is regulated by the first, (3) neither ParaHox gene is directly regulated through RAREs and instead both are regulated by an upstream gene which is RA responsive. These mechanisms are not mutually exclusive and a combination could be acting on the *AmphiXlox*/*AmphiCdx* boundary.

Under scenario one, both *AmphiCdx* and *AmphiXlox* would be directly regulated by RA. This could be through a single RARE such as the identified DR5c upstream of *AmphiXlox*, or there may be additional atypical or distant RAREs regulating these genes not identified in this study.

Regulation of a boundary between two genes is possibly more likely to occur through cross-regulation between the genes (for example Toresson et al., 2000; Tour et al., 2002). This makes scenario two more likely and only one gene needs to be regulated by RA, with this gene subsequently regulating the A/P limit of the other gene. Although not definitive we postulate *AmphiXlox* is more likely to be directly regulated than *AmphiCdx* as *AmphiCdx* is repressed by ectopic RA, and RAR/RXR heterodimers normally mediate transcriptional activation rather than repression upon RA binding (Bastien and Rochette-Egly, 2004).

It is also possible that neither ParaHox gene is directly regulated by RA but instead an upstream, RA-responsive gene might regulate the *AmphiXlox*/*AmphiCdx* boundary. A good candidate for this would be *AmphiTR2/4* as it is RA responsive, competitively binds to RAREs and is expressed posteriorly in amphioxus embryos. In the *AmphiXlox*/*AmphiCdx* boundary region there are low–mid levels of both *AmphiRAR* and *AmphiTR2/4*. Under excess RA conditions, *AmphiRAR* is up-regulated throughout the posterior gut, while *AmphiTR2/4* decreases in the posterior endoderm (Escriva et al., 2002). In contrast, with BMS009 treatment, *AmphiRAR* levels are dramatically decreased in the posterior endoderm and *AmphiTR2/4* is up-regulated throughout the gut. Alteration of RA levels therefore clearly changes the balance between AmphiTR2/4 and AmphiRAR/AmphiRXR binding to RAREs in the posterior endoderm implicating AmphiTR2/4 as a candidate for regulating the *AmphiCdx*/*AmphiXlox* boundary.

In *Xenopus*, *Xlox*/*Xlhbox8* and *Cdx2* are expressed in abutting locations, respectively in the stomach/pancreas and small intestine (Horb and Slack, 2001). Zebrafish *cad1* (which is unaffected by exogenous RA) is expressed immediately posterior to the pancreas (where *Xlox*/*Pdx1* is expressed) (Stafford and Prince, 2002) and zebrafish *cdx4* initially does not overlap with *pdx1*, though by 16 h of development there are a few cells expressing both genes (Kinkel et al., 2008). In the mouse endoderm there appears to be a very small (if any) overlap of *Xlox*/*Pdx1* and *Cdx2* expression (Fang et al., 2006). The same is true for sea urchins, with expression of *Sp-Xlox* and *Sp-Cdx* meeting at the junction of the midgut sphincter (Arnone et al., 2006). One of the difficulties with examining boundary levels between Xlox and Cdx is the presence of three functionally redundant Cdx genes in vertebrates. This was partly addressed in zebrafish, where *Cdx4* null/MO-*Cdx1a* fish displayed an expansion of *Pdx1*, indicating a regulatory connection between these ParaHox genes (Kinkel et al., 2008). Altogether taking into account all the ParaHox genes of a single organism, an abutting boundary between Xlox and Cdx was probably the ancestral condition for at least all deuterostomes.

RA regulates early neural expression of AmphiGsx

Unlike the relatively late RA endodermal sensitivity of *AmphiXlox*/*AmphiCdx*, the early domain of *AmphiGsx* is RA responsive. The lack of change in late *AmphiGsx* expression in the CV is consistent with the hypothesis that gene expression in the amphioxus CV is protected against the influence of RA signalling, possibly through the activity of competitive RAR/RXR inhibitors, such as AmphiTR2/4 (Schubert et al., 2006b). Exogenous RA shifts the early *AmphiGsx* domain anteriorly

and induces additional *AmphiGsx*-expressing cells, whereas BMS009 abolishes *AmphiGsx* expression. This loss of detectable expression in response to BMS009 treatment also occurs in epidermal sensory cells for *AmphiERR* and *AmphiHox1, 3, 4* and *6* (Schubert et al., 2004). AmphiHox genes are also expressed in the developing neural tube, with *AmphiHox1-3* expression overlapping the earlier *AmphiGsx* expression (Wada et al., 1999; Schubert et al., 2006b). Indeed, *AmphiHox3* has an anterior boundary that approximately coincides with the anterior boundary of *AmphiGsx*. All of the Hox genes examined so far in the amphioxus neural tube are up-regulated and shifted anteriorly by RA, in a similar manner to *AmphiGsx*, in addition to being down-regulated and shifted posteriorly by BMS009 (Schubert et al., 2006b). Various studies have also suggested that *AmphiHox1* and *AmphiHox3* are likely to be directly regulated by RA (Manzanares et al., 2000; Schubert et al., 2005, 2006b; Wada et al., 2006) and perhaps the RA-induced anterior expansion of *AmphiGsx* expression is mediated by one of these Hox genes. The loss of the early *AmphiGsx* expression domain in BMS009-treated embryos could also be explained by the reduction and posterior shift of Hox expression in these embryos, as AmphiHox protein levels may have dropped below the threshold required to activate *AmphiGsx*. Intriguingly, a putative Hoxb1/Pbx binding site (AGATGGATGG) (Popperl et al., 1995) is located 1916 bp upstream of *AmphiGsx* (data not shown). In contrast to the *AmphiGsx* anterior limit, the posterior boundary shifts rostrally by only a single somite length, indicating the Hox gene pathway is unlikely to be specifying this posterior limit, as we would expect to see a more dramatic rostral shift after treatment with excess RA. Thus, there appear to be separate regulatory mechanisms for the anterior and posterior limits of *AmphiGsx* expression as these boundaries shift differentially in response to RA. We cannot however exclude the possibility that *AmphiGsx* is directly regulated by RAR/RXR and under this scenario the decrease of *AmphiRAR* expression levels in BMS009-treated embryos could explain the loss of *AmphiGsx* expression.

It is intriguing that only *AmphiGsx* is regulated by RA in the neural tube despite all three ParaHox genes being expressed in this tissue. This implies the regulatory apparatus of *AmphiXlox* and *AmphiCdx* is modular, with an endodermal regulatory module and a neural tube regulatory module. Under this mechanism only the endodermal modules of both genes will include a RA responsive pathway. The lack of neural *AmphiXlox* RA responsiveness is especially important as *AmphiXlox* and *AmphiGsx* expression overlap in wild-type embryos. After treatment with RA however, *AmphiGsx* shifts anteriorly whilst neural *AmphiXlox* expression remains unaffected, thereby reducing the likelihood of shared *AmphiGsx/AmphiXlox* enhancers constraining the chordate ParaHox cluster.

Direct regulation of ParaHox genes by RA

We postulate that at least some of the RA mediated effects on ParaHox expression are mediated directly by the RAR/RXR heterodimer. Consistent with this suggestion, we show by gel shift experiments that the amphioxus RAR/RXR heterodimer weakly binds to a putative RARE located between *AmphiGsx* and *AmphiXlox*. In addition, this binding site alone or an 8 kb genomic region surrounding this site both mediate transcriptional activation of a reporter gene upon RA stimulation. Although the induction was modest, these in vitro and cell transfection results imply the possibility of amphioxus RAR/RXR binding within the ParaHox cluster. The identification of this potential direct RA regulation on the ParaHox cluster is important for understanding both the regulation of the ParaHox genes and some of the constraints on the cluster. It should be noted that the amphioxus RAR/RXR heterodimer may not be directly activating ParaHox gene expression, but instead could be required for the remodelling of the chromatin, thus allowing other transcription factors access to the ParaHox cluster and hence allowing RA responsive regulation to occur. There may also be additional RAREs contributing to the regulation of the amphioxus ParaHox cluster that were not detected by our bioinformatic searches. However, if there are additional RAREs within the ParaHox cluster they must be highly divergent from the consensus RARE sequence.

One final alternative scenario that should not be ignored is the possibility of 9-*cis* RA mediating this effect. 9-*cis* RA is easily converted from all-*trans* RA and is likely to be present in small amounts during exogenous RA treatments with all-*trans* RA. Also, amphioxus contains relatively high levels of endogenous 9-*cis* RA, though still at half the level of all-*trans* RA (Dalfo et al., 2002). The physiological role of 9-*cis* RA in vertebrates is still unclear however, and also needs to be established in amphioxus (reviewed in Simões-Costa et al., 2008). It has recently been shown that amphioxus RXR can both bind and be activated by 9-*cis* RA, albeit with lower efficiency than vertebrate RXRs (Tocchini-Valentini et al., in press), and it is thus conceivable that the RA signal in chordates might be transduced by as yet undefined RAREs that would have been missed in the present work.

Evolution of ParaHox regulation by RA

Studies of RA regulation of ParaHox genes from vertebrates are limited. No studies have examined Gsx in response to RA and direct regulation of Xlox by RA has never been identified. RA does however induce *Pdx1* in mouse ES cells (Micallef et al., 2005; Shiraki et al., 2008), and RA from mesoderm explants is sufficient to induce Pdx1 in anterior endoderm (Kumar et al., 2003). In addition, reduction of RA in both mice and zebrafish results in a loss of *Pdx1* expression in the dorsal pancreatic anlage, and replacement of RA restores mouse *Pdx1* expression (Stafford and Prince, 2002; Martin et al., 2005; Molotkov et al., 2005). In contrast to Xlox and Gsx, RA has been demonstrated to directly regulate both mouse and chicken *Cdx1/CdxA* genes through an atypical RARE upstream of mouse *Cdx1* and also through a conserved RARE within the intron of tetrapod *Cdx1/CdxA* genes (Houle et al., 2000, 2003; Gaunt et al., 2003). RA also causes a reduction of anterior *Cdx4* expression in mouse neural tissue and mesoderm (endodermal *Cdx4* expression was not examined) (Iulianella et al., 1999) and mouse *Cdx2* may also be slightly reduced (Roelen et al., 2002). Clear similarities exist for RA regulation between *AmphiCdx* and tetrapod *Cdx2* and *Cdx4* genes, with all of these genes being down-regulated in response to RA. However, changes in the regulation of Cdx by RA have obviously occurred, as *AmphiCdx* is not induced like the mouse *Cdx1* genes; and the zebrafish *Cdx4* gene appears to confer immunity against RA signalling in posterior endoderm (Kinkel et al., 2008). Moreover, the RAREs controlling the RA response of mouse and chicken *Cdx1/CdxA* genes are not conserved in the amphioxus ParaHox cluster (data not shown).

Here we have demonstrated all three amphioxus ParaHox genes are regulated by RA. These results are the first demonstration of all three ParaHox genes being affected by RA in a single species. In fact, no signalling pathway has previously been shown to affect the complete set of ParaHox genes from any animal. We also provide evidence that some of this regulation may be directly mediated through a RARE in the ParaHox cluster. Although there are differences between the regulation of vertebrate and amphioxus ParaHox genes by RA, it is likely that RA is upstream of all three ParaHox genes in both taxa (although vertebrate Gsx genes clearly need to be examined in more detail). Quite probably then, at the origin of the chordates the ParaHox cluster was already regulated by RA. It is clear however, that a substantial amount of work on chordate ParaHox regulation is still required, with consideration of the ParaHox cluster as a whole being paramount. It is also intriguing that the evolutionary sister of the ParaHox cluster (the Hox cluster) is also regulated by RA in both amphioxus and vertebrates (Wada et al., 2006). Recent discoveries of RAR genes from protostomes (Campo-Paysaa et al., 2008) have also pushed back the likely origin of RA signalling in animal evolution, allowing the intriguing hypothesis that perhaps RA regulated the original ProtoHox cluster. Further work on

the control of ParaHox and Hox genes by RA within the animal kingdom, should establish whether regulation by RA signalling is a conserved trait of these sister gene clusters.

Note added in proof

The organisation of the *Xenopus tropicalis* ParaHox cluster mentioned as a personal communication from H.V. Isaacs (lines 406-407) has recently been published, Iles, J.C., Winterbottom, E. and Isaacs, H.V. 2008. Cloning and expression analysis of the anterior ParaHox genes, *Gsh1* and *Gsh2* from *Xenopus tropicalis*. Devel. Dyn. 238, 194–203.

Acknowledgments

The authors are indebted to John M. Lawrence from the University of South Florida, USA, for providing laboratory space during the amphioxus spawning season and to Linda Z. Holland for help during amphioxus field work. The authors would also like to acknowledge Gérard Triqueneaux and Mathilde Paris for technical assistance with gel shift experiments, Ben Raymond for advice on statistical methods, Jordi Garcia-Fernàndez for providing the *AmphiCdx* clone and Hinrich Gronemeyer for providing the BMS009 compound. This work was supported by grants from the BBSRC to DEKF, the ANR and the CNRS to MS, the MENRT to MS and VL as well as by CRESCENDO, a European Union Integrated Project of FP6. PWO is funded by the Clarendon, ORS and EPA Cephalosporin scholarships. The authors would like to thank the anonymous reviewers for their comments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2008.11.027.

References

Arnone, M.I., Rizzo, F., Annunciata, R., Cameron, R.A., Peterson, K.J., Martinez, P., 2006. Genetic organization and embryonic expression of the ParaHox genes in the sea urchin *S. purpuratus*: insights into the relationship between clustering and colinearity. Dev. Biol. 300, 63–73.

Bastien, J., Rochette-Egly, C., 2004. Nuclear retinoid receptors and the transcription of retinoid-target genes. Gene 328, 1–16.

Brooke, N.M., Garcia-Fernàndez, J., Holland, P.W.H., 1998. The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. Nature 392, 920–922.

Campo-Paysaa, F., Marlétaz, F., Laudet, V., Schubert, M., 2008. Retinoic acid signaling in development: tissue-specific functions and evolutionary origins. Genesis 46, 640–656.

Cañestro, C., Albalat, R., Escriva, H., Gonzalez-Duarte, R., 2001. Endogenous β-galactosidase activity in amphioxus: a useful histochemical marker for the digestive system. Dev. Genes Evol. 211, 154–156.

Coulier, F., Burtey, S., Chaffanet, M., Birg, F., Birnbaum, D., 2000. Ancestrally-duplicated paraHOX gene clusters in humans. Int. J. Oncol. 17, 439–444.

Dalfo, D., Albalat, R., Molotkov, A., Duester, G., Gonzalez-Duarte, R., 2002. Retinoic acid synthesis in the prevertebrate amphioxus involves retinol oxidation. Dev. Genes Evol. 212, 388–393.

Duboule, D., 2007. The rise and fall of Hox gene clusters. Development 134, 2549–2560.

Escriva, H., Holland, N.D., Gronemeyer, H., Laudet, V., Holland, L.Z., 2002. The retinoic acid signaling pathway regulates anterior/posterior patterning in the nerve cord and pharynx of amphioxus, a chordate lacking neural crest. Development 129, 2905–2916.

Fang, R.X., Olds, L.C., Sibley, E., 2006. Spatio-temporal patterns of intestine-specific transcription factor expression during postnatal mouse gut development. Gene Expression Patterns 6, 426–432.

Ferrier, D.E.K., Holland, P.W.H., 2001a. Sipunculan ParaHox genes. Evolut. Develop. 3, 263–270.

Ferrier, D.E.K., Holland, P.W.H., 2001b. Ancient origin of the Hox gene cluster. Nat. Rev., Genet. 2, 33–38.

Ferrier, D.E.K., Holland, P.W.H., 2002. *Ciona intestinalis* ParaHox genes: evolution of Hox/ParaHox cluster integrity, developmental mode, and temporal colinearity. Mol. Phylogenet. Evol. 24, 412–417.

Ferrier, D.E.K., Minguillón, C., 2003. Evolution of the Hox/ParaHox gene clusters. Int. J. Dev. Biol. 47, 605–611.

Ferrier, D.E.K., Dewar, K., Cook, A., Chang, J.L., Hill-Force, A., Amemiya, C., 2005. The chordate ParaHox, cluster. Curr. Biol. 15, R820–R822.

Finnerty, J.R., Martindale, M.Q., 1999. Ancient origins of axial patterning genes: Hox genes and ParaHox genes in the Cnidaria. Evolut. Develop. 1, 16–23.

Furlong, R.F., Younger, R., Kasahara, M., Reinhardt, R., Thorndyke, M., Holland, P.W.H., 2007. A degenerate ParaHox gene cluster in a degenerate vertebrate. Mol. Biol. Evol. 24, 2681–2686.

Gaunt, S.J., Drage, D., Cockley, A., 2003. Vertebrate caudal gene expression gradients investigated by use of chick *cdx-A/lacZ* and mouse *cdx-1/lacZ* reporters in transgenic mouse embryos: evidence for an intron enhancer. Mech. Dev. 120, 573–586.

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

Holland, L.Z., Holland, P.W.H. and Holland, N.D., (1996). Revealing homologies between body parts of distantly related animals by *in situ* hybridization to developmental genes: amphioxus versus vertebrates. In Molecular Zoology. Advances, Strategies, and Protocols (eds Ferraris, J. D. and Palumbi, S. R.), pp. pp. 267–282; 473–483. New York, NY: Wiley-Liss.

Horb, M.E., Slack, J.M., 2001. Endoderm specification and differentiation in *Xenopus* embryos. Dev. Biol. 236, 330–343.

Houle, M., Prinos, P., Iulianella, A., Bouchard, N., Lohnes, D., 2000. Retinoic acid regulation of Cdx1: an indirect mechanism for retinoids and vertebral specification. Mol. Cell. Biol. 20, 6579–6586.

Houle, M., Sylvestre, J.R., Lohnes, D., 2003. Retinoic acid regulates a subset of Cdx1 function in vivo. Development 130, 6555–6567.

Iulianella, A., Beckett, B., Petkovich, M., Lohnes, D., 1999. A molecular basis for retinoic acid-induced axial truncation. Dev. Biol. 205, 33–48.

Kinkel, M.D., Eames, S.C., Alonzo, M.R., Prince, V.E., 2008. Cdx4 is required in the endoderm to localize the pancreas and limit β-cell number. Development 135, 919–929.

Kumar, M., Jordan, N., Melton, D., Grapin-Botton, A., 2003. Signals from lateral plate mesoderm instruct endoderm toward a pancreatic fate. Dev. Biol. 259, 109–122.

Langeland, J.A., Tomsa, J.M., Jackman, W.R., 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.

Lickert, H., Kemler, R., 2002. Functional analysis of *cis*-regulatory elements controlling initiation and maintenance of early *Cdx1* gene expression in the mouse. Dev. Dyn. 225, 216–220.

Maden, M., 2007. Retinoic acid in the development, regeneration and maintenance of the nervous system. Nat. Rev., Neurosci. 8, 755–765.

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

Mark, M., Ghyselinck, N.B., Chambon, P., 2006. Function of retinoid nuclear receptors: lessons from genetic and pharmacological dissections of the retinoic acid signaling pathway during mouse embryogenesis. Annu. Rev. Pharmacol. Toxicol. 46, 451–480.

Marlétaz, F., Holland, L.Z., Laudet, V., Schubert, M., 2006. Retinoic acid signaling and the evolution of chordates. Int. J. Biol. Sci. 2, 38–47.

Martin, M., Gallego-Llamas, J., Ribes, V., Kedinger, M., Niederreither, K., Chambon, P., Dolle, P., Gradwohl, G., 2005. Dorsal pancreas agenesis in retinoic acid-deficient *Raldh2* mutant mice. Dev. Biol. 284, 399–411.

Micallef, S.J., Janes, M.E., Knezevic, K., Davis, R.P., Elefanty, A.G., Stanley, E.G., 2005. Retinoic acid induces Pdx1-positive endoderm in differentiating mouse embryonic stem cells. Diabetes 54, 301–305.

Molotkov, A., Molotkova, N., Duester, G., 2005. Retinoic acid generated by *Raldh2* in mesoderm is required for mouse dorsal endodermal pancreas development. Dev. Dyn. 232, 950–957.

Mulley, J.F., Chiu, C.H., Holland, P.W.H., 2006. Breakup of a homeobox cluster after genome duplication in teleosts. Proc. Natl. Acad. Sci. U. S. A. 103, 10369–10372.

Perlmann, T., Rangarajan, P.N., Umesono, K., Evans, R.M., 1993. Determinants for selective RAR and TR recognition of direct repeat HREs. Genes Dev. 7, 1411–1422.

Pilon, N., Oh, K., Sylvestre, J.R., Savory, J.G., Lohnes, D., 2007. Wnt signaling is a key mediator of *Cdx1* expression in vivo. Development 134, 2315–2323.

Popperl, H., Bienz, M., Studer, M., Chan, S.K., Aparicio, S., Brenner, S., Mann, R.S., Krumlauf, R., 1995. Segmental expression of *Hoxb-1* is controlled by a highly conserved autoregulatory loop dependent upon *exd/pbx*. Cell 81, 1031–1042.

Prohaska, S.J., Stadler, P.F., 2006. Evolution of the vertebrate ParaHox clusters. J. Exp. Zool. Part B — Mol. Dev. Evol. 306B, 481–487.

Rebeiz, M., Posakony, J.W., 2004. GenePalette: a universal software tool for genome sequence visualization and analysis. Dev. Biol. 271, 431–438.

Roelen, B.A., de Graaff, W., Forlani, S., Deschamps, J., 2002. *Hox* cluster polarity in early transcriptional availability: a high order regulatory level of clustered *Hox* genes in the mouse. Mech. Dev. 119, 81–90.

Sandelin, A., Wasserman, W.W., 2005. Prediction of nuclear hormone response elements. Mol. Endocrinol. 19, 595–606.

Schubert, M., Holland, N.D., Escriva, H., Holland, L.Z., Laudet, V., 2004. Retinoic acid influences anteroposterior positioning of epidermal sensory neurons and their gene expression in a developing chordate (amphioxus). Proc. Natl. Acad. Sci. U. S. A. 101, 10320–10325.

Schubert, M., Yu, J.K., Holland, N.D., Escriva, H., Laudet, V., Holland, L.Z., 2005. Retinoic acid signaling acts via *Hox1* to establish the posterior limit of the pharynx in the chordate amphioxus. Development 132, 61–73.

Schubert, M., Escriva, H., Xavier-Neto, J., Laudet, V., 2006a. Amphioxus and tunicates as evolutionary model systems. Trends Ecol. Evol. 21, 269–277.

Schubert, M., Holland, N.D., Laudet, V., Holland, L.Z., 2006b. A retinoic acid-*Hox* hierarchy controls both anterior/posterior patterning and neuronal specification in the developing central nervous system of the cephalochordate amphioxus. Dev. Biol. 296, 190–202.

Shiraki, N., Yoshida, T., Araki, K., Umezawa, A., Higuchi, Y., Goto, H., Kume, K., Kume, S., 2008. Guided differentiation of ES cells into Pdx1-expressing regional specific definitive endoderm. Stem Cells 26, 874–885.

Siegel, N., Hoegg, S., Salzburger, W., Braasch, I., Meyer, A., 2007. Comparative genomics of ParaHox clusters of teleost fishes: gene cluster breakup and the retention of gene sets following whole genome duplications. BMC Genomics 8, 312.

Simões-Costa, M.S., Azambuja, A.P., Xavier-Neto, J., 2008. The search for non-chordate retinoic acid signaling: lessons from chordates. J. Exp. Zool. B — Mol. Dev. Evol. 310, 54–72.

Stafford, D., Prince, V.E., 2002. Retinoic acid signaling is required for a critical early step in zebrafish pancreatic development. Curr. Biol. 12, 1215–1220.

Tocchini-Valentini, G.D., Rochel, N., Escriva, H., Germain, P., Peluso-Iltis, C., Paris, M., Sanglier-Cianferani, S., Van Dorsselaer, A., Moras, D. and Laudet, V., in press. Structural and functional insights into the ligand binding domain of a non-duplicated RXR from the invertebrate chordate amphioxus. J. Biol. Chem. doi:10.1074/jbc.M805692200.

Toresson, H., Potter, S.S., Campbell, K., 2000. Genetic control of dorsal–ventral identity in the telencephalon: opposing roles for Pax6 and Gsh2. Development 127, 4361–4371.

Tour, E., Pillemer, G., Gruenbaum, Y., Fainsod, A., 2002. *Gbx2* interacts with *Otx2* and patterns the anterior–posterior axis during gastrulation in *Xenopus*. Mech. Dev. 112, 141–151.

Tulachan, S.S., Doi, R., Kawaguchi, Y., Tsuji, S., Nakajima, S., Masui, T., Koizumi, M., Toyoda, E., Mori, T., Ito, D., et al., 2003. All-*Trans* retinoic acid induces differentiation of ducts and endocrine cells by mesenchymal/epithelial interactions in embryonic pancreas. Diabetes 52, 76–84.

Vanacker, J.M., Pettersson, K., Gustafsson, J.A., Laudet, V., 1999. Transcriptional targets shared by estrogen receptor-related receptors (ERRs) and estrogen receptor (ER) α, but not by ER. EMBO J. 18, 4270–4279.

Wada, H., Garcia-Fernàndez, J., Holland, P.W.H., 1999. Colinear and segmental expression of amphioxus Hox genes. Dev. Biol. 213, 131–141.

Wada, H., Escriva, H., Zhang, S.C., Laudet, V., 2006. Conserved RARE localization in amphioxus *Hox* clusters and implications for *Hox* code evolution in the vertebrate neural crest. Dev. Dyn. 235, 1522–1531.

Wang, Z., Benoit, G., Liu, J., Prasad, S., Aarnisalo, P., Liu, X., Xu, H., Walker, N.P., Perlmann, T., 2003. Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors. Nature 423, 555–560.