Developmental Biology 321 (2008) 455–469

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/developmentalbiology

Genomes & Developmental Control

cis-Regulatory sequences driving the expression of the *Hbox12* homeobox-containing gene in the presumptive aboral ectoderm territory of the *Paracentrotus lividus* sea urchin embryo

Vincenzo Cavalieri [a], Maria Di Bernardo [b], Letizia Anello [b], Giovanni Spinelli [a,*]

[a] *Dipartimento di Biologia Cellulare e dello Sviluppo A. Monroy, Università di Palermo, Viale delle Scienze Edificio 16, 90128 Palermo, Italy*
[b] *Istituto di Biomedicina e Immunologia Molecolare A. Monroy-CNR, via Ugo La Malfa 153, 90146 Palermo, Italy*

ARTICLE INFO

Article history:
Received for publication 25 October 2007
Revised 23 May 2008
Accepted 3 June 2008
Available online 13 June 2008

Keywords:
Homeodomain
Aboral ectoderm
cis-Regulatory module
Gene transfer
Otx
Sox
Myb
Sea urchin embryo
Chromatin ImmunoPrecipitation

ABSTRACT

Embryonic development is coordinated by networks of evolutionary conserved regulatory genes encoding transcription factors and components of cell signalling pathways. In the sea urchin embryo, a number of genes encoding transcription factors display territorial restricted expression. Among these, the zygotic *Hbox12* homeobox gene is transiently transcribed in a limited number of cells of the animal-lateral half of the early *Paracentrotus lividus* embryo, whose descendants will constitute part of the ectoderm territory. To obtain insights on the regulation of *Hbox12* expression, we have explored the *cis*-regulatory apparatus of the gene. In this paper, we show that the intergenic region of the tandem *Hbox12* repeats drives GFP expression in the presumptive aboral ectoderm and that a 234 bp fragment, defined aboral ectoderm (AE) module, accounts for the restricted expression of the *transgene*. Within this module, a consensus sequence for a Sox factor and the binding of the Otx activator are both required for correct *Hbox12* gene expression. Spatial restriction to the aboral ectoderm is achieved by a combination of different repressive sequence elements. Negative sequence elements necessary for repression in the endomesoderm map within the most upstream 60 bp region and nearby the Sox binding site. Strikingly, a Myb-*like* consensus is necessary for repression in the oral ectoderm, while down-regulation at the gastrula stage depends on a *GA*-rich region. These results suggest a role for *Hbox12* in aboral ectoderm specification and represent our first attempt in the identification of the gene regulatory circuits involved in this process.

Introduction

Development of animal embryos proceeds as a progression of states of spatially defined regulatory gene expression (Angerer and Angerer, 2003; Davidson et al., 1998; Goldstein and Freeman, 1997). Through this progression, specification occurs when groups of cells of the embryo come to express a given set of genes. Correct patterning of cell fates along the embryonic axes requires differential inheritance of maternal regulatory molecules and signalling interactions among cells. A large part of the genetic program and the fundamental laws directing cell specification are encoded in the genomic DNA. However, it is still not quite clear how genes are activated or silenced in an orchestrated manner in any cell type or embryonic territory and how gene activities are maintained through cell generations to ensure the proper development.

It has been proposed that the information within the genomic regulatory sequences might constitute a code for development that is interpreted by the specific binding of transcription factors to the proper target sites. The genomic regulatory code for development generates a system of evolutionary conserved interactions that has the architecture of a network (Davidson et al., 2002a,b). Stated in simple terms, the outputs of these interactions result in a gene being turned on or off at the appropriate developmental time, cell lineages or in different areas of the developing embryo. Transcription factors play a key role at the nodes of such a gene network: if a gene encodes a transcriptional regulator, the output influences other *cis*-regulatory elements that are target sites for that regulator. The linkages of a gene regulatory network (GRN) can be tested by performing experiments that perturb gene expression and can be verified by identifying the *cis*-regulatory control elements and their key target sites. Such an analysis has been extensively performed with the purple sea urchin embryo *Strongylocentrotus purpuratus* and has led to the formulation of an explicit model of GRN directing the specification of the endomesodermal cell types during the initial 30 h of development and including almost 50 genes (Oliveri and Davidson, 2004a,b). The explanatory power as well as the predictive properties of a GRN can ultimately be revealed through the systematic identification of the key fragments of genomic DNA that execute the *cis*-regulatory interactions. The recently available genome sequence of *S. purpuratus* (Sodergren et al., 2006) provides a crucial data set for this purpose,

* Corresponding author. Fax: +39 091 6577 410.
E-mail address: spinelli@unipa.it (G. Spinelli).

0012-1606/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2008.06.006

allowing a unique opportunity to investigate the function and the evolutionary process also through a comparative point of view.

Notwithstanding the extraordinary level of molecular details available on sea urchin embryo patterning along the animal–vegetal axis (Ettensohn et al., 2003; Logan et al., 1999; Oliveri et al., 2002; Weitzel et al., 2004), less is known about the specification and regionalization along the oral–aboral (OA) axis of polarity which runs from the mouth to the opposite side of the larva. OA polarity is not firmly established in the unfertilized egg, but rather relies on a combination of inherited maternal information and inductive interactions among early blastomeres, becoming morphologically recognizable from the gastrula stage onward (Angerer and Angerer, 2003; Brandhorst and Klein, 2002; Davidson et al., 1998; Wikramanayake et al., 1998). It has been recently shown that mitochondria are asymmetrically distributed in unfertilized eggs of *S. purpuratus* (Coffman and Denegre, 2007; Coffman et al., 2004). The polarity of this anisotropic distribution does not change significantly in the early embryo, leading to an unequal apportioning of mitochondria to the blastomeres. The maternal mitochondrial asymmetry correlates with OA polarity, with the blastomeres inheriting the highest density of mitochondria tending to give rise to the oral pole of the embryo (Coffman et al., 2004). Moreover, microinjection of purified mitochondria can bias the orientation of the OA axis (Coffman et al., 2004). Thus, specification of OA polarity appears to be entrained, at least in part, by a maternally specified anisotropy in mitochondrial distribution. At gastrula stage the embryonic ectoderm territory is noticeably partitioned into a thickened (oral) and a squamous (aboral) epithelium, separated by a belt of cuboidal ciliated cells, namely the ciliary band. The TGF-β growth factors have emerged as a major family of paracrine signalling molecules regulating oral–aboral polarization in early development (Angerer and Angerer, 2000; Duboc et al., 2004; Range et al., 2007). Among the TGF-β gene super-family, *Nodal* is expressed on the oral side of the early embryo, where it plays fundamental roles by establishing the oral–aboral and left–right asymmetries (Duboc and Lepage, 2008; Duboc et al., 2004, 2005). Hypoxia appears to radialize embryos by suppressing *Nodal* expression (Coffman et al., 2004), suggesting that *Nodal* expression is redox-regulated. Accordingly, the *cis*-regulatory apparatus of Nodal contains consensus target sequences for the redox-sensitive bZIP transcription factors (Nam et al., 2007; Range et al., 2007).

Though recent evidence confirmed the *Nodal cis*-regulatory system as a necessary key for the progression of the oral ectoderm network, unveiling the architecture of such a network requires primary knowledge of the whole set of transcription factors that are active in the presumptive ectoderm and their functional interactions. An increasing number of genes encoding transcription regulators are known to be specifically activated in cells that become ectoderm (Amore et al., 2003; Howard-Ashby et al., 2006a,b; Tu et al., 2006), and among these is the *Hbox12* homeobox-containing gene (Di Bernardo et al., 1995). *PlHbox12* transcription occurs transiently during the very early cleavage stages, immediately preceding the specification process leading to the segregation of broad regions of embryonic territory precursors. Following its initial activation at the 4/8-cell stage embryo, *Hbox12* expression increases up to morula/early blastula stage, fading out and becoming silenced after hatching (Di Bernardo et al., 1995). Whole mount *in situ* hybridization revealed that *Hbox12* transcripts are asymmetrically distributed along both the animal–vegetal and the oral–aboral axes, in some blastomeres of the animal hemisphere of fourth to sixth cleavage embryos (Di Bernardo et al., 1995). Undoubtedly, descendants of the *Hbox12*-expressing cells constitute part of the ectoderm territory, but their allocation to either the oral or aboral lineage remained too long uncharacterized. The observed pattern of expression is unique in sea urchin development, and indicates that the *cis*-regulatory sequences that control the expression of this gene receive precocious input from the newborn gene regulatory network controlling sea urchin development.

In order to obtain insights on the regulation of *Hbox12* expression in the sea urchin *Paracentrotus lividus*, we have explored its *cis*-regulatory apparatus. Hence, we identified a compact module, containing both positive and negative *cis*-regulatory sequences, responsible for the restricted expression of the GFP *transgene*. Interestingly, within this module we demonstrated that a consensus sequence for a Sox factor and the binding of the Otx transcriptional activator are both required for the correct *Hbox12* gene expression. Spatial restriction to the aboral ectoderm is achieved by a combination of different repressive sequence elements. Among them, we identified a Myb-*like* consensus, which prevents ectopic expression in the oral ectoderm cells, and a 60 bp DNA fragment necessary for repression in the endomesoderm. Furthermore, a multiple *GA* repeat-containing sequence is involved in the down-regulation of the gene at gastrula stage.

Materials and methods

Isolation of sea urchin Hbox12 cDNAs and sequence alignment

Three cDNAs, named *Hbox12-c* to *-e*, corresponding to the complete coding sequences of *Hbox12* were obtained by screening a cDNA library from morula *P. lividus* embryos using ^{32}P-labeled probes generated from full-length *Hbox12* cDNA (Di Bernardo et al., 1995). Full nucleotide sequences were verified and translated using the pDRAW32 software (http://www.acaclone.com). Multiple sequence alignment of homeodomains was generated using ClustalW version 1.83 (Chenna et al., 2003; Thompson et al., 1994) and the alignment output file was formatted using BioEdit version 7.0.8. Accession numbers for sequences used in alignment were listed below. Sequences of *Sppmar* 1.1 to 1.4 were obtained from Genscan analysis on the nucleotide sequence of two *S. purpuratus* BAC clones (Genbank accessions: NW_001304149 and NW_001292071.1).

Preparation of promoter-GFP transgene constructs

A *Hbox12* lambda-clone of about 12 kb was retrieved from a *P. lividus* genomic library and an insert of almost 3 kb, corresponding to the complete *Hbox12-a* gene, was subcloned into the pGEM-4Z vector (Clontech). The 1.45GFP construct was generated by PCR fusion as follows. A 1.45 kb fragment abutting at the 3′-end the ATG start codon of *Hbox12* was PCR-amplified with *Pfu* DNA polymerase (Promega). Primers used for the amplification were pHbox12-EcoRV-up, 5′-GGGGGGATATCGGATGAGAAAATGAGTGT-3′; and pHbox12-BglII-down, 5′-GGGGCAGATCTCGATGAGACTTGGGTGAT-3′.The underlined sequences are the restriction enzyme sites created for ligation. The PCR fragment was cut with EcoRV and BglII and ligated upstream and *in frame* to the Green Lantern *GFP* reporter gene into a pGL3-modified vector. Derivative constructs were obtained by using standard molecular biology techniques. Prior to use for microinjection experiments, the sequence of each construct was confirmed. Hence, the resulting plasmids were linearized at the *KpnI* site of the plasmid polylinker, upstream to the 5′-flanking region of *Hbox12*. The functional 0.84 kb promoter-GFP DNA fragment was obtained from the 1.45GFP construct by PCR reaction. Nucleotide sequences of primers were as follows: pHbox12-up (5′-TAGTCAGAAAGAGAAAAAGAGATG-3′) and pGL-down (5′-CCTCTTCGCTATTACGCCAG-3′). Amplification reaction was performed using the “Elongase Amplification System” (Invitrogen).

Binding sites search and mutagenesis

Putative binding site sequences were searched using TESS (url: http://www.cbil.upenn.edu/tess) and MatInspector (Cartharius et al., 2005; Quandt et al., 1995) software packages. A couple of oligonucleotide primers containing five mutated bases were generated for the conserved predicted site for Otx. Two additional primers were generated to eliminate a 29 bp sequence containing the consensus

site for Sox. In both cases, mutations were introduced by PCR using Pfu polymerase (Promega). Ten nanograms of double-strand template were used per reaction. Following temperature cycling (95 °C 30 s, then 18 cycles of 95 °C 30 s, then 55 °C 1 min, 68 °C 6 min), the products were treated with DpnI (Invitrogen), to remove the template DNA molecules. The nicked DNA constructs were then transformed into XL1-Blue competent cells and mutations confirmed by restriction digestions and sequencing. Primers used were as follows:

Otx mutation:
for 5′-ACAATGTAATTTTTTATTAATCGATATCATCATAATAGGCCTATTAATAT-3′;
rev 5′-ATATTAATAGGCCTATTATGATGATATCGATTAATAAAAAATTACATTGT-3′.
Δ29(Sox) deletion:
for 5′-TTACAGATTCAGAATTATAGATCTATTAATGAGATTAATCATAAT-3′;
rev 5′-ATTATGATTAATCTCATTAATAGATCTATAATTCTGAATCTGTAA-3′.

Microinjection of DNA constructs and synthetic RNA

Microinjection was conducted as described (Cavalieri et al., 2003, 2007). Approximately 5,000 molecules of either the desired plasmid DNA or the appropriate PCR product were injected into the zygote, together with Texas Red-conjugated dextran (in some selected trials) added at a concentration of 5% in a 2 pl volume of 30% glycerol. In the competition experiments, double-stranded *GAGA* oligonucleotides were ligated with T4 DNA ligase (Biolabs) and fractionated onto polyacrylamide gel. DNA fragments containing four to six tandem copies were eluted from the gel and mixed with the plasmid solution to be microinjected, at the ratio of 50:1.

For mRNA injection, *En-OtxHD* (Li et al., 1999) and *CS2+nlsEn* (Cavalieri et al., 2003) constructs were linearized and transcribed *in vitro* using the Sp6 or T3 mMessage mMachine kit (Ambion). Capped mRNAs were resuspended in ultrapure RNase-free water (Gibco) and 2 pl, corresponding to the amounts described in the legend to figures, were injected.

Injected embryos at the desired stage were harvested, mounted on glass slides and examined under an epifluorescence Leica DM-4500B microscope. DIC, bright-field, or fluorescence images were captured with a Leica DC 300F digital camera and processed using the Adobe Photoshop 7.0 software.

RNA extraction, reverse transcription and real-time quantitative PCR

The amounts of GFP transcription driven by either the 1.45GFP construct or its mutated forms, at morula (6 h) and late gastrula (26 h) stages were evaluated as follows. Total RNA from batches of 150 microinjected embryos was extracted by using the High Pure RNA Isolation kit (Roche). To fully eliminate any residual DNA contamination, RNA samples were treated with reagents provided by the Turbo DNA-*free* kit (Ambion), according to the conditions suggested by the manufacturer, and resuspended in a final volume of 30 μl. Reverse transcription into cDNA was performed in a 80 μl reaction using random hexamers and the TaqMan Reverse Transcription Reagents Kit (Applied Biosystems). The resulting cDNA sample was further diluted and the equivalent amount corresponding to two to five embryos was used as template for Q-PCR analysis. Primer sets were chosen to amplify products of 100–150 bp in length. Q-PCR experiments were performed from two different batches and all reactions were run in triplicate on the 7300 Real-Time PCR system (Applied Biosystems) using SYBR Green chemistry (Applied Biosystems). ROX was used as a measure of background fluorescence and MBF-1 mRNA, which is known to be expressed at a constant level during development (Alessandro et al., 2002), was used to normalize all data, in order to account for fluctuations among different preparations. At the end of the amplification reactions we run a 'melting curve analysis' to confirm the homogeneity of all Q-PCR products. Calculations from Q-PCR raw data were performed by the RQ Study software version 1.2.3 (Applied Biosystem), using the comparative Ct method (ΔΔCt). The oligonucleotide sequences were as follows:

GFP amplicon (152 bp): 5′-AGGGCTATGTGCAGGAGAGA-3′ (forward) and 5′-CTTGTGGCCGAGAATGTTTC-3′ (reverse); *Hbox12* amplicon (124 bp): 5′-ACGTCTTCGTCGAGCATCTC-3′ (forward) and 5′-GCATGGTGTCTTTCGCTTACG-3′ (reverse); *MBF-1* amplicon (102 bp): 5′-ATGACACAGCCTGGAGCT-3′ (forward) and 5′-TACCAAGGAAGTGGGTGT-3′ (reverse).

Chromatin ImmunoPrecipitation

Chromatin ImmunoPrecipitation (ChIP) assay was performed essentially as described (Di Caro et al., 2007), with minor modifications. Briefly, formaldehyde cross-linked sea urchin embryos at morula and gastrula stages were washed three times with cold PBS, collected by centrifugation and incubated in cell lysis buffer (10 mM hydroxyethyl piperazine-ethane-sulfonic acid, pH 8.0 and 85 mM KCl, 0.5% NP40, containing the following protease inhibitors: 1 μg/ml leupeptin, 1 μg/ml aprotinin, 1 mM PMSF), for 10 min on ice. Nuclei were pelleted by centrifugation at 2000 *g* for 5 min, resuspended in nuclear lysis buffer (50 mM Tris pH 8.1, 10 mM EDTA and 1% SDS, containing the same protease inhibitors as in cell lysis buffer) and incubated on ice for 10 min. Chromatin was sonicated using a microtip on a Branson Sonifier to an average fragment size of 0.5 kb, as determined by agarose gel electrophoresis. To reduce non-specific background, the samples were diluted into five volumes of ChIP dilution buffer (16.7 mM Tris–HCl, pH 8.1, 167 mM NaCl, 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, plus proteinase inhibitors) and incubated with 100 μl of a salmon sperm DNA/protein A-sepharose slurry (Sigma) for 1 h at 4 °C with mixing. Ten percent of chromatin, cleared by centrifugation, was withdrawn (input control) and processed as the immunoprecipitated chromatin. For each ChIP experiment aliquots of chromatin containing 25 μg of DNA were incubated with the anti-*Sp*Otx or the anti-MBF-1 serum overnight at 4 °C. As a negative control, the same amount of chromatin was incubated in the absence of antibodies. The immune complexes were adsorbed to protein A-sepharose. The beads were washed for 5 min, on a rotating platform, with a low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.1, 150 mM NaCl), a high salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.1, 500 mM NaCl), a LiCl buffer (0.25 M LiCl, 1% NP40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris pH 8.0) and twice in 1×TE buffer (10 mM Tris–HCl, 0.1 mM EDTA pH 8.0). The immunocomplexes were eluted with the elution buffer (1% SDS, 0.1 M $NaHCO_3$), digested with RNase at 37 °C and treated with proteinase K in 0.3 M NaCl at 65 °C for 4 h to reverse the cross-links. DNA from chromatin samples was extracted with phenol/chloroform, precipitated with ethanol and dissolved in 50 μl of Ultrapure water (Gibco). DNA samples were then quantified by readings in a Qubit Fluorometer (Invitrogen) using the Quant-iT dsDNA HS assay kit (Invitrogen), following the manufacturer's recommendations. The enrichment of *Hbox12* regulatory sequences in genomic DNA purified from the precipitated chromatin fractions was examined by PCR amplification in the linear range. For PCR reactions, 100 pg of the immunoprecipitated chromatin were used as template with the following sets of primers. *Hbox12-ChIP* (forward) 5′-GGAGAGAAGTTGTGAGAGAGC-3′ and *Hbox12-ChIP* (reverse) 5′-AGGCCTATTATGATTAATCTCAT-3′. PCR reaction conditions were: 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min. Amplifications were performed in the linear range for 30 cycles and the products were analyzed in 2% agarose gels.

Results

Genomic structure of the PlHbox12 locus

Unpublished evidence indicated that multiple gene copies of *Hbox12* exist in the *P. lividus* genome. Indeed, sequencing of a lambda-genome clone revealed that it included two complete tandem repeated copies of

the gene, named *Hbox12-a* and *Hbox12-b* respectively. The gene organization was derived either by aligning the sequence of the lambda-clone with that of the full-length *Hbox12* cDNA and by analyzing it with the Genscan software. A diagram of the *Hbox12* gene anatomy reported in Fig. 1A shows that both genes share an identical structure. Indeed, the entire DNA sequence of a single gene unit encompasses two exons, split by a single intron, together extending over almost 1 kb of the genome. Both genes are transcribed in the same direction and the open reading frame begins within exon 1, which comprehends codons 1–66 and is preceded by a 75 bp of leader sequence. The coding region is interrupted by the unique intron of 300 bp, placed between codons 66 and 67, corresponding to amino acids 46–47 of the homeodomain. The homeobox of several members of the pair-class, such as *ceh-10*, *goosecoid*, and *otx*, is analogously interrupted by an intron at the same position (Burglin, 1994). Exon 2 also includes the carboxyl-end of the coding region and is followed by a 130 bp of 3′-UTR and by the polyadenilation site. In a *in silico* search for binding sites along the ~1.45 kb upstream sequence of each *Hbox12* gene unit, a non-canonical TATA-box and two CCAAT-boxes (one of which inverted) were found, respectively 30 bp, 85 bp and 105 bp upstream of the putative transcription initiation site (Fig. 1A). The sequence similarities of the two genes were close to 99%, except for the 5′-flanking region (95%). Importantly, a third partial *Hbox12* copy is located downstream of the *Hbox12-b* in the lambda-genome fragment, just interrupted by the cloning site. It only contains ~300 bp of 5′-flanking region closely related to that of the other two genes. It follows, that at least three tandem arrayed *Hbox12* genes exist in the *P. lividus* genome and their structural features suggest that duplications have occurred at this locus in the course of evolution.

An extensive search into the available data banks failed to find the true orthologs of *Pl*Hbox12 in other sea urchin species. To date,

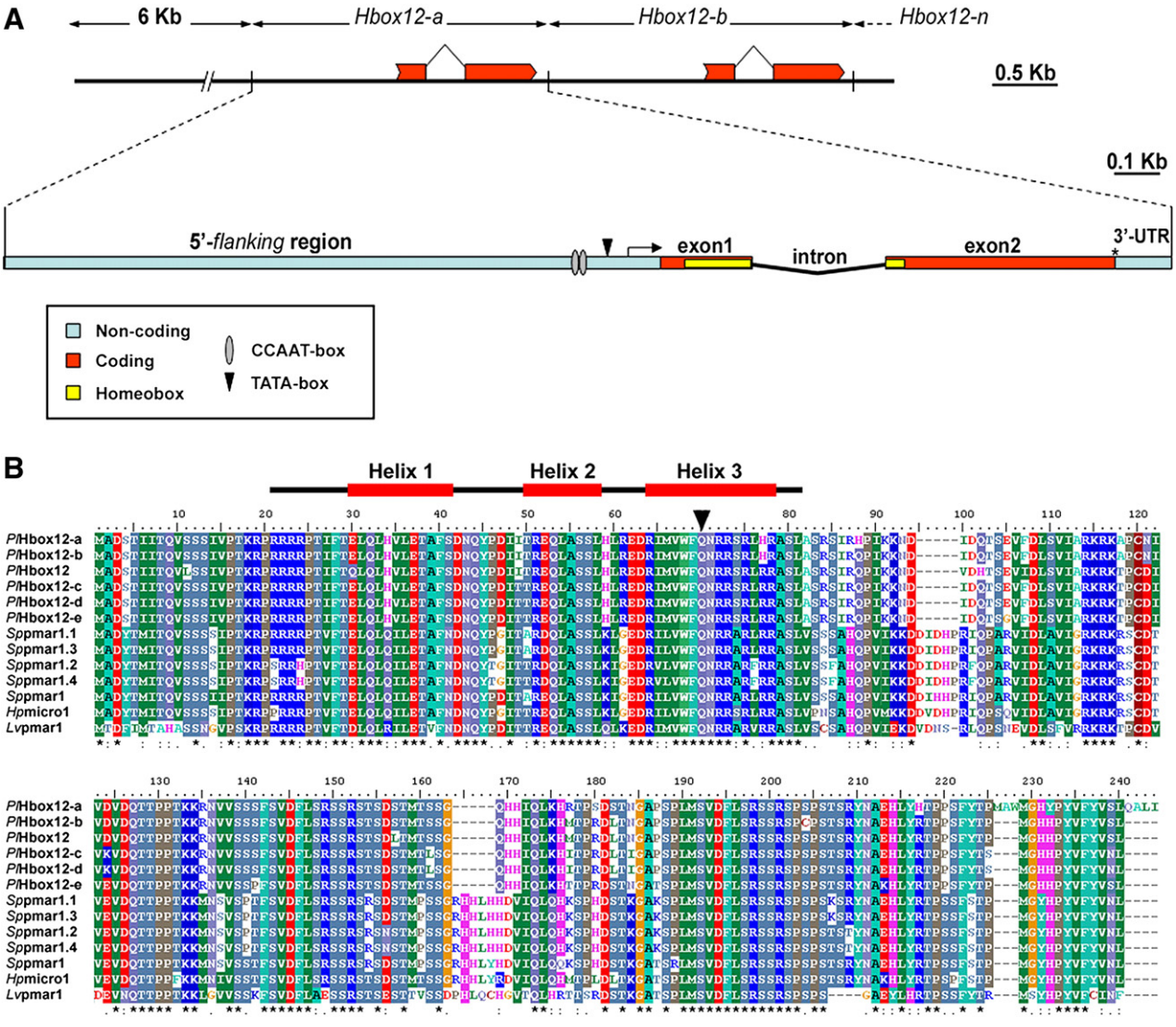

Fig. 1. (A) Diagrammatic representation of the *Hbox12* gene structure. The annotation of a lambda-genome clone shows the tandem array of two *Hbox12* gene copies, namely *Hbox12-a* and *Hbox12-b*. Note that a third partial copy, *Hbox12-n*, is located downstream. The horizontal black line and red boxes respectively represents the genomic DNA and the two exons. The bent arrow close to exon1 denotes the putative transcription start site. The size and position of the homeodomain is shown in yellow-coloured box, extending between exon1 and exon2. A black arrowhead indicates the TATA-box sequence, while the two light grey ovals indicate CCAAT-boxes. (B) Multiple ClustalW sequence alignment of the *Pl*Hbox12, *Sp*pmar1, *Lv*pmar1 and *Hp*micro1 proteins. Identical residues in all of the proteins are pointed by asterisks, while double dots and single dots indicate decreasing degrees of conservation. Divergent amino acids are indicated by blank spaces. Filled boxes indicate amino acids that are identical in at least six of the aligned proteins. Dashes represent the gaps inserted for maximal alignment. The Glutamine at position 50, conserved in the paired-*like* class homeodomain, is marked by an arrowhead. The red boxes indicate the position of the three helices of the homeodomain. *Pl*Hbox12 amino acidic sequence was derived from a translated cDNA clone, accession number X83675 (Di Bernardo et al., 1995). Sequences used for alignment, and their accession numbers, were: *Sp*pmar1, NM_214508 (Oliveri et al., 2002); *Lv*pmar1, DQ667003 (Wu et al., 2007); *Hp*micro1, AB180907 (Nishimura et al., 2004); *Sp*pmar1.1–1.4 sequences were retrieved by a Genscan analysis of the following BAC items: NW_001304149 and NW_001292071.1.

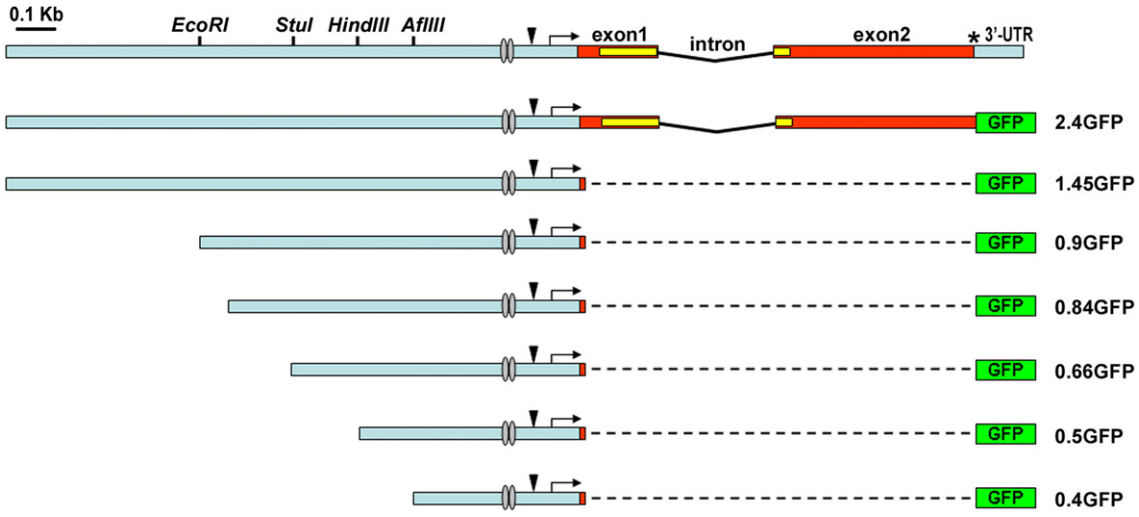

Fig. 2. Schematic drawing of the GFP *transgene* constructs driven by the *Hbox12 cis*-regulatory apparatus and 5′-deletions. At the top is a map of the *Hbox12* gene structure, in which unique sites for restriction enzymes, in the 5′-flanking region, are indicated. A black arrowhead indicates the TATA-box sequence, while the two light grey ovals indicate CCAAT-boxes. The bent arrow denotes the putative transcription start site. The name of each construct is listed on the right.

proteins that show the higher sequence identity are encoded both by the *pmar1* genes identified in *S. purpuratus* (Oliveri et al., 2002) and *L. variegatus* (Wu and McClay, 2007), and by the *pmar1*-related *micro-1* genes of *H. pulcherrimus* (Kitamura et al., 2002) and *A. crassispina* (Nishimura et al., 2004). All these homeobox genes belong to the Q50 paired-*like* class, which is defined as having a homeobox similar to that of the paired-class genes, but lacking the paired box that encodes a second DNA-binding domain (Burglin, 1994). An alignment of the deduced amino acidic sequences of *Hbox12 a–e* (two genomic and four cDNA sequences) shown in Fig. 1B revealed that they were 94–100% identical to each other. Minor differences were found in the sequences external to the homeodomain, probably due to polymorphisms among the individual urchins used to generate the libraries. For instance, the *Hbox12-a* gene product is globally seven amino acids longer than *Hbox12-b*. It contains an *in frame* tripeptide inserted after residue 196, and an additional C-terminal tetrapeptide (Fig. 1B), due to a single point mutation of the first position of the stop codon. On the contrary, there was only an average of 74% of sequence identity between *Hbox12* and *pmar1*/*micro1* proteins. Notably, significant differences were found even in the homeodomain (Fig. 1B).

The 5′-flanking region from the Hbox12 gene drives expression of a GFP reporter gene in aboral ectoderm founder cells

Due to the tandem repeat organization, the *cis*-regulatory apparatus of *Hbox12* is likely to reside in the gene unit either within the 1.45 kb of genomic upstream sequence and/or in the intron. To identify putative *cis*-regulatory modules, a computational analysis by inter-specific sequence comparison has been successfully used by other authors (Amore and Davidson, 2006; Lee et al., 2007; Nam et al., 2007; Ransick and Davidson, 2006). This approach, however, does not seem to be applicable for *Hbox12* because of the apparent lack of true orthologs in other sea urchin species. Hence, in order to identify DNA elements required for the proper regulation of *Hbox12* in the embryo, we prepared a series of reporter constructs for gene transfer experiments. A diagram of these *transgenes* is depicted in Fig. 2. As a first step, we PCR-amplified a single fragment of 1.45 kb in length, abutting at the 3′-end the ATG start codon of the *Hbox12-a* gene, and fused it *in frame* with the Green Fluorescent Protein (GFP) reporter gene. The construct is referred to as 1.45GFP (Fig. 2). The linearized DNA was then microinjected into sea urchin zygotes, embryos were allowed to develop and scored for GFP expression. The quality of each batch of eggs used was monitored with respect to their ability to develop normally. As a control of the expression of the injected *transgenes*, we used an actively transcribed GFP construct driven by the full promoter of the hatching enzyme gene. A summary of the results of three different microinjection experiments shows that the 1.45GFP construct was expressed in an average of 60% of injected embryos (Table 1), in close agreement with the localization of *Hbox12* transcripts at the 60-cell stage and the lineage map. This percentage is a bit lower than the number of positive embryos (almost 85%) expressing the control construct (not shown). Such a difference can be easily explained by the reduced number of blastomeres expressing

Table 1
Spatial expression of *Hbox12* promoter-GFP *transgene* constructs in microinjected *P. lividus* embryos

Injected *transgenes*	% GFP-expressing embryos[a,b]	Territory of expression[c]					
		Aboral ectoderm	Oral ectoderm	Ciliary band	Endoderm	SMCs	PMCs
1.45GFP	60.6 (±1.3)	78.4 (±2.1)	1.9 (±0.6)	8.2 (±1.8)	19.8 (±2.5)	–	1.8 (±0.7)
0.9GFP	59.3 (±1.8)	77.8 (±2.4)	2.1 (±0.8)	8.8 (±1.8)	21.9 (±1.9)	1.0 (±0.6)	1.5 (±0.6)
0.66GFP	68.7 (±1.8)	41.8 (±2.2)	11.4 (±1.8)	14.6 (±2.9)	43.7 (±1.5)	13.3 (±0.8)	12.3 (±1.6)
0.5GFP	67.9 (±2.2)	38.6 (±4.6)	12.2 (±2.0)	14.0 (±1.2)	41.9 (±1.8)	14.1 (±0.8)	15.6 (±2.1)
0.4GFP	4.7 (±2.5)[d]	26.1 (±16.2)	33.9 (±28.2)	10.0 (±20.0)	10.0 (±20.0)	8.4 (±16.7)	11.7 (±13.3)

All values are mean percentages of three independent trials, with standard error of the mean in parentheses. Each experiment was carried out on about 150 eggs of a single batch. Abbreviations: SMCs, secondary mesenchyme cells; PMCs, primary mesenchyme cells.

[a] Embryos were scored as GFP-expressing only if two or more cells were fluorescent.

[b] % GFP-expressing=(total number of GFP-expressing embryos/total number of injected embryos)×100.

[c] Fraction of total embryos that displayed GFP fluorescence in the indicated cell types, i.e. (embryos expressing GFP in indicated cell type/total number of GFP-expressing embryos)×100. Values for each cell type are calculated independently of GFP expression in other cell types.

[d] Insignificant level of expression was obtained with the 0.4GFP construct.

Hbox12 with respect to the hatching enzyme gene and by the mosaic incorporation of the exogenous DNA in the sea urchin embryo (Flytzanis et al., 1985; Franks et al., 1988; Hough-Evans et al., 1988; Lepage et al., 1992). Approximately 78% of positive embryos expressed the *transgene* in the aboral ectoderm (Table 1). Fig. 3 shows GFP expression in a number of representative living embryos at different stages of development. As expected from the timing and abundance of the endogenous transcripts, coupled to the delay of fluorescence detection due to the kinetics of GFP accumulation, patches of clonally related *transgene* expression were detected from the early blastula stage (Fig. 3A). At this stage of development it is unfeasible to distinguish among different cell types in the embryo only on the basis of morphological features. At the mesenchyme blastula stage, as PMCs begin to ingress the blastocoel, green spots of GFP fluorescence were observed at the animal hemisphere of the embryos, in cells that are clearly part of the embryonic ectoderm territory (Figs. 3B, C). At gastrula stage, in which all cell types have been specified and can be easily recognized, expression of GFP was seen almost exclusively in the aboral ectoderm (Fig. 3D; Table 1). This is clearly evident in embryos at later development. At prism and pluteus stages patches of GFP-expressing cells were frequently seen at or close to the vertex (Figs. 3E–J), either on the dorsal (Figs. 3F, G) or the anal side (Fig. 3H) or on both sides (Figs. 3I, J), with respect to the left–right axis. Some sibling embryos (~8%), like that shown in Fig. 3K, expressed the construct even in cells of the ciliary band, which include intercalated cells of aboral ectodermal origin, as shown by lineage tracing (Cameron et al., 1993). In other embryos (~20%), like those shown in Figs. 3L, M, expression of the 1.45GFP construct also occurred in few cells located at the aboral side of the invaginated *archenteron*. To correctly interpret this result it should be emphasized that the allocation of veg1 cells to ectoderm and endoderm during cleavage is highly variable and does not occur predictably (Logan and McClay, 1997). In fact, it seems to correlate with their proximity to the underlying veg2 cells and/or the overlying mesomere progeny (Cameron and Davidson, 1997). Furthermore, mesomeres of 16/32-cell stage embryos also contribute, although at a low frequency (16–20%), to the endoderm (Logan and McClay, 1997; Sherwood and McClay, 2001). Altogether, these phenomena depend, at least in part, on the variation of the third cleavage plane along the animal–vegetal axis, which may vary in different egg batches (Cameron and Davidson, 1997; Logan and McClay, 1997). In light of these evidence and since *Hbox12* transcripts have shown to be restricted toward one side of the animal cap and veg1 tier (Di Bernardo et al., 1995), it is not too surprising to find patches of expression even at the aboral side of the hindgut (Fig. 3L) or midgut (Fig. 3M) of some embryos. Ectopic expression was rarely observed in oral ectoderm or mesenchyme cells, but remained below significant levels in all cases examined during this analysis (Table 1).

In a subsequent screen for additional potential *cis*-regulatory modules localized into the intron sequences, a segment of *Hbox12-a* genomic DNA, spanning all the 5′-flanking and the entire coding region, was joined to the GFP reporter gene, to give the 2.4GFP

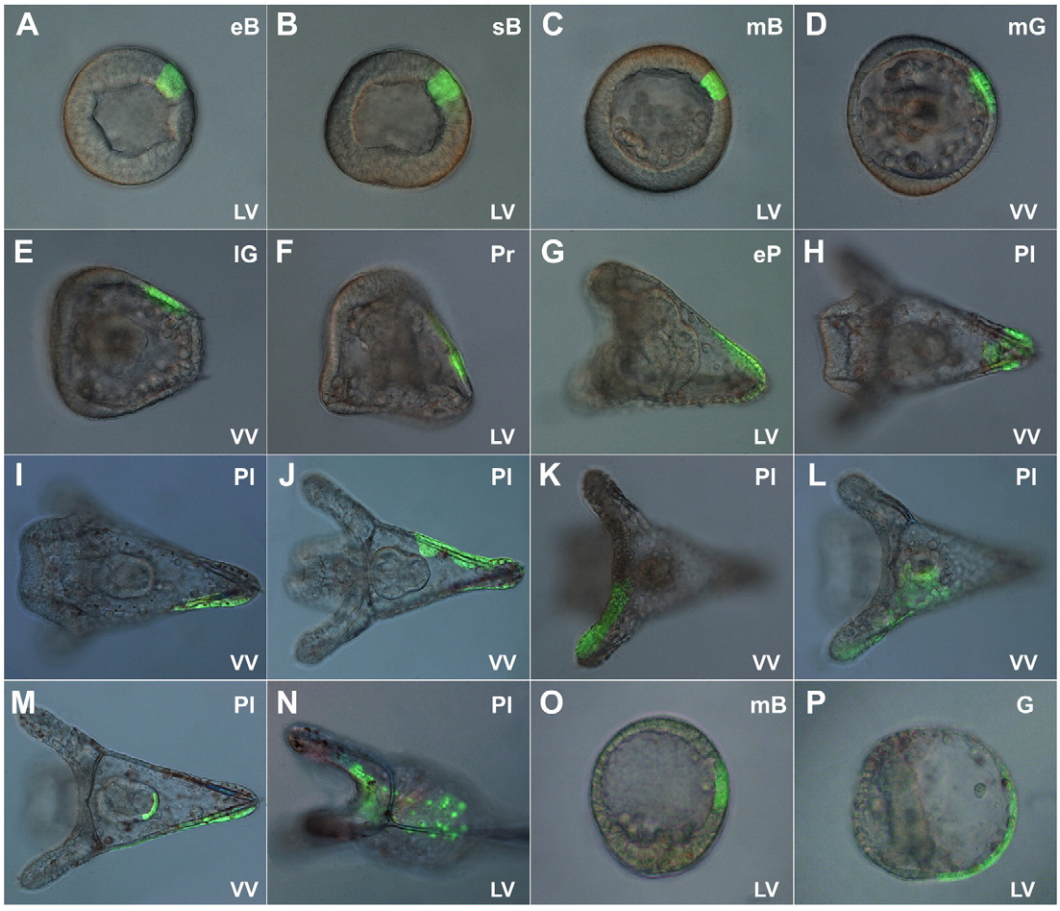

Fig. 3. Expression of the 1.45GFP and 2.4GFP *transgene* constructs during embryonic development of *P. lividus* (A–N) and *S. purpuratus* (O–P). GFP fluorescence image overlays from live *transgene* injected embryos are shown. The upper right corner of each image is labelled with its developmental stage, and the orientation is indicated on the lower right corner. In embryos microinjected with 1.45GFP construct and observed at early stages of development (A–C, O), GFP green spots are always localized clonally in groups of cells which are fated to constitute the aboral ectoderm epithelium of the larvae (D–J, P). (K) Expression of the *transgene* in the ciliary band, which comprehends both oral and aboral ectoderm cell type precursors. (L–M) Pluteus stage embryos expressing the *transgene* both in aboral ectoderm and respectively in the aboral side of midgut (L) and hindgut (M). (N) Expression of the 2.4GFP construct in the aboral ectoderm territory of a pluteus stage embryo. Abbreviations: eB, early blastula; hB, blastula; sB, swimming blastula; mB, mesenchyme blastula; mG, mid-gastrula; G, gastrula; lG, late gastrula; Pr, prism; eP, early pluteus; Pl, pluteus; LV, lateral view with the animal pole at the top; VV, ventral view with the oral side to the left.

construct. Embryos injected with this *transgene* construct showed no differences in the spatial distribution of GFP with respect to that described for 1.45GFP construct (Fig. 3N). The only difference concerns the intracellular localization, which is nuclear for the Hbox12-GFP fusion protein. An identical GFP localization was observed in embryos bearing analogous *transgene* constructs corresponding to the *Hbox12-b* gene (not shown). These results strongly indicate that *Hbox12-a* and *-b* genes are co-expressed and, more importantly, that the intron sequences do not contain *cis*-regulatory elements required for the proper spatial expression.

Interestingly, gene transfer assays of the 1.45GFP construct conducted in *S. purpuratus* embryos showed a spatial *transgene* expression preferentially restricted to some cell clones contributing to the aboral ectoderm, both at mesenchyme blastula (O) and gastrula stage (P). This finding suggests the evolutionary conservation in *S. purpuratus* of the pool of *trans*-acting factors that bind the *cis*-regulatory apparatus of the *P. lividus Hbox12* gene.

An aboral ectoderm cis-regulatory module is required for the proper Hbox12 expression

As a first approach to dissect promoter elements required for the transcriptional regulation of the *Hbox12* gene, we created a series of deletions from the 5′-end of the 1.45GFP *transgene* using convenient restriction sites or PCR amplification reactions (Fig. 2). Once again, each linearized construct was microinjected into three distinct batches of fertilized eggs and embryos were scored for GFP expression. Values reported in Table 1 give a statistical essence of the results. Embryos representative of the spatial GFP localization are shown in Fig. 4. The 0.9GFP construct, bearing a 5′-truncation that removed 500 bp, consistently reproduced the aboral ectoderm-specific regulation pattern observed with the 1.45GFP construct (Table 1). Next we tested the effect of some 5′-deletions in the 0.9 kb promoter fragment. Two constructs lacking respectively 234 bp (0.66GFP) and 395 bp (0.5GFP), showed a similar behaviour. With both we observed a slight increase, up to 8%, of GFP-expressing embryos (Table 1). Although spots of GFP fluorescence were still detected in the aboral epithelium lineage (~40% of embryos), a broad ectopic appearance of GFP-expressing cell patches was concomitantly observed in inappropriate embryonic tissues throughout development (Fig. 4; Table 1). In particular, about 43% of embryos expressed the *transgene* in the vegetal plate territory during the pregastrular phase (Fig. 4A) and later, GFP fluorescence was observed in large cell patches of the *archenteron* (Figs. 4B–E), either at the oral (Fig. 4C) or aboral side (Fig. 4D). Notably, vegetal-expression was extended even to the foregut compartment (Figs. 4C, D) and in some SMCs (Figs. 4C, E). In sibling embryos (~12%) expression of the *transgenes* was detected in either the apical plate at mesenchyme blastula stage (Fig. 4F) or in the ventrolateral ectoderm, strictly close to the triradiate spicule elements, at gastrula stage (Figs. 4G, H). These ectoderm territories give rise to the various oral areas and indeed, GFP expression persisted in the ectoderm just surrounding the mouth (Fig. 4J) and at the tips of the arms of plutei (Fig. 4I). In other embryos (~14%) GFP fluorescence was initially seen only in few of the ingressing PMCs of the mesenchyme blastula stage (Fig. 4K), but diffused to all skeletogenic cells at later stages (Figs. 4L–O), because of the establishment of syncytial cables among cells. Taken together, this randomized pattern of *transgene* expression suggests that the 0.66 kb and 0.5 kb promoter fragments probably lack negative regulatory sequence elements that maintain the highly restricted *Hbox12* expression in the aboral ectoderm founder cells. Removing all, but the 355 bp closer to the putative transcription start site (0.4GFP), almost abolished the expression of the *transgene* (see below). From this

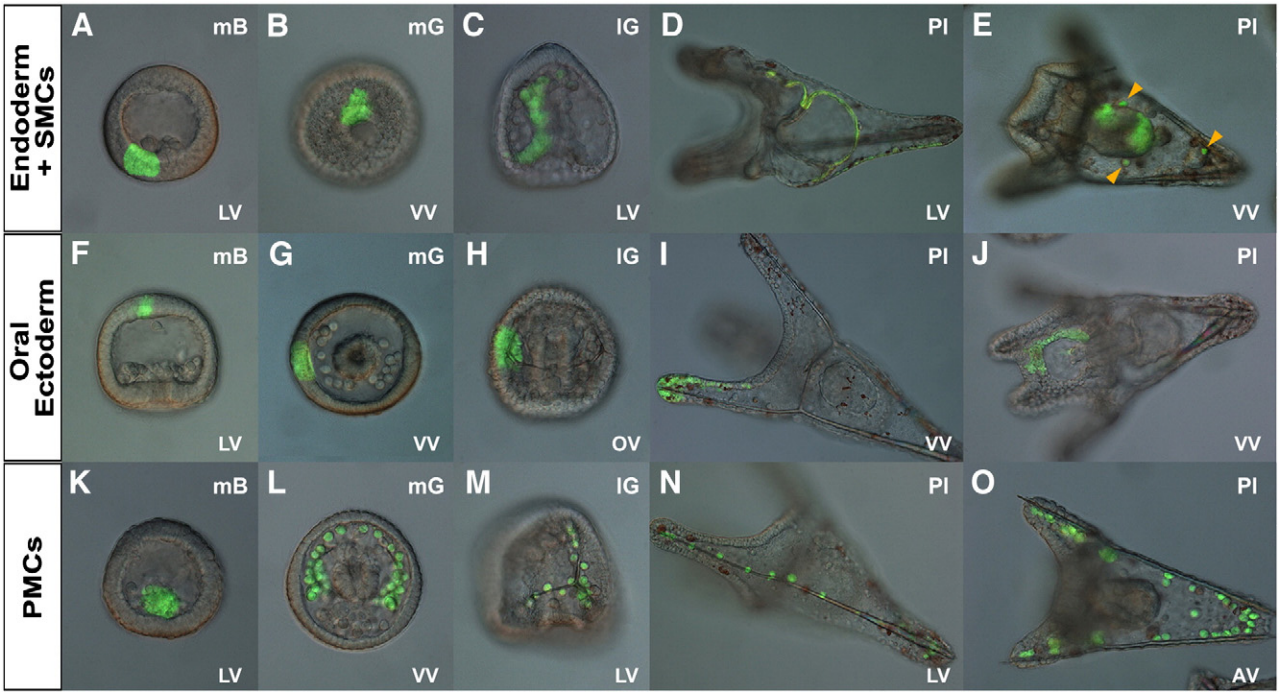

Fig. 4. Classification of the spatial expression patterns observed throughout embryonic development, following injection of 0.66GFP and 0.5GFP DNA constructs. Spatial expression profiles beyond aboral ectoderm (not shown) are classified into three major groups, indicated in each row as: Endoderm+SMCs (A–E), Oral Ectoderm (F–J), and PMCs (K–O). The upper right corner of each image is labelled with its developmental stage, and the orientation is indicated on the lower right corner. (A–E), *Transgenic* embryos expressing GFP in the vegetal plate during the pregastrular phase (A), and at both oral (C) and aboral (D) sides of the *archenteron* at later stages. (E) Pluteus stage embryo in which GFP expression was found even in some SMCs, as indicated by orange arrowheads. (F–J) Microinjected embryos showing GFP localization in oral ectoderm territories. (I) Pluteus stage embryo with GFP fluorescence in clonal cells at the tip of one of the post-oral arms. A focus plane on the mouth in (J) allows identifying of GFP fluorescence in some cells of the oral ectoderm epithelium, just surrounding the mouth. (K–O) Expression of the *transgenes* in skeletogenic cells. At later stages, GFP diffused to all PMCs due to the establishment of syncytial cables among cells, and the green fluorescence was clearly visible along the spicule elements of these embryos. Abbreviations: mB, mesenchyme blastula; mG, mid-gastrula; lG, late gastrula; Pl, pluteus; LV, lateral view with the animal pole at the top; VV, ventral view; OV, view from the oral ectoderm.

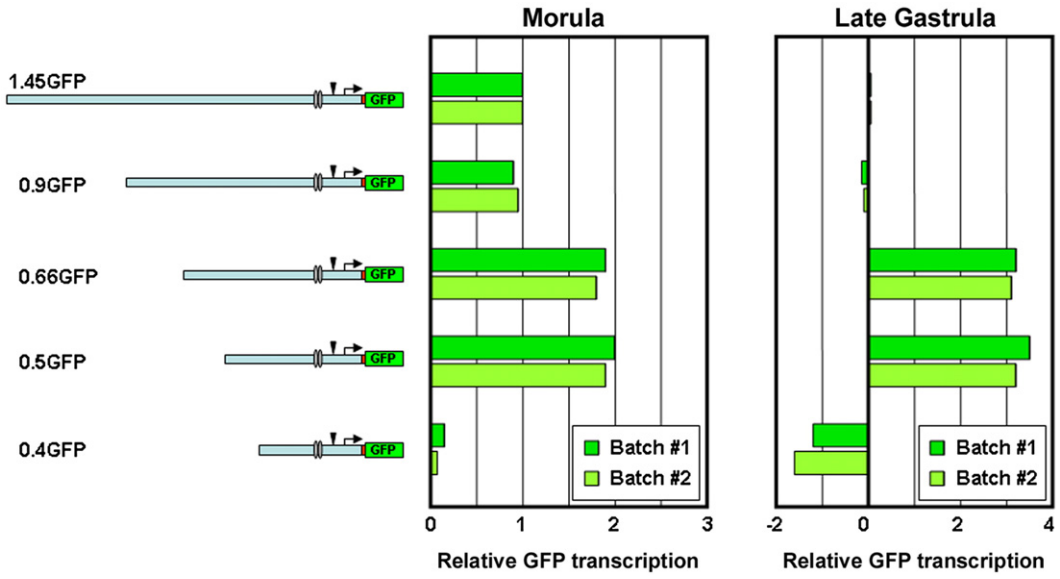

Fig. 5. Q-PCR analysis of the relative transcriptional activity generated by the 5′-truncated *Hbox12* promoter *transgenes*. Graphs show *n*-fold changes in mRNA expression level of *GFP* based on the threshold cycle number (Ct) of 5′-truncated constructs compared to that of the 1.45GFP control *transgene*. Ct numbers were normalized for the endogenous *MBF-1* in the same sample. Data were derived from two independent microinjection experiments and each bar represents the average of triplicate samples from a given batch of embryos.

observation, we can argue that most of the important positive *cis*-regulatory elements were lost in the 0.4GFP construct.

To further determine the effect of deletions of the *Hbox12* promoter on the expression of the *transgene*, we measured the GFP mRNA amount by Quantitative RT-PCR (Q-PCR) in two independent batches of embryos at morula and late gastrula stages (Fig. 5). The kinetics of both 1.45GFP and 0.9GFP constructs qualitatively resembled that of the endogenous *Hbox12*. Indeed, *transgene* transcription was detected at morula stage, the time in which *Hbox12* peaks its expression, while small amounts of mRNA were present in embryos at late gastrula stage, when *Hbox12* transcription has been turned off. Since the turnover rate of GFP mRNA is not known in these cells, we cannot be exactly sure when the transcriptional activity of the *transgene* terminates. Nevertheless, we can conclude that the 0.9 kb promoter fragment accurately recapitulates, both spatially and temporally, the early aboral-specific expression pattern of *Hbox12*.

At morula stage, injection of both the 0.66GFPand 0.5GFP constructs led to a nearly two-fold higher expression of GFP than that of the 1.45GFP and 0.9GFP constructs (Fig. 5). This result can be probably ascribed to the broad ectopic expression described above. Moreover, the 0.66 kb and 0.5 kb *transgene* constructs were not down-regulated at late gastrula (Fig. 5), demonstrating that the fundamental elements required for the temporal regulation lay within the −0.9 to −0.66 kb promoter region. In agreement with microscopic observations carried out throughout development, the abundance of the GFP RNA in the 0.4GFP injected embryos dropped precipitously and remained very low. Taken together these results strongly suggest that the genomic region comprised between −0.9 and −0.66 kb contains a compact *cis*-regulatory module, defined as *A*boral *E*ctoderm (AE) module. It should contain most, if not all, of the sequence elements necessary for proper temporal and spatial expression of the *Hbox12* gene.

Identification on the AE module of the Otx and putative Sox binding site as positive inputs for Hbox12 gene expression

A search of transcription factor binding sites within the AE *cis*-regulatory module using the MatInspector and TESS softwares (Cartharius et al., 2005; Quandt et al., 1995) revealed the presence of several potential binding sites for transcription factors. In addition, we noticed the presence of a purine region containing a stretch of sixteen *GA* tandem repeats (Fig. 6). In the minus strand of the 3′-most

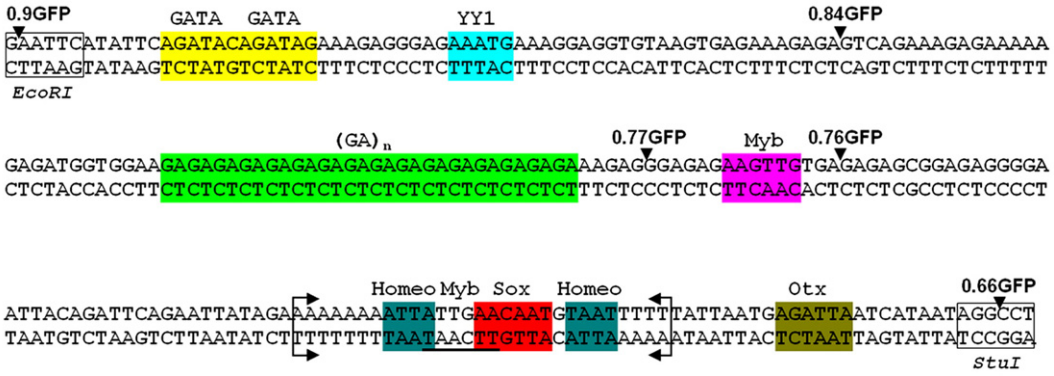

Fig. 6. The entire 234 bp nucleotide sequence of the AE *cis*-regulatory module is shown and annotated. It extends from the EcoRI and StuI restriction sites, indicated by hatched boxes, of the 5′-flanking region of the *Hbox12* gene. Relative positions of various putative *cis*-regulatory elements are indicated by coloured boxes, with the exception of a Myb-*like* consensus, which is underlined. The termini of 5′-deletion mutants are pointed by black triangles and construct names indicated in bold characters. Bent arrows on both strands delimitate the internal deletion employed to eliminate the Sox site.

downstream region there is a *TAATCT* motif. This sequence element has been demonstrated to bind the Otx factor with high affinity (Hanes and Brent, 1989, 1991; Treisman et al., 1989; Wilson et al., 1993). To assess the functional role of this putative *cis*-regulatory sequence element, we preliminarily disrupted the Otx function by co-injecting into the sea urchin zygotes an *in vitro* transcribed mRNA encoding for a forced En-OtxHD repressor, in which the homeodomain of *Sp*Otx was joined to the repression domain of *Drosophila* Engrailed (Li et al., 1999), along with the 0.9GFP construct. Injection of a control RNA encoding for the En-grailed repressor domain at the same doses of En-OtxHD had no effect on *transgene* expression and embryonic development (Fig. 7A). Conversely, and according to what described by other authors (Li et al., 1999), the ectopic expression of the En-OtxHD repressor had drastic effects on aboral ectoderm and endoderm differentiation, as judged morphologically. Embryos observed at late gastrula stage (Fig. 7B) appeared to be constituted by a uniformly thickened epithelium that had oral ectoderm characteristics and no discernible ciliary band was identified. The putative mesenchyme cells adopted a radial distribution and the *archenteron* failed to connect the ectoderm. As expected, these embryos did not produce detectable GFP fluorescence (Fig. 7B′). We next performed Q-PCR analysis at morula stage, to determine both *transgene* and *Hbox12* expression in the En-OtxHD injected embryos. Strikingly, overexpression of En-OtxHD caused a dose-dependent attenuation in the level of GFP and endogenous *Hbox12* mRNA (Fig. 7C). Injection of 0.08 pg of En-OtxHD RNA almost completely impaired gene expression.

To definitively prove the effective binding of Otx(α) to the AE module of the *Hbox12* promoter *in vivo*, we performed Chromatin ImmunoPrecipitation (ChIP) assay. To this end, DNA-binding proteins of morula and gastrula stage embryos were covalently linked to genomic DNA by formaldehyde treatment, and equal amounts of sheared soluble chromatin purified from both stages were immune-precipitated by a polyclonal anti-Otx antibody (Gan et al., 1995; Mao et al., 1996) or, as a control, anti-MBF-1 antibodies (Alessandro et al., 2002). Genomic DNA purified from the precipitated chromatin fractions and DNA input were PCR-amplified in the linear range with two oligonucleotide primers flanking the Otx binding site. As shown in Fig. 7D, we found that the Otx transcription factor binds to the specific consensus site on the promoter of the *Hbox12* gene, and there is a tight correspondence between *Hbox12* transcription at morula stage and occupancy of the promoter region by the Otx transcription factor. Conversely, down-regulation of the *Hbox12* gene at gastrula stage well correlated with little or no association of this regulator with its binding site. As expected, the antiserum against the *α-H2A* histone gene activator MBF-1 (Alessandro et al., 2002; Di Caro et al., 2007) did not precipitate any detectable sequences of the *Hbox12* promoter from the chromatin of both developmental stages.

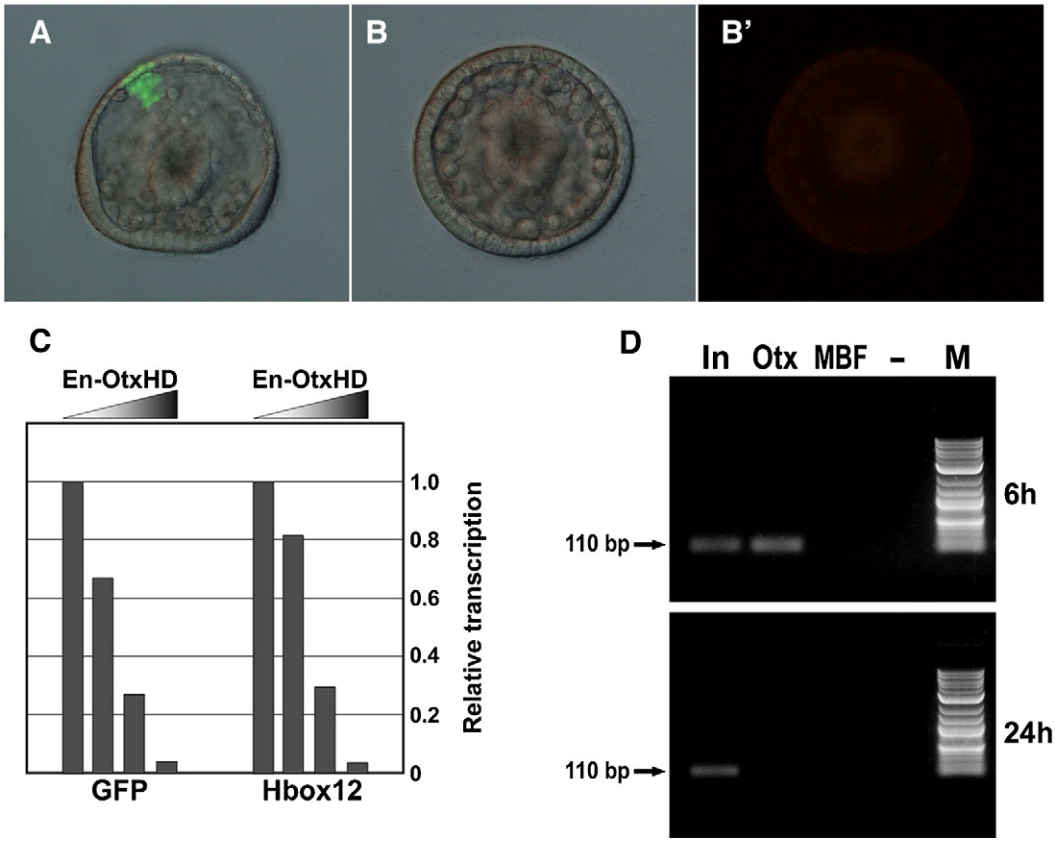

Fig. 7. Transcriptional output of the binding of Otx-Engrailed forced repressor to the *Hbox12* promoter. (A) Superimposed view of a prism stage control embryo co-injected with 0.08 pg of a *in vitro* transcribed mRNA encoding the En repressor domain and the 0.9GFP construct. The embryonic ectoderm is correctly partitioned and the reporter gene is expressed in the aboral epithelium. (B–B′) Bright-Field and epifluorescent images of a representative embryo, of the same age as in panel A, co-injected with the 0.9GFP *transgene*, the *in vitro* transcribed mRNA encoding the forced repressor En-OtxHD and Texas Red-conjugated dextran. Neither GFP fluorescence nor distinguishable ectoderm polarization is detectable in this embryo. Endoderm failed to connect the ectoderm and PMCs are radialized around the *archenteron*. (C) Q-PCR analysis of En-OtxHD co-injected embryos. Injection of 0.01 to 0.08 pg of En-OtxHD mRNA caused a precipitous drop, with a dose-dependent effect, in the expression level of both 0.9GFP *transgene* and endogenous *Hbox12* gene. (D) Chromatin ImmunoPrecipitation (ChIP) experiment. Equal amounts of soluble chromatin from cross-linked embryos at early blastula (6 h) and late gastrula (26 h) stages were precipitated with polyclonal antibodies against Otx or MBF-1, or incubated without adding antibodies (−). After reversion of the cross-link, DNA from the immunoprecipitates or input chromatin (In) was purified and amplified with *Hbox12* specific primers. The ethidium bromide staining of the agarose gel shows that the association of the Otx transcription factor to its binding site in the *Hbox12* AE module well correlates with *Hbox12* transcription. The anti-MBF-1 antibody did not show any PCR signal at both developmental stages.

Table 2
Variation of spatial expression of the *Hbox12* promoter-GFP *transgene* by perturbation of sequence elements in the *AE* module

Injected *transgenes*	% GFP-expressing embryos[a,b]	Territory of expression[c]					
		Aboral ectoderm	Oral ectoderm	Ciliary band	Endoderm	SMCs	PMCs
1.45GFP	59.8 (±1.5)	79.2 (±1.3)	1.5 (±1.0)	8.5 (±1.2)	18.8 (±1.8)	1.5 (±0.8)	1.8 (±0.7)
0.84GFP	57.9 (±1.6)	69.3 (±1.7)	2.2 (±0.8)	7.3 (±1.5)	25.7 (±1.6)	6.5 (±0.9)	24.8 (±1.6)
0.77GFP	60.1 (±1.8)	67.0 (±1.4)	1.8 (±0.5)	6.2 (±1.5)	27.5 (±1.4)	7.4 (±0.7)	22.1 (±1.8)
0.76GFP	83.1 (±2.2)	51.8 (±1.8)	33.0 (±1.6)	16.8 (±1.6)	29.5 (±1.9)	16.9 (±1.2)	16.9 (±1.4)
1.45(Δ29)GFP[d]	51.4 (±2.1)	34.4 (±2.2)	47.3 (±3.9)	8.9 (±0.7)	18.4 (±1.6)	22.4 (±4.1)	30.8 (±3.8)
Otx-mut[d]	59.7 (±1.8)	56.9 (±1.7)	1.5 (±0.9)	7.3 (±1.4)	12.2 (±1.1)	0.9 (±0.2)	2.0 (±0.6)

All values are mean percentages of three independent trials, with standard error of the mean in parentheses. Each experiment was carried out on about 150 eggs of a single batch.
[a] Embryos were scored as GFP-expressing only if two or more cells were fluorescent.
[b] % GFP-expressing=(total number of GFP-expressing embryos/total number of injected embryos)×100.
[c] Fraction of total embryos that displayed GFP fluorescence in the indicated cell types, i.e. (embryos expressing GFP in indicated cell type/total number of GFP-expressing embryos)×100. Values for each cell type are calculated independently of GFP expression in other cell types.
[d] Perturbation of the indicated sequence element was introduced in the AE module of the 1.45GFP construct.

To further confirm the involvement of the Otx activator in the expression of the *Hbox12* gene, we mutated its binding site on the *transgene*. We observed that substitution of five bases did not cause an apparent decrease in the number of the reporter-expressing embryos during development (Table 2). Nevertheless, we detected a sharp drop of the reporter distribution in the aboral ectoderm. In addition, Q-PCR measurements revealed a ~40% decrease of GFP transcripts level, when compared with the 1.45GFP wild type construct (Fig. 8). Altogether, these results fully support a direct regulatory role of Otx in *Hbox12* activation in the aboral ectoderm lineage cells.

However, the *Otx*-mutated *transgene* still drove relatively high levels of reporter transcription, suggesting that it contains other transcription factor binding sites necessary for full *Hbox12* expression. The above mentioned computational search of binding sites revealed the presence of a candidate motif, *AACAAT*, in the plus strand of the AE module (Fig. 6), that is known to be recognized by the members of the Sox factor family (van Beest et al., 2000). Such a binding site is embedded in a region of very simple sequences that resulted difficult to mutagenize. For this reason we deleted a 29 bp DNA fragment which includes the Sox site. Following injection in zygotes, we noticed a decrease in the number of GFP-expressing embryos, with respect to the 1.45GFP controls (Table 2). Strikingly, we detected similar statistic distribution in both ectoderm domains as a result of a substantial decrease in the aboral cells, coupled to an ectopic increase in the oral cells. We interpret this as a cumulative effect due to the deletion of additional negative sequences yet to be identified. Potential candidates are a non-canonical Myb-binding sequence, *ATTGAA* (Nicolaides et al., 1991), and two *TAAT* homeodomain binding sites, adjacent to the Sox site (Fig. 6). In agreement with the lower fraction of GFP-expressing embryos, Q-PCR analysis demonstrates that such a deletion produced a strong effect on the transcriptional level of the *transgene*, which decreased to ~25% of its original value (Fig. 8).

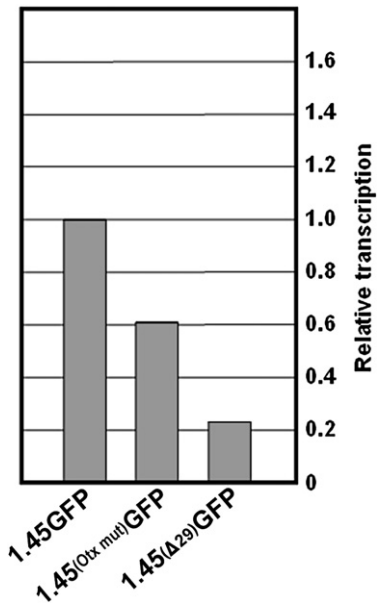

Fig. 8. Quantitative effects of site mutations and deletions on the transcriptional activity of the 0.9GFP *transgene* in embryos at morula stage. Mutation of the Otx element, as well as deletion of the 29 bp ($1.45_{(\Delta 29)}$GFP construct) sequence containing the Sox binding site cause a drastic decrease of the *transgene* transcription. Mutation and deletion of target sites contained in the AE module were tested for expression in gene transfer experiments as described in the text. Graphs show *n*-fold changes in mRNA expression level of *GFP*, based on the threshold cycle number (Ct) of the mutated constructs, compared to that of the intact 1.45GFP control. Ct numbers were normalized for the endogenous *MBF-1* in the same sample. Data were derived from two independent microinjection experiments and each bar represents the average of triplicate samples from a given batch of microinjected embryos.

Negative cis-regulatory sequences for both temporal and spatial expression

In order to identify DNA elements of the AE module responsible for the restricted spatial expression of the gene, we deleted a 60 bp sequence from the 5′-end of the 0.9GFP construct and tested the effect of this truncation on the transcriptional activity of the *Hbox12* promoter in microinjected embryos. A summary of numerical results is given in Table 2. As expected, almost 60% of embryos injected with the 1.45GFP control showed measurable GFP expression, and 79% of them expressed the reporter in aboral ectoderm cells. Only less than 4% of the *GFP*-expressing embryos gave ectopic expression confined to mesenchyme cells. The deletion of the 60 bp sequence (0.84GFP construct), had little influence on the number of positive embryos, but, noteworthy, we observed a distinct shift in the spatial expression pattern. In particular, aboral ectoderm expression decreased to ~69%, while mesenchyme cell expression surprisingly rose to ~31%. Furthermore, a similar number of embryos (19–25%) expressed the GFP reporter in the *archenteron* with both constructs, but notably we noticed that the foregut and SMCs were frequently scored as positive for GFP with the 0.84GFP *transgene* (Figs. 9A–C). These results suggest that the removal of the first 60 bp sequence of the 0.9 kb promoter altered the function of the AE control region such that ectopic expression in other territories was enhanced.

Q-PCR measurement revealed that at morula stage the GFP RNA abundance from the 0.84GFP was three fold higher than that generated by the 0.9GFP construct (Fig. 9D). This can be explained by the additional expression of the *transgene* in the ectopic vegetal lineages. No detectable RNA was found at late gastrula stage with both clones (Fig. 9D). Altogether, these results demonstrate the capability of the 60 bp sequence to repress transcription of *Hbox12* in the endomesoderm lineages. In addition, they exclude the presence in

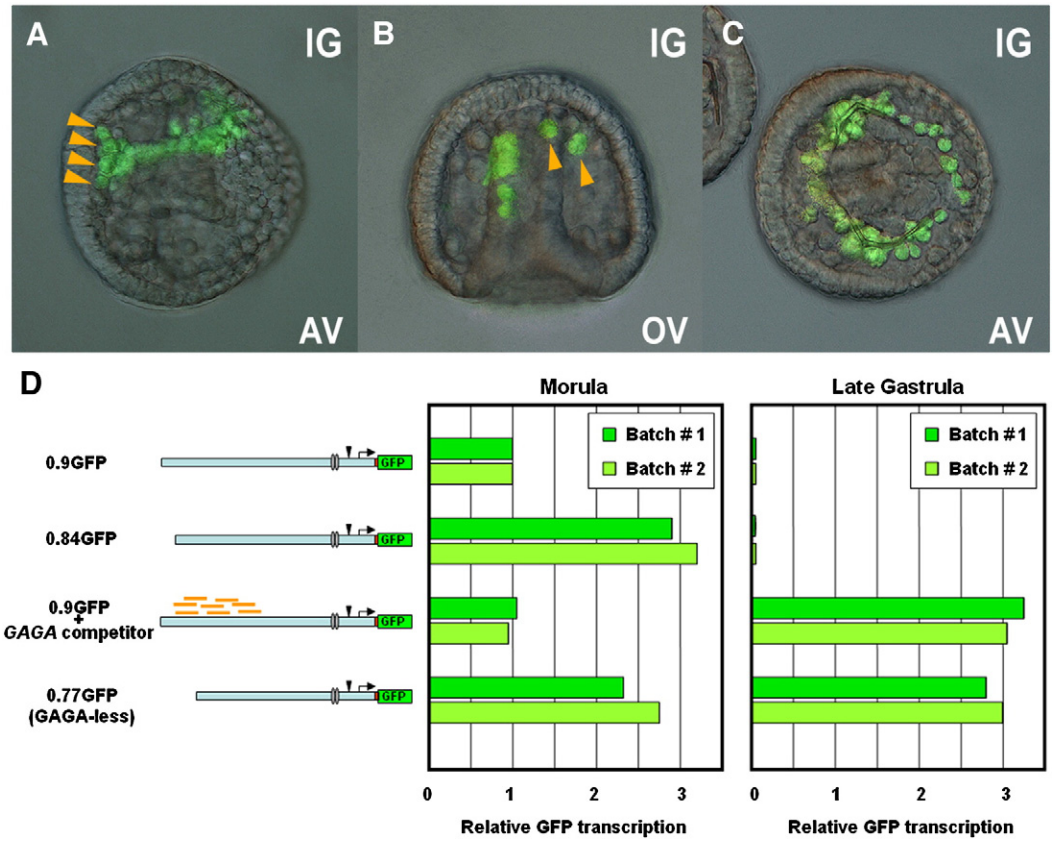

Fig. 9. Variation of the *Hbox12* promoter-GFP *transgene* expression by deletion of negative spatial and temporal sequence elements from the AE module. (A–C) Expression of the 0.84GFP construct in embryos at late gastrula stage. GFP fluorescence image overlays from live *transgene* injected embryos are shown. The upper right corner of each image is labelled with its developmental stage, and the orientation is indicated on the lower right corner. (A) In this embryo GFP expression occurred in endoderm cells located along one side, from the blastopore to the upper third of the invaginating *archenteron*, and also comprehended some SMCs. (orange arrows in panel B); (C) An embryo expressing GFP in PMCs. Abbreviations: lG, late gastrula; AV, view from the animal pole with the oral side to the left; OV, view from the oral ectoderm. (D) Effects of the *in vivo* competition of proteins binding to the *GA* repeats and deletion of the *GA*-rich region on the temporal regulation of *Hbox12* promoter driven *transgene* expression. Q-PCR analysis was carried at morula and late gastrula stages on the *transgene* transcripts from the 0.9GFP, 0.84GFP and 0.77GFP constructs. The 0.84GFP and 0.77GFP constructs derived from the 0.9GFP after the deletion, respectively, of the 5′-most 60 bp and 130 bp sequences. *GAGA* competitor consisted of ligated *GA* polymers. Graphs show *n*-fold changes in mRNA expression level of *GFP*, based on the threshold cycle number (Ct) of the 0.84GFP, the *in vivo GA*-competed 0.9GFP and the 0.77GFP constructs, compared to that of the intact 0.9GFP control. Ct numbers were normalized for the endogenous *MBF-1* in the same sample. Data were derived from two independent microinjection experiments and each bar represents the average of triplicate samples from a given batch of microinjected embryos.

such a region of negative regulatory sequences for down-regulation of *Hbox12* at gastrula stage.

To obtain some insights on the regulatory sequences involved in temporal repression, we concentrated our attention on the purine-rich region containing sixteen *GA* tandem repeats located in the AE module (Fig. 6). As described, a stretch of *GA* sequences, very similar to those found in *Hbox12*, is located upstream the enhancer of the *P. lividusα-H2A* histone gene and we have demonstrated that this sequence element is essential for the down-regulation at gastrula stage (Di Caro et al., 2004). Because the two genes have a very similar temporal expression profile, we hypothesised that the *GA* repeats are needed for the silencing also of *Hbox12*. To address this issue, we performed an *in vivo* competition assay that we previously used to define the function of the *cis*-regulatory sequences responsible for the *α-H2A* histone gene expression (Di Caro et al., 2004). Ligated *GAGA* oligonucleotides were microinjected into fertilized sea urchin eggs together with the 0.9GFP construct at the molar ratio of 50 to 1, and *transgene* expression was determined by Q-PCR. As indicated by the results shown in Fig. 9D, in contrast to the non-competed construct, injection of molar excess of the *GAGA* polymers up-regulated the expression of the *transgene* at gastrula stage.

To enforce this evidence we deleted the sequence region containing the *GA* repeats from the 0.84GFP construct and injected the resulting 0.77GFP *transgene* into zygotes. Such a deletion did not consistently change the spatial distribution of the reporter in the transgenic embryos, with respect to the 0.84GFP construct (Table 2), indicating that no additional negative spatial *cis*-elements are located within the deleted region. Obviously, the relative GFP transcription is 2–3 times higher at morula stage for the 0.77GFP compared to the 0.9GFP construct, because the former lacks the above described 60 bp sequence that provides a negative spatial input at morula stage. Nevertheless, Q-PCR analysis showed that the elimination of the *GA* repeat region up-regulated the expression of the *transgene* at gastrula stage (Fig. 9D). Altogether, these results strongly suggest that the *GA* repeats are the site of binding of negative factor(s) required for down-regulation.

Interestingly, computational analysis conducted on the AE module revealed an almost perfect conserved motif, *CAACTT*, in the minus strand, immediately downstream to the *GA* repeats (Fig. 6). The sequence of this motif well matches the canonical consensus binding site, *YAACG/TG*, for the Myb factor (Luscher and Eisenman, 1990). It as been shown that the Myb protein is expressed in cells of the oral ectoderm and endomesoderm territories of *S. purpuratus* at the prism stage, where it acts as a spatial repressor that define the aboral ectoderm-specific *CyIIIa* gene (Coffman et al., 1997). Because the Myb-*like* site is close to the 5′-end of the 0.77 kb promoter fragment (Fig. 6), which drives very little *transgene* expression in the oral ectoderm (Table 2), we deleted it by a further 5′-truncation and assayed the spatial expression of the resulting 0.76GFP *transgene* at late

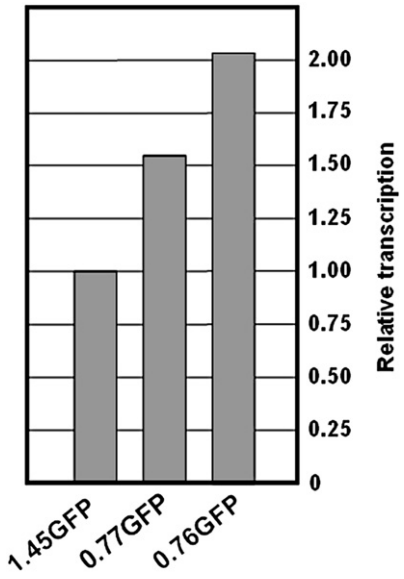

Fig. 10. Deletion of the Myb-*like* consensus from the 5′-end of the 0.77GFP construct causes an additional increase of the *transgene* transcription. Graphs show *n*-fold changes in mRNA expression level of *GFP*, based on the threshold cycle number (Ct) of the mutated constructs, compared to that of the intact 1.45GFP control. Ct numbers were normalized for the endogenous *MBF-1* in the same sample. Data were derived from two independent microinjection experiments and each bar represents the average of triplicate samples from a given batch of microinjected embryos.

developmental stages. Indeed, the percentage of embryos displaying restricted expression of GFP in the aboral territory decreased significantly, from ~79% to ~52%, when this site was eliminated. Conversely, the percentage of embryos expressing GFP in oral ectoderm increased considerably. The overall fraction of embryos expressing the *transgene* in the endomesoderm remained similar to that obtained with the 0.77GFP and 0.84GFP constructs, although, the GFP positive embryos stained in the SMCs did more than double their number (Table 2). Furthermore, the fraction of GFP-expressing embryos was almost 23% higher than that obtained for the 1.45GFP-injected embryos. We reasoned that this could be a cumulative effect due to the sole removal of negative *cis*-regulatory elements from the *AE* module. Definitely, the Myb-less 0.76GFP construct retains both positive inputs given by the Sox and Otx sites.

As expected, Q-PCR measurements showed that removal of the Myb-*like* binding site increased the expression level of the *transgene* to a further ~60% with respect to the 0.77GFP construct (Fig. 10). We conclude that the Myb-*like* binding site is a key negative *cis*-regulatory element required to prevent the ectopic expression of *Hbox12* in the oral ectoderm territory.

Discussion

Hbox12 is a regulator with a homeodomain that, as shown more than a decade ago, represents the first zygotic transcription factor with a highly restricted temporal and spatial expression in the animal hemisphere of the very early sea urchin embryo (Di Bernardo et al., 1994, 1995). This paper concerns the analysis of the *cis*-regulatory region of this gene, aiming at the identification of the upstream regulators. This represents the first step towards the knowledge of its function.

A balance of positive and negative elements is responsible for the aboral ectoderm-specific expression of the Hbox12 gene

Functional *cis*-regulatory dissection of the *Hbox12* gene revealed the AE module, which appeared to be necessary for the specific activity of *Hbox12* promoter in the presumptive aboral ectoderm. Within the AE module, an *in silico* search for putative binding of transcriptional regulators identified a *TAATCT* element, which is a high-affinity binding site for the Otx factor (Hanes and Brent, 1989, 1991; Treisman et al., 1989; Wilson et al., 1993). Although *Otx* is a single copy gene in the sea urchin genome, at least four distinct transcripts exist, due to differential promoter utilization and alternative RNA splicing (Kiyama et al., 1998; Li et al., 1997; Mitsunaga-Nakatsubo et al., 1998). They encode two Otx orthodenticle-related proteins, (α) and (β), identical in the homeodomain, but differing only in their N-terminal sequences (Li et al., 1997). Intriguingly, the Otx(α) protein is a key transcriptional activator during early embryonic development and appears to function selectively in different cell types (Gan et al., 1995; Mao et al., 1996; Oliveri et al., 2002; Wei et al., 1997; Yuh et al., 1998). A role for Otx(α) has been demonstrated in the direct activation of some genes, such as *Spec2a*, whose expression is activated when the aboral ectoderm founder cells arise, and whose mRNAs accumulate exclusively in aboral ectoderm cells (Hardin et al., 1988; Mao et al., 1996; Tomlinson and Klein, 1990; Tomlinson et al., 1990). Most of the Otx(α) molecules reside in the cytoplasm of eggs and early cleavage stage embryos but translocate into the nuclei at 60-cell stage (Li et al., 1997; Mao et al., 1996). The time of Otx(α) nuclear translocation strictly correlates with the peak of *Hbox12* transcription, so we were not surprised to find that the Otx(α) activator is associated with the *Hbox12* promoter at morula stage and constitutes a positive input within the AE module (Fig. 11).

A second positive *cis*-regulatory element is constituted by a perfectly conserved *AACAAT* sequence, probably recognized by a Sox factor. Of particular interest, SoxB proteins are maternally expressed and broadly distributed in the presumptive animal hemisphere of the early sea urchin embryo (Kenny et al., 1999; Kenny et al., 2003). They act as positive regulators and are required for the expression of some ectoderm-specific genes, such as *Nodal* and *Univin* (Range et al., 2007). Unfortunately several attempts to mutagenize the Sox binding sequence failed, so we could not directly prove its involvement in *Hbox12* transcription. However, the effects of the internal deletion of the 29 bp DNA fragment containing the Sox consensus on the spatial expression of the *transgene*, strongly suggest that a Sox factor could

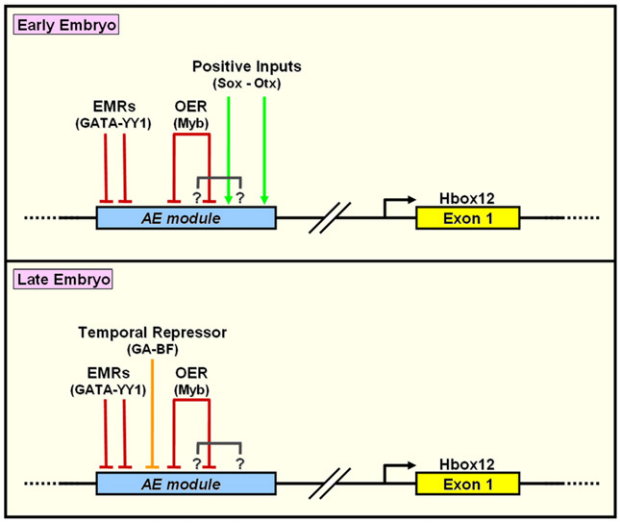

Fig. 11. A model for the spatial and temporal expression of the*Hbox12*gene in the sea urchin embryo. Integration of transcription factor inputs occurs by *cis*-regulatory elements located in the AE module. In the early embryo, *Hbox12* expression is initiated in the presumptive aboral ectoderm by combinatorial positive inputs from Otx and probably Sox. Unidentified repressors prevent *Hbox12* transcription in the endomesoderm, while the Myb repressor probably negatively regulates *Hbox12* expression in the presumptive oral ectoderm. At late cleavage stages, transcription of *Hbox12* is abolished by the loss of the positive inputs coupled to the appearing of a temporal repressor input (GA-binding factor, GA-BF) acting on the *GA*-rich sequence element. Question marks indicate additional inputs of unidentified factors acting on the two predicted *TAAT* homeodomain-consensus adjacent to the Sox site.

participate in conferring the strong boost of transcription that characterizes the temporal activation of *Hbox12*.

The presence of Otx and Sox in nuclei of other than aboral ectoderm cells at the time of *Hbox12* transcription indicates that these factors cannot be solely responsible for the aboral ectoderm-specific expression. For example, it is well known that several genes are activated by Otx at different times and/or in different territories during embryogenesis, including *pmar1* in the micromeres (Oliveri et al., 2002) and *endo16* in the vegetal plate (Yuh et al., 2001, 2004). Consequently, the aboral ectoderm-exclusive expression of *Hbox12* should be guaranteed through interaction of ancillary repressors with negative DNA elements of the AE module that preclude activity in the surrounding territories. We found that this was indeed the case. Deletion analysis indicated a negative action of a 60 bp region located at the 5′-most of the AE module sequence. When this region is missing, *transgene* expression spreads to the vegetal domain, *viz* micromere and veg2 derived cells, implying the existence of a still unidentified repressor(s) functioning in vegetal cells to which the AE module can respond. We noticed that DNA sequence of such a region contains at least two potential consensus binding sites for GATA factors and a site for YY1 (Fig. 11). It should be emphasized that both factors have been characterized to act as strong transcriptional repressors in several contexts (Letting et al., 2004; Shi et al., 1991). In particular, in the sea urchin embryo, GATA-E factor acts as direct territorial repressor of *Spec2a* gene and serves to confine *Spec2a* expression in aboral ectoderm founder cells (Kiyama et al., 2005). In spite of that, the timing of GATA-E appearance (Kiyama and Klein, 2007; Lee and Davidson, 2004) excludes any role of this repressor in *Hbox12* territorial restriction, at least at very early stages of development. Of some interest, the *Drosophila* YY1 (the sea urchin ortholog has still to be identified) has been involved in the recruitment of Polycomb group proteins for achievement and maintaining of transcriptional repression of developmentally important genes (Atchison et al., 2003; Wilkinson et al., 2006).

Mechanisms for spatial and temporal transcriptional repression

There is another broad territory in which *Hbox12* expression is repressed, *viz* the oral ectoderm. It is of great relevance that the *Myb* transcripts are localized in the oral ectoderm and endomesoderm territories where the encoded protein acts as a repressor to restrict the expression of the *CyIIIa* gene in the aboral ectoderm cells of *S. purpuratus* (Coffman et al., 1997). Although nothing is known about the spatial distribution of Myb in the early embryo, a repressive function is carried out even on the *cis*-regulatory apparatus of the *Nodal* gene (Range et al., 2007). The evidence presented in this paper strongly suggest that a Myb or a closely related factor acts as an absolutely necessary negative spatial input to which the AE module can respond (Fig. 11). In fact, when the Myb-*like* binding site is missing, the expression of the *GFP transgene* spreads towards the oral domain, indicating that a Myb-*like* factor plays a prominent role as a transcriptional repressor of the *Hbox12* gene in this territory. In addition, the results summarized in Table 2 suggest a possible contribution of a Myb-*like* factor to prevent the ectopic expression even in the endomesoderm.

Additional repressor binding sites seem also localized more downstream from the putative Myb-binding sequence. An *ATTGAA* sequence element partially overlaps the Sox consensus. It has been demonstrated that such a sequence constitutes a non-canonical binding site for the autoregulation of the *c-Myb* gene expression in human cells (Nicolaides et al., 1991). We speculate that this putative DNA element could be utilized in combination with the canonical *CAACTT* site by the same repressor, probably Myb, to prevent *Hbox12* expression in both oral ectoderm and endomesoderm territories (Fig. 11). Alternatively, distinct repressors could act independently on the two consensus sequences. Moreover, we cannot distinguish between a positive, negative or neutral functional role of the two predicted *TAAT* homeodomain-consensus adjacent to the Sox site (Fig. 11). Further experiments aiming at the identification of the transcriptional repressors would elucidate the mechanism(s) of territorial restriction of the *Hbox12* gene.

A binding site for a putative temporal repressor is located in a region containing the *GA* repeated sequences (Fig. 11). Indeed, deletion of this region, as well as co-injection of excess of the *GAGA* oligonucleotide, impaired the capability of the AE module to down-regulate the *transgene* expression at gastrula stage without affecting its transcription at earlier stages. The same *cis*–regulatory sequence is needed for the silencing of the α-*H2A* histone gene whose temporal expression overlaps with that of *Hbox12* (Di Caro et al., 2004). Of some interest, *GA* sequence repeats are present also in the lower strand of the 5′-flanking region of the *pmar1* genes (not shown). Altogether, these evidence highlight a potential common mechanism of shutting off very early genes during sea urchin development.

In summary, we are fairly confident that we have identified the main regulatory region for *Hbox12*, although we cannot entirely exclude the possibility that minor additional *cis*-regulatory elements located in the −0.66 to −0.4 region may contribute to a refinement of the expression profile.

Hbox12 and gene regulatory network of the embryonic ectoderm

The C-terminal region of the Hbox12 protein contains two almost perfectly repeated 11-amino acid long peptides (FSVDFLSRSSR and MSVDFLSRSSR) that are present in other homeodomain transcriptional repressors (Galliot et al., 1999; Mailhos et al., 1998; Smith and Jaynes, 1996). Although we have little knowledge, the highly restricted expression of Hbox12 in the aboral ectoderm founder cells suggests an involvement in the specification of this territory. This hypothesis is supported by the finding that Otx and Sox are positive inputs for *Hbox12* expression, and at least for one of these, *viz* Otx, we demonstrated that it is a direct input. In agreement with earlier observations (Li et al., 1999), the ectopic expression of a dominant-negative Otx fusion had drastic defects in aboral ectoderm specification. Interestingly, embryos expressing such an obligate repressor appeared radialized and promoter activity of both *transgene* and endogenous *Hbox12* were severely reduced with almost identical kinetics. From these results we speculate that *Hbox12* might occupy a peculiar position within the ectoderm gene regulatory network. Additional experiments are in progress to confirm this possibility.

Acknowledgments

We thank W.H. Klein for the Otx antibody and the En-OtxHD plasmid, and C. Gache for the Hatching-Enzyme-GFP construct. We are particularly grateful to G. Amore for his hospitality to perform microinjections in *S. purpuratus* embryos. We acknowledge the insightful comments of an anonymous reviewer that further improved this paper. This work was supported by grants from the University of Palermo (ex 60%) and MIUR (Programmi di Ricerca Scientifica di Interesse Nazionale).

References

Alessandro, C., Di Simone, P., Buscaino, A., Anello, L., Palla, F., Spinelli, G., 2002. Identification of the enhancer binding protein MBF-1 of the sea urchin modulator alpha-H2A histone gene. Biochem Biophys. Res. Commun. 295, 519–525.

Amore, G., Davidson, E.H., 2006. *cis*-Regulatory control of cyclophilin, a member of the ETS-DRI skeletogenic gene battery in the sea urchin embryo. Dev. Biol. 293, 555–564.

Amore, G., Yavrouian, R.G., Peterson, K.J., Ransick, A., McClay, D.R., Davidson, E.H., 2003. Spdeadringer, a sea urchin embryo gene required separately in skeletogenic and oral ectoderm gene regulatory networks. Dev. Biol. 261, 55–81.

Angerer, L.M., Angerer, R.C., 2000. Animal–vegetal axis patterning mechanisms in the early sea urchin embryo. Dev. Biol. 218, 1–12.

Angerer, L.M., Angerer, R.C., 2003. Patterning the sea urchin embryo: gene regulatory networks, signaling pathways, and cellular interactions. Curr. Top Dev. Biol. 53, 159–198.
Atchison, L., Ghias, A., Wilkinson, F., Bonini, N., Atchison, M.L., 2003. Transcription factor YY1 functions as a PcG protein *in vivo*. Embo J. 22, 1347–1358.
Brandhorst, B.P., Klein, W.H., 2002. Molecular patterning along the sea urchin animal–vegetal axis. Int. Rev. Cytol. 213, 183–232.
Burglin, T.R., 1994. A *Caenorhabditis elegans* prospero homologue defines a novel domain. Trends Biochem. Sci. 19, 70–71.
Cameron, R.A., Britten, R.J., Davidson, E.H., 1993. The embryonic ciliated band of the sea urchin, *Strongylocentrotus purpuratus* derives from both oral and aboral ectoderm. Dev. Biol. 160, 369–376.
Cameron, R.A., Davidson, E.H., 1997. LiCl perturbs ectodermal veg1 lineage allocations in *Strongylocentrotus purpuratus* embryos. Dev. Biol. 187, 236–239.
Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M., Werner, T., 2005. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 21, 2933–2942.
Cavalieri, V., Bernardo, M.D., Spinelli, G., 2007. Regulatory sequences driving expression of the sea urchin Otp homeobox gene in oral ectoderm cells. Gene Expr. Patterns 7, 124–130.
Cavalieri, V., Spinelli, G., Di Bernardo, M., 2003. Impairing Otp homeodomain function in oral ectoderm cells affects skeletogenesis in sea urchin embryos. Dev. Biol. 262, 107–118.
Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., Thompson, J.D., 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. 31, 3497–3500.
Coffman, J.A., Denegre, J.M., 2007. Mitochondria, redox signaling and axis specification in metazoan embryos. Dev. Biol. 308, 266–280.
Coffman, J.A., Kirchhamer, C.V., Harrington, M.G., Davidson, E.H., 1997. SpMyb functions as an intramodular repressor to regulate spatial expression of CyIIIa in sea urchin embryos. Development 124, 4717–4727.
Coffman, J.A., McCarthy, J.J., Dickey-Sims, C., Robertson, A.J., 2004. Oral-aboral axis specification in the sea urchin embryo II. Mitochondrial distribution and redox state contribute to establishing polarity in *Strongylocentrotus purpuratus*. Dev. Biol. 273, 160–171.
Davidson, E.H., Cameron, R.A., Ransick, A., 1998. Specification of cell fate in the sea urchin embryo: summary and some proposed mechanisms. Development 125, 3269–3290.
Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim, O., Brown, C.T., Livi, C.B., Lee, P.Y., Revilla, R., Rust, A.G., Pan, Z., Schilstra, M.J., Clarke, P.J., Arnone, M.I., Rowen, L., Cameron, R.A., McClay, D.R., Hood, L., Bolouri, H., 2002a. A genomic regulatory network for development. Science 295, 1669–1678.
Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim, O., Brown, C.T., Livi, C.B., Lee, P.Y., Revilla, R., Schilstra, M.J., Clarke, P.J., Rust, A.G., Pan, Z., Arnone, M.I., Rowen, L., Cameron, R.A., McClay, D.R., Hood, L., Bolouri, H., 2002b. A provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo. Dev. Biol. 246, 162–190.
Di Bernardo, M., Russo, R., Oliveri, P., Melfi, R., Spinelli, G., 1994. Expression of homeobox-containing genes in the sea urchin (*Parancentrotus lividus*) embryo. Genetica 94, 141–150.
Di Bernardo, M., Russo, R., Oliveri, P., Melfi, R., Spinelli, G., 1995. Homeobox-containing gene transiently expressed in a spatially restricted pattern in the early sea urchin embryo. Proc. Natl. Acad Sci. U. S. A. 92, 8180–8184.
Di Caro, D., Melfi, R., Alessandro, C., Serio, G., Di Caro, V., Cavalieri, V., Palla, F., Spinelli, G., 2004. Down-regulation of early sea urchin histone H2A gene relies on *cis* regulative sequences located in the 5¢ and 3¢ regions and including the enhancer blocker sns. J. Mol. Biol. 342, 1367–1377.
Di Caro, V., Cavalieri, V., Melfi, R., Spinelli, G., 2007. Constitutive promoter occupancy by the MBF-1 activator and chromatin modification of the developmental regulated sea urchin alpha-H2A histone gene. J. Mol. Biol. 365, 1285–1297.
Duboc, V., Lepage, T., 2008. A conserved role for the nodal signaling pathway in the establishment of dorso-ventral and left–right axes in deuterostomes. J. Exp. Zoolog B. Mol. Dev. Evol. 310, 41–53.
Duboc, V., Rottinger, E., Besnardeau, L., Lepage, T., 2004. Nodal and BMP2/4 signaling organizes the oral–aboral axis of the sea urchin embryo. Dev. Cell 6, 397–410.
Duboc, V., Rottinger, E., Lapraz, F., Besnardeau, L., Lepage, T., 2005. Left–right asymmetry in the sea urchin embryo is regulated by nodal signaling on the right side. Dev. Cell 9, 147–158.
Ettensohn, C.A., Illies, M.R., Oliveri, P., De Jong, D.L., 2003. Alx1, a member of the Cart1/Alx3/Alx4 subfamily of paired-class homeodomain proteins, is an essential component of the gene network controlling skeletogenic fate specification in the sea urchin embryo. Development 130, 2917–2928.
Flytzanis, C.N., McMahon, A.P., Hough-Evans, B.R., Katula, K.S., Britten, R.J., Davidson, E.H., 1985. Persistence and integration of cloned DNA in postembryonic sea urchins. Dev. Biol. 108, 431–442.
Franks, R.R., Hough-Evans, B.R., Britten, R.J., Davidson, E.H., 1988. Direct introduction of cloned DNA into the sea urchin zygote nucleus, and fate of injected DNA. Development 102, 287–299.
Galliot, B., de Vargas, C., Miller, D., 1999. Evolution of homeobox genes: Q50 paired-like genes founded the paired class. Dev. Genes Evol. 209, 186–197.
Gan, L., Mao, C.A., Wikramanayake, A., Angerer, L.M., Angerer, R.C., Klein, W.H., 1995. An orthodenticle-related protein from *Strongylocentrotus purpuratus*. Dev. Biol. 167, 517–528.
Goldstein, B., Freeman, G., 1997. Axis specification in animal development. Bioessays 19, 105–116.
Hanes, S.D., Brent, R., 1989. DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. Cell 57, 1275–1283.
Hanes, S.D., Brent, R., 1991. A genetic model for interaction of the homeodomain recognition helix with DNA. Science 251, 426–430.
Hardin, P.E., Angerer, L.M., Hardin, S.H., Angerer, R.C., Klein, W.H., 1988. Spec2 genes of *Strongylocentrotus purpuratus*. Structure and differential expression in embryonic aboral ectoderm cells. J. Mol. Biol. 202, 417–431.
Hough-Evans, B.R., Britten, R.J., Davidson, E.H., 1988. Mosaic incorporation and regulated expression of an exogenous gene in the sea urchin embryo. Dev. Biol. 129, 198–208.
Howard-Ashby, M., Materna, S.C., Brown, C.T., Chen, L., Cameron, R.A., Davidson, E.H., 2006a. Gene families encoding transcription factors expressed in early development of *Strongylocentrotus purpuratus*. Dev. Biol. 300, 90–107.
Howard-Ashby, M., Materna, S.C., Brown, C.T., Chen, L., Cameron, R.A., Davidson, E.H., 2006b. Identification and characterization of homeobox transcription factor genes in *Strongylocentrotus purpuratus*, and their expression in embryonic development. Dev. Biol. 300, 74–89.
Kenny, A.P., Kozlowski, D., Oleksyn, D.W., Angerer, L.M., Angerer, R.C., 1999. SpSoxB1, a maternally encoded transcription factor asymmetrically distributed among early sea urchin blastomeres. Development 126, 5473–5483.
Kenny, A.P., Oleksyn, D.W., Newman, L.A., Angerer, R.C., Angerer, L.M., 2003. Tight regulation of SpSoxB factors is required for patterning and morphogenesis in sea urchin embryos. Dev. Biol. 261, 412–425.
Kitamura, K., Nishimura, Y., Kubotera, N., Higuchi, Y., Yamaguchi, M., 2002. Transient activation of the micro1 homeobox gene family in the sea urchin (*Hemicentrotus pulcherrimus*) micromere. Dev. Genes Evol. 212, 1–10.
Kiyama, T., Akasaka, K., Takata, K., Mitsunaga-Nakatsubo, K., Sakamoto, N., Shimada, H., 1998. Structure and function of a sea urchin orthodenticle-related gene (HpOtx). Dev. Biol. 193, 139–145.
Kiyama, T., Klein, W.H., 2007. SpGataE, a *Strongylocentrotus purpuratus* ortholog of mammalian Gata4/5/6: protein expression, interaction with putative target gene spec2a, and identification of friend of Gata factor SpFog1. Dev. Genes Evol. 217, 651–663.
Kiyama, T., Zhang, N., Dayal, S., Yun Lee, P., Liang, S., Villinski, J.T., Klein, W.H., 2005. *Strongylocentrotus purpuratus* transcription factor GATA-E binds to and represses transcription at an Otx-Goosecoid *cis*-regulatory element within the aboral ectoderm-specific spec2a enhancer. Dev. Biol. 280, 436–447.
Lee, P.Y., Davidson, E.H., 2004. Expression of Spgatae, the *Strongylocentrotus purpuratus* ortholog of vertebrate GATA4/5/6 factors. Gene Expr. Patterns 5, 161–165.
Lee, P.Y., Nam, J., Davidson, E.H., 2007. Exclusive developmental functions of gatae *cis*-regulatory modules in the *Strongylocentrorus purpuratus* embryo. Dev. Biol. 307, 434–445.
Lepage, T., Sardet, C., Gache, C., 1992. Spatial expression of the hatching enzyme gene in the sea urchin embryo. Dev. Biol. 150, 23–32.
Letting, D.L., Chen, Y.Y., Rakowski, C., Reedy, S., Blobel, G.A., 2004. Context-dependent regulation of GATA-1 by friend of GATA-1. Proc. Natl. Acad. Sci. U. S. A. 101, 476–481.
Li, X., Chuang, C.K., Mao, C.A., Angerer, L.M., Klein, W.H., 1997. Two Otx proteins generated from multiple transcripts of a single gene in *Strongylocentrotus purpuratus*. Dev. Biol. 187, 253–266.
Li, X., Wikramanayake, A.H., Klein, W.H., 1999. Requirement of SpOtx in cell fate decisions in the sea urchin embryo and possible role as a mediator of beta-catenin signaling. Dev. Biol. 212, 425–439.
Logan, C.Y., McClay, D.R., 1997. The allocation of early blastomeres to the ectoderm and endoderm is variable in the sea urchin embryo. Development 124, 2213–2223.
Logan, C.Y., Miller, J.R., Ferkowicz, M.J., McClay, D.R., 1999. Nuclear beta-catenin is required to specify vegetal cell fates in the sea urchin embryo. Development 126, 345–357.
Luscher, B., Eisenman, R.N., 1990. New light on Myc and Myb. Part II. Myb. Genes Dev. 4, 2235–2241.
Mailhos, C., Andre, S., Mollereau, B., Goriely, A., Hemmati-Brivanlou, A., Desplan, C., 1998. *Drosophila* Goosecoid requires a conserved heptapeptide for repression of paired-class homeoprotein activators. Development 125, 937–947.
Mao, C.A., Wikramanayake, A.H., Gan, L., Chuang, C.K., Summers, R.G., Klein, W.H., 1996. Altering cell fates in sea urchin embryos by overexpressing SpOtx, an orthodenticle-related protein. Development 122, 1489–1498.
Mitsunaga-Nakatsubo, K., Akasaka, K., Sakamoto, N., Takata, K., Matsumura, Y., Kitajima, T., Kusunoki, S., Shimada, H., 1998. Differential expression of sea urchin Otx isoform (hpOtxE and HpOtxL) mRNAs during early development. Int. J. Dev. Biol. 42, 645–651.
Nam, J., Su, Y.H., Lee, P.Y., Robertson, A.J., Coffman, J.A., Davidson, E.H., 2007. *Cis*-regulatory control of the nodal gene, initiator of the sea urchin oral ectoderm gene network. Dev. Biol. 306, 860–869.
Nicolaides, N.C., Gualdi, R., Casadevall, C., Manzella, L., Calabretta, B., 1991. Positive autoregulation of c-myb expression via Myb binding sites in the 5′ flanking region of the human c-myb gene. Mol. Cell Biol. 11, 6166–6176.
Nishimura, Y., Sato, T., Morita, Y., Yamazaki, A., Akasaka, K., Yamaguchi, M., 2004. Structure, regulation, and function of micro1 in the sea urchin *Hemicentrotus pulcherrimus*. Dev. Genes Evol. 214, 525–536.
Oliveri, P., Carrick, D.M., Davidson, E.H., 2002. A regulatory gene network that directs micromere specification in the sea urchin embryo. Dev. Biol. 246, 209–228.
Oliveri, P., Davidson, E.H., 2004a. Gene regulatory network analysis in sea urchin embryos. Methods Cell. Biol. 74, 775–794.
Oliveri, P., Davidson, E.H., 2004b. Gene regulatory network controlling embryonic specification in the sea urchin. Curr. Opin. Genet. Dev. 14, 351–360.

Quandt, K., Frech, K., Karas, H., Wingender, E., Werner, T., 1995. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. 23, 4878–4884.

Range, R., Lapraz, F., Quirin, M., Marro, S., Besnardeau, L., Lepage, T., 2007. *Cis*-regulatory analysis of nodal and maternal control of dorsal–ventral axis formation by Univin, a TGF-{beta} related to Vg1. Development 134, 3649–3664.

Ransick, A., Davidson, E.H., 2006. *cis*-Regulatory processing of Notch signaling input to the sea urchin glial cells missing gene during mesoderm specification. Dev. Biol. 297, 587–602.

Sherwood, D.R., McClay, D.R., 2001. LvNotch signaling plays a dual role in regulating the position of the ectoderm-endoderm boundary in the sea urchin embryo. Development 128, 2221–2232.

Shi, Y., Seto, E., Chang, L.S., Shenk, T., 1991. Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein. Cell 67, 377–388.

Smith, S.T., Jaynes, J.B., 1996. A conserved region of engrailed, shared among all en-, gsc-, Nk1-, Nk2- and msh-class homeoproteins, mediates active transcriptional repression *in vivo*. Development 122, 3141–3150.

Sodergren, E., Weinstock, G.M., Davidson, E.H., Cameron, R.A., Gibbs, R.A., Angerer, R.C., Angerer, L.M., Arnone, M.I., Burgess, D.R., Burke, R.D., Coffman, J.A., Dean, M., Elphick, M.R., Ettensohn, C.A., Foltz, K.R., Hamdoun, A., Hynes, R.O., Klein, W.H., Marzluff, W., McClay, D.R., Morris, R.L., Mushegian, A., Rast, J.P., Smith, L.C., Thorndyke, M.C., Vacquier, V.D., Wessel, G.M., Wray, G., Zhang, L., Elsik, C.G., Ermolaeva, O., Hlavina, W., Hofmann, G., Kitts, P., Landrum, M.J., Mackey, A.J., Maglott, D., Panopoulou, G., Poustka, A.J., Pruitt, K., Sapojnikov, V., Song, X., Souvorov, A., Solovyev, V., Wei, Z., Whittaker, C. A., Worley, K., Durbin, K.J., Shen, Y., Fedrigo, O., Garfield, D., Haygood, R., Primus, A., Satija, R., Severson, T., Gonzalez-Garay, M.L., Jackson, A.R., Milosavljevic, A., Tong, M., Killian, C.E., Livingston, B.T., Wilt, F.H., Adams, N., Belle, R., Carbonneau, S., Cheung, R., Cormier, P., Cosson, B., Croce, J., Fernandez-Guerra, A., Geneviere, A.M., Goel, M., Kelkar, H., Morales, J., Mulner-Lorillon, O., Robertson, A.J., Goldstone, J.V., Cole, B., Epel, D., Gold, B., Hahn, M.E., Howard-Ashby, M., Scally, M., Stegeman, J.J., Allgood, E.L., Cool, J., Judkins, K.M., McCafferty, S.S., Musante, A.M., Obar, R.A., Rawson, A.P., Rossetti, B.J., Gibbons, I.R., Hoffman, M.P., Leone, A., Istrail, S., Materna, S.C., Samanta, M.P., Stolc, V., Tongprasit, W., et al., 2006. The genome of the sea urchin *Strongylocentrotus purpuratus*. Science 314, 941–952.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic. Acids Res. 22, 4673–4680.

Tomlinson, C.R., Klein, W.H., 1990. Temporal and spatial transcriptional regulation of the aboral ectoderm-specific Spec genes during sea urchin embryogenesis. Mol. Reprod. Dev. 25, 328–338.

Tomlinson, C.R., Kozlowski, M.T., Klein, W.H., 1990. Ectoderm nuclei from sea urchin embryos contain a Spec-DNA binding protein similar to the vertebrate transcription factor USF. Development 110, 259–272.

Treisman, J., Gonczy, P., Vashishtha, M., Harris, E., Desplan, C., 1989. A single amino acid can determine the DNA binding specificity of homeodomain proteins. Cell 59, 553–562.

Tu, Q., Brown, C.T., Davidson, E.H., Oliveri, P., 2006. Sea urchin Forkhead gene family: phylogeny and embryonic expression. Dev. Biol. 300, 49–62.

van Beest, M., Dooijes, D., van De Wetering, M., Kjaerulff, S., Bonvin, A., Nielsen, O., Clevers, H., 2000. Sequence-specific high mobility group box factors recognize 10–12-base pair minor groove motifs. J. Biol. Chem. 275, 27266–27273.

Wei, Z., Angerer, L.M., Angerer, R.C., 1997. Multiple positive *cis* elements regulate the asymmetric expression of the SpHE gene along the sea urchin embryo animal–vegetal axis. Dev. Biol. 187, 71–78.

Weitzel, H.E., Illies, M.R., Byrum, C.A., Xu, R., Wikramanayake, A.H., Ettensohn, C.A., 2004. Differential stability of beta-catenin along the animal–vegetal axis of the sea urchin embryo mediated by dishevelled. Development 131, 2947–2956.

Wikramanayake, A.H., Huang, L., Klein, W.H., 1998. beta-Catenin is essential for patterning the maternally specified animal–vegetal axis in the sea urchin embryo. Proc. Natl. Acad. Sci. U. S. A. 95, 9343–9348.

Wilkinson, F.H., Park, K., Atchison, M.L., 2006. Polycomb recruitment to DNA *in vivo* by the YY1 REPO domain. Proc. Natl. Acad. Sci. U. S. A. 103, 19296–19301.

Wilson, D., Sheng, G., Lecuit, T., Dostatni, N., Desplan, C., 1993. Cooperative dimerization of paired class homeo domains on DNA. Genes Dev. 7, 2120–2134.

Wu, S.Y., McClay, D.R., 2007. The Snail repressor is required for PMC ingression in the sea urchin embryo. Development 134, 1061–1070.

Yuh, C.H., Bolouri, H., Davidson, E.H., 1998. Genomic *cis*-regulatory logic: experimental and computational analysis of a sea urchin gene. Science 279, 1896–1902.

Yuh, C.H., Bolouri, H., Davidson, E.H., 2001. *cis*-Regulatory logic in the endo16 gene: switching from a specification to a differentiation mode of control. Development 128, 617–629.

Yuh, C.H., Dorman, E.R., Howard, M.L., Davidson, E.H., 2004. An otx *cis*-regulatory module: a key node in the sea urchin endomesoderm gene regulatory network. Dev. Biol. 269, 536–551.