MOLECULAR EMBRYOLOGY OF A LARVACEAN UROCHORDATE, *OIKOPLEURA DIOICA*, AND THE ORIGIN OF CHORDATE INNOVATIONS

by

SUSAN LEE BASSHAM

A DISSERTATION

Presented to the Department of Biology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2002

"Molecular embryology of a larvacean urochordate, *Oikopleura dioica*, and the origin of chordate innovations," a dissertation prepared by Susan Lee Bassham in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Biology. This dissertation has been approved and accepted by:

_____

Dr. Nora Terwilliger, Chair of the Examining Committee

_____

Date

Committee in charge:

Dr. Nora Terwilliger, Chair
Dr. John Postlethwait
Dr. Charles Kimmel
Dr. Richard Emlet
Dr. William Orr

Accepted by:

_____

Dean of the Graduate School

An Abstract of the Dissertation of

Susan Lee Bassham for the degree of Doctor of Philosophy

in the Department of Biology to be taken June 2002

Title: MOLECULAR EMBRYOLOGY OF A LARVACEAN UROCHORDATE, *OIKOPLEURA DIOICA*, AND THE ORIGIN OF CHORDATE INNOVATIONS

Approved: _____
Dr. John H. Postlethwait

Phyla are recognized by unifying characters, some of which are unique evolutionary innovations, but how do new structures and body plans evolve? Defining characters of the phylum Chordata include the notochord, dorsal hollow nerve cord, and post-anal tail. Despite these commonalities, the chordate subphyla Vertebrata, Cephalochordata, and Urochordata span a broad range of morphological diversity. Vertebrates are distinguished from the other chordates by an elaborated brain, paired sense organs, and a skeleton. Two embryonic tissues, neural crest and placodes, are essential for these features. It became widely accepted that crest and placodes are vertebrate innovations because no clear homologues of these tissues or their derivatives had been observed in other chordates. Further speculation suggested that acquisition of neural crest and placodes permitted or propelled the evolution of raptorial vertebrates from a sessile, filter-feeding ancestor. By comparing ontogenies of extant chordates and

their outgroups, we hope to reconstruct the timing and mechanism of the origin of chordate novelties.

Within the most basal chordate lineage, Urochordata, larvaceans uniquely allow us to analyze development of adult characters in a chordate body plan that does not become reconfigured by sweeping metamorphic events. This dissertation examines the evolutionary history of two chordate innovations, notochord and placodes. I present DNA sequences and expression patterns for larvacean homologues of several vertebrate genes important in notochord and placode development: *T (Brachyury)*, *Pitx*, *Eya*, and members of the *Pax* and *Six* gene families. Larvacean *T* is expressed in notochord, but also in posterior endoderm. Because deuterostome and protostome outgroups display a similar pattern, chordate *T* may have a conserved role in endoderm patterning more ancient than its role in the notochord. I also show that several genes important for the development of a subset of vertebrate placodes are expressed in the developing ciliary funnel and ventral organ in larvaceans. Ultrastructure and topography had previously suggested that these organs are candidate homologues of pituitary and olfactory placodes. My results are consistent with an origin of these placodes in the common ancestor of modern chordates, predating the evolution of vertebrates.

This dissertation includes both my previously published and my co-authored materials.

CURRICULUM VITAE

NAME OF AUTHOR: Susan Lee Bassham

PLACE OF BIRTH: Honolulu, Hawaii

DATE OF BIRTH: June 17, 1969

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon
Reed College

DEGREES AWARDED:

Doctor of Philosophy in Biology, 2002, University of Oregon
Bachelor of Arts in Biology, 1992, Reed College

AREAS OF SPECIAL INTEREST:

Evolution of Development

PROFESSIONAL EXPERIENCE:

Graduate Teaching Fellow, Department of Biology, University of Oregon, Eugene, 1994-95

Research Assistant, Department of Biology, Reed College, Portland, 1992-94

AWARDS AND HONORS:

Appointed to the "Integrated Training in Evolution of Development" training grant (NSF DGE-9972830), University of Oregon, 1999-2002

"Best Poster" , 31st Annual N.W. Regional Developmental Biology Conference, Washington, 1998

Appointed to the Developmental Biology Training Program training grant (NIH 5 T32 HD07348), University of Oregon, 1995-98

PUBLICATIONS:

Bassham, S., and Postlethwait, J. (2000). *Brachyury* (*T*) expression in embryos of a larvacean urochordate, *Oikopleura dioica*, and the ancestral role of *T*. *Dev. Biol.* 220, 322-332.

Bassham, S., Beam, A., and Shampay, J. (1998). Telomere variation in *Xenopus laevis*. *Mol. Cell Biol.* 18, 269-275.

Shampay, J., Schmitt, M., and Bassham, S. (1995). A novel minisatellite at a cloned hamster telomere. *Chromosoma* 104, 29-38.

ACKNOWLEDGEMENTS

I am indebted to my advisor, Dr. John Postlethwait for his steadfast support and encouragement, and for his valuable contributions to the development of these projects. I also thank Dr. Angel Amores, Dr. William Cresko, Dr. Allan Force, and Dr. Yi-Lin Yan for their helpful contributions, and Burley Young, skipper of the Charming Polly, for his generous assistance with collecting larvaceans. This research was funded in part by an NIH Developmental Biology Training Program grant (5 T32 HD07348), by an NSF IGERT grant in Evolution of Development (DGE-9972830), and by other NIH grants (R01 RR10715, P01 HD22486).

For my parents, James A. Bassham and Leslie A. Bassham

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

The divergence of the three chordate subphyla is ancient, perhaps not much younger than the radiation of the bilaterians. Fossils of cephalochordate, vertebrate, and urochordate-like animals occur in Lower Cambrian strata (Shu *et al.*, 1996, Shu *et al.*, 2000; Shu *et al.*, 2001). If these taxonomic assignments are correct, not only the chordates, but also the distinct body plans of the three chordate subphyla have been in existence for more than 530 million years. Because chordate fossils are rare and because the oldest known fossils do not appear to bridge the gulfs between modern chordate body plans, the study of extant chordates is an essential thread to understanding the nature of the chordate common ancestor and the origin of the vertebrates. Traits shared by all chordate subphyla are likely to have been present in their common ancestor. Near the base of the chordate tree, Urochordata diverged from the lineage that would, not long after, give rise to the cephalochordates and vertebrates, and within the urochordates, the larvaceans may represent the most deeply branching lineage (Wada, 1998). Garstang's famous insight, that ontogeny does not recapitulate phylogeny, but rather creates it, expresses the importance of comparing embryonic over adult characters in reconstructing the evolutionary history of morphologies. This dissertation investigates the evolution of chordate innovations through an analysis of larvacean embryogenesis.

The morphology of most adult urochordates is not obviously chordate-like. Ascidians, for example, lose most of the definitive chordate traits of their motile larvae at metamorphosis, and become sessile filter feeders. It wasn't until 1867 that ascidians were recognized to be chordates rather than molluscs. The discovery of the chordate nature of ascidians then extended to the other urochordates, including larvaceans, which had only 16 years earlier been recognized as adult urochordates (reviewed in Bone, 1998). Unlike their ascidian relatives, larvaceans retain a chordate morphology into adulthood; their tadpole-like body has a notochord, a dorsal hollow nerve cord, and a post-anal tail. Larvaceans are planktonic filter-feeders that secrete a complex, extracellular "house", a particle filter for trapping unicells (Flood and Deibel, 1998). The development of one larvacean, *O. dioica*, has been studied in some detail (Delsman, 1910, 1912; Galt, 1972; Fenaux, 1976). In this species, development from zygote to juvenile is rapid, lasting only 15 hours at 12°C, and a full generation is accomplished in under two weeks. These characteristics, in addition to optical clarity of the animals at all stages, year-round reproduction, and relative hardiness to laboratory culture, has made *O. dioica* a good system in which to study larvacean embryogenesis. In comparison to their ascidian relatives, however, larvaceans in general are delicate and this has discouraged much analysis of their embryology.

In this study, I have investigated the evolutionary history of two chordate innovations, the notochord and the placodes. The notochord, a rod of tissue that stiffens the axis of non-vertebrate chordates and is an axial organizing tissue in vertebrate embryos, is one of the unifying characters of the chordates. The placodes, embryonic

structures that contribute to a variety of paired sensory structures and to the anterior pituitary, are widely held to be an innovation of the vertebrates.

Though the notochord is the organ for which Chordata is named, and though its function is essential to proper development of the vertebrate axis, its evolutionary origin is not understood. Chapter II describes the expression of *T* or *Brachyury*, a gene important for notochord development in all chordates. *T* stands for tailless or tail, which is kinked or absent in mice mutated at this locus. In ascidian embryonic development, *T* is reportedly expressed only in the cells that will make the notochord (Yasuo and Satoh, 1994), while in vertebrates (Wilkinson *et al.*, 1990; Schulte-Merker *et al.*, 1992) and cephalochordates (Holland *et al.*, 1995) *T* is expressed in additional cells involuting through the blastopore during gastrulation. The restriction of *T* expression to the notochordal lineage in ascidians could represent the ancestral state for the *T* gene in the chordate common ancestor, or the ascidian pattern might represent a loss in the urochordate lineage of broader ancestral *T* expression in involuting mesendoderm. To weigh these alternative hypotheses, I analyzed the expression of *T* in larvacean embryos and found that it is expressed not only in notochord, but also in posterior endoderm, suggesting that the ancestral role of this notochord gene also included a capacity in posterior endoderm patterning.

The influential "new head" hypothesis (Gans and Northcutt, 1983; Northcutt and Gans, 1983) postulates that the appearance of placodes and neural crest in an ancient vertebrate propelled the evolution and diversification of predatory vertebrates from filter-feeding ancestors. The introduction of these ideas in 1983 largely predated information

on the molecular mechanisms of crest and placode development, and little beyond classical observation was known about the embryology of non-vertebrate chordates. The lack of intermediates among living vertebrates, which all appear to have inherited a full complement of crest- and placode-derived tissues, and the apparent absence of homologous tissues in non-vertebrate chordates led to the conclusion that crest and placodes were invented together in an early vertebrate. New data from the study of ascidian and cephalochordate embryos is challenging this hypothesis, and there is increasing discussion of whether some types of placodes could have predated the origin of the vertebrates (see reviews by: Baker and Fraser, 1997; Graham and Begbie, 2000; Shimeld and Holland, 2000).

In experiments reported in Chapter III of this dissertation, I examined the possibility that placode homologues exist in larvacean development. On the basis of their location and ultrastructure, two larvacean organs had been proposed as homologues to the vertebrate olfactory epithelium and anterior pituitary, both placode derivatives. The ventral organ, with its potentially chemosensitive ciliary receptors and its paired innervation to the forebrain resembles the vertebrate olfactory system (Bollner *et al.*, 1986), and the ciliary funnel which connects the brain to the pharynx is like Rathke's pouch, the anlage of the vertebrate adenohypophysis or anterior pituitary (Holmberg, 1982). I analyzed the expression of larvacean *Eya*, *Six*, *Pax*, and *Pitx* genes whose orthologues in vertebrates are expressed in the olfactory and adenohypophyseal placodes, and found that several of these genes are expressed in the developing ventral organ and ciliary funnel.

In ascidian embryos, the primordia of the atrial siphon were suggested to be otic placode homologues on the strength of their expression of ascidian *Pax2/5/8* whose orthologues mark the vertebrate otic placodes. (Wada *et al.*, 1998.) The atrium is the excurrent opening of the ascidian branchial chamber. I analyzed the expression of larvacean *Pax2/5/8* and found it to be expressed in the developing spiracles or gills. If the larvacean and ascidian expression patterns are homologous, the expression of *Pax2/5/8* in the ascidian atrial primordia could reflect an ancient role for this gene in patterning the gill, rather than an indication of urochordate otic placode homologues.

This dissertation is concerned with the origin of evolutionary innovations. In addressing questions of origin, it is important to determine when an apparent novelty first appeared, but also to understand how the novel character develops in the various organisms that share it. The compared ontogenies of chordates tell us how their common ancestor might have developed, and the embryos of non-chordates help us infer how genes now important for the development of chordate novelties were anciently employed. In Chapter II, outgroup organisms that lack a notochord provide important information about the ancestry of a gene essential to notochord development, and larvacean embryos appear to provide a link between the outgroup and chordate patterns.

The following chapter contains material that was previously published (Bassham and Postlethwait, 2000). I conceived and carried out the described experiments with the support of my advisor, John Postlethwait.

CHAPTER II

BRACHYURY (*T*) EXPRESSION IN EMBRYOS OF A LARVACEAN UROCHORDATE, *OIKOPLEURA DIOICA*, AND THE ANCESTRAL ROLE OF *T*

The *Brachyury,* or *T,* gene is required for notochord development in animals occupying all three chordate subphyla, and probably also had this role in the last common ancestor of the chordate lineages. In two chordate subphyla (vertebrates and cephalochordates), *T* is also expressed during gastrulation in involuting endodermal and mesodermal cells, and in vertebrates at least, this expression domain is required for proper development. In the basally diverging chordate subphylum Urochordata, animals in the class Ascidiacea do not employ *T* during gastrulation in endodermal or non-axial mesodermal cells, and it has been suggested that non-notochordal roles for *T* were acquired in the cephalochordate-vertebrate lineage after it split with Urochordata. To test this hypothesis, we cloned *T* from *Oikopleura dioica,* a member of the urochordate class Appendicularia (or Larvacea), which diverged basally in the subphylum. Investigation of the expression pattern in developing *Oikopleura* embryos showed early expression in presumptive notochord precursor cells, in the notochord, and in parts of the developing gut and cells of the endodermal strand. We conclude that the ancestral role of *T* likely included

expression in the developing gut, and became necessary in chordates for construction of the notochord.

INTRODUCTION

The phylum Chordata gets its name from the notochord, but the evolutionary origin of this mesodermal rod is as yet unclear (Peterson *et al.*, 1999; Harada *et al.*, 1995). The development of the notochord in vertebrate embryos requires the action of the *Brachyury* gene (or *T*), as mutations in this gene in mouse (Herrmann *et al.*, 1990) and in zebrafish (Schulte-Merker *et al.*, 1994) block or reduce the development of the notochord. Besides its role in notochord development, in vertebrate embryos, *T* is also required for correct internalization and migration of posterior mesoderm during gastrulation (Wilson *et al.*, 1995). Vertebrate embryos express *T* transiently around the blastopore in presumptive endoderm and mesoderm cells as they involute and migrate, and more persistently in the notochord (e.g., Wilkinson *et al.*, 1990; Schulte-Merker *et al.*, 1992). Similarly, in cephalochordates, expression of at least one of two genes homologous to *T* (*Ambra-1* or -2) occurs in a ring around the blastopore during gastrulation, in the presomitic mesoderm, and in the notochord (Holland *et al.*, 1995). In ascidian urochordates, the role of *T* in notochord development is conserved (Yasuo and Satoh, 1998). In contrast, though, expression of the ascidian *T* orthologue was found *only* in cells of notochord lineage, and not broadly in involuting endoderm and mesoderm or in mesoderm lateral to the presumptive notochord (Yasuo and Satoh, 1993, 1994; Corbo

et al., 1997). What was the original role of *T* in chordate development, and how was it recruited to its current roles in vertebrate embryos?

At least two hypotheses have been suggested to explain the difference in *T* expression between representatives of the cephalochordate-vertebrate clade and ascidians in its urochordate sister group. In one model, the primary ancestral function of the *T* gene in the chordates was in notochord specification (Yasuo and Satoh, 1994; Holland *et al.*, 1995). This hypothesis implies that a function for *T* in gastrulation and paraxial mesoderm formation was acquired, presumably by the evolution of positive regulatory elements driving *T* expression, in the cephalochordate-vertebrate clade after its divergence from the urochordates. Alternatively, both functions of *T* (mesendoderm formation and notochord specification) may have been present in the ancestral chordate, and ascidians retained only the notochord function. Consistent with this proposal is the identification of cis-acting regulatory regions that suppress the activation of *T* in non-notochordal cell lineages, specifically in mesenchyme and muscle of ascidian embryos (Corbo *et al.*, 1997); this may indicate an evolutionary gain of function in the ascidian lineage which eliminates *T* expression in those tissues in ascidians compared to an ancestral state.

Because only one of the four classes of the subphylum Urochordata has been investigated with respect to *T* gene function, it is difficult to generalize about the role of *T* in early chordates. What is needed is an investigation of *T* in a basally diverging urochordate, namely a member of the urochordate class Appendicularia, also called larvaceans (Fig. 1). Appendicularia are a sister class of the clade including the sessile,

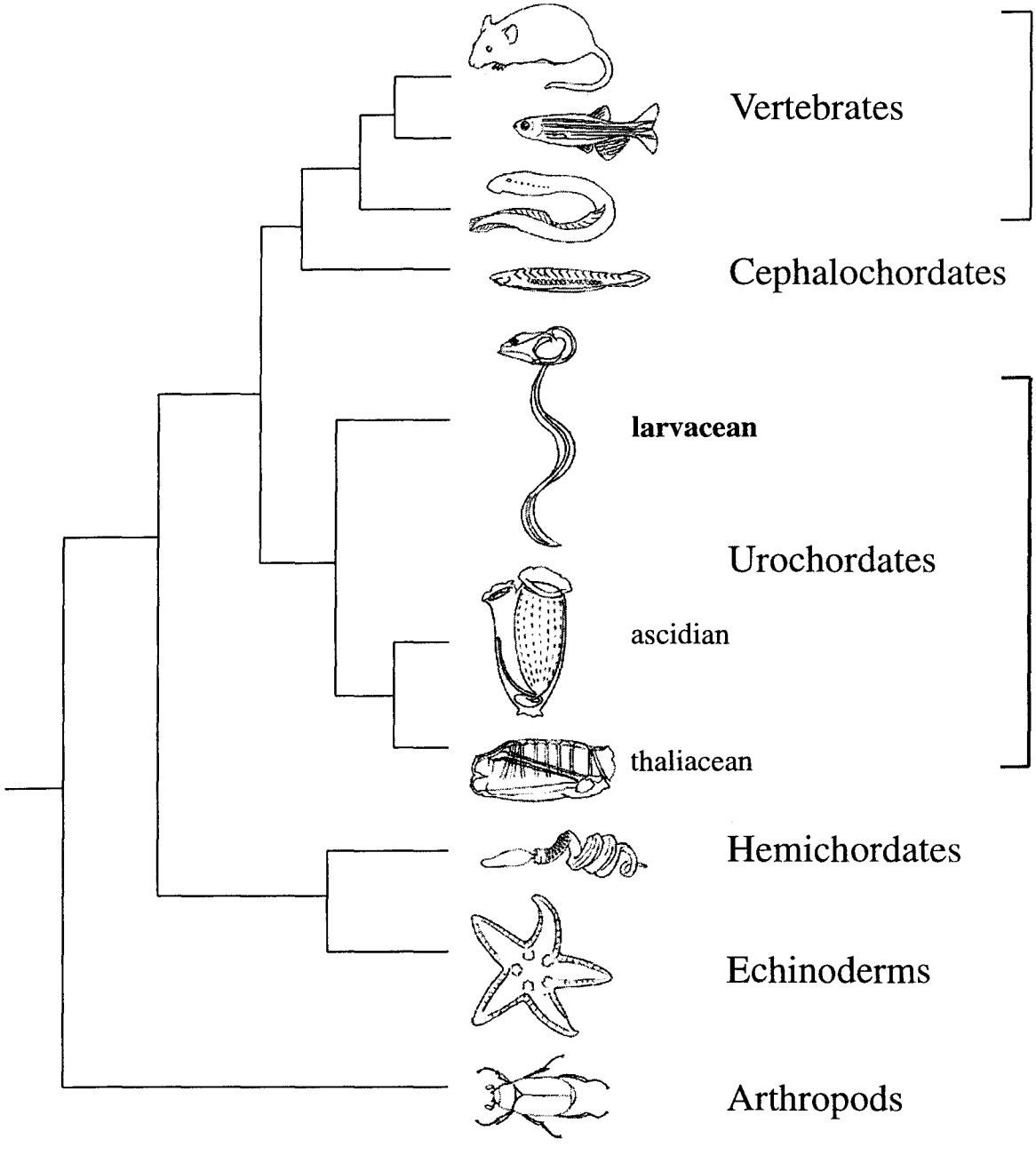


FIG. 1. Phylogenetic relationships among chordates and the other deuterostomes, with a protostome outgroup. This tree topology, which shows larvaceans branching basally in the urochordates, is modeled on rDNA phylogenies from Wada and Satoh (1994). Additional studies (Wada, 1998) suggest that Ascidiacea may be paraphyletic, in which case, ascidian and thaliacean subsets group together to the exclusion of pleurogonid ascidians; the basal position of larvaceans, however, is unchanged. The hemichordates and echinoderms are drawn as sister groups in accordance with gene trees from Wada and Satoh (1994) and Lynch (1999).

filter-feeding ascidians, the abyssal, ascidian-like sorberaceans, and the pelagic thaliaceans. Larvaceans may be more deeply rooted in the urochordate clade than these other classes, and retain a chordate body plan throughout their life history (Holland *et al.*, 1988). In contrast, metamorphosis reconfigures the ascidian into a sessile adult that little resembles the chordate design of its larva. Early authors in the field of chordate evolution (e.g., Willey, 1894) argued that a free-swimming morphology like that of larvaceans, cephalochordates and vertebrates was the primitive state of the chordate ancestor and that the radical metamorphosis of ascidians is evolutionarily derived. Recent phylogenetic analyses of rRNA sequences agree with this view, suggesting that the larvacean lineage diverged earliest within the urochordates (Christen and Braconnot, 1998; Wada, 1998; Wada and Satoh, 1994). In order to test whether the notochord-restricted expression of *T* is a pattern that is generalizable to the urochordate subphylum, and therefore to address the hypothesis that notochord specification is the ancestral function for *T* in the chordates, we cloned a *T* homologue from the larvacean *Oikopleura dioica* and analyzed its expression during embryogenesis.

In this first investigation of the molecular embryology of any larvacean, we found that during gastrulation of *O. dioica* embryos, *T* is expressed in a pattern similar to that in ascidians: in cells of the notochord lineage but not in mesoderm with lateral fates. When formation and differentiation of the larvacean tail is nearly complete, however, *T* is employed in non-notochordal expression domains, notably in regions of the forming gut and cells of the endodermal strand. This may reflect a deeper metazoan ancestry of this

gene's function in the patterning of posterior endoderm, and suggests that in the earliest chordate *T* was employed in both notochord and hindgut development.

MATERIALS AND METHODS

Animal Culture and Embryo Collection

Adult and juvenile *O. dioica* were collected from the plankton in coastal waters near Coos Bay, Oregon, sorted into 12°-14°C sea water in glass jars (2 liter), and fed until mature on a diet of natural phytoplankton and cultured algal strains (*Dunaliela tertiolecta*, *Isochrysis galbana*, *Rhodamonas lens*, *Nanochloropsis* sp., and *Micromonas* sp. (strain Dw-8)). Sperm was used immediately upon spawning or was collected by rupturing ripe males by pipette and stored undiluted at 12°C for up to 24 hours. Females were transferred into small fingerbowls to spawn and eggs were collected for synchronous fertilization in 0.45 micron-filtered seawater. Embryos were kept at 12°C.

Cloning and Phylogenetic Analysis

mRNA extracted from pooled embryos ranging from 2-celled to late tailbud stages was converted by Superscript reverse transcriptase (Gibco/BRL) to cDNA. A larvacean *T* fragment from this cDNA was amplified with nested degenerate PCR primers corresponding to the following amino acid sequences: NEMIVTK and WKYVNGEW (sense) and VTAYQNEE and NPFAKAF (antisense), corresponding to nucleotide sequences 5'AA(T/C)GA(A/G)ATGAT(T/C/A)GT(I/C)AC(I/C)AA3', 5'TGGAA(A/G)TA(C/T)GT(I/C)AA(T/C)GG(I/C)GA(A/G)TGG3', 5'TC(T/C)TC(A/G)TT(T/C)TG(A/G)TA(I/C)GC(I/C)GT(I/C)AC3', and

5'GAA(I/C)GC(T/C)TT(I/C)GC(A/G)AA(I/C)GG(A/G)TT3', respectively. mRNA extracted from tailbud stage embryos was reverse transcribed to make pooled cDNA for RACE PCR using the Marathon cDNA Amplification Kit (Clontech), and overlapping 3' and 5' RACE products of a full-length transcript were cloned into TA-Cloning vector pCRII (Invitrogen). The open reading frame was identified by the criteria of Kozak (1996). The 3' fragment, from corresponding amino acid position SVYMHP through the end of the 3' UTR was used to generate DIG-labeled riboprobes (Boehringer Mannheim DIG/Genius RNA Labeling Kit).

Nucleotide sequences were aligned using MegAlign software (version 3.08, DNASTAR) and trimmed to only unambiguously alignable regions. Gaps and any ambiguous adjacent regions were excluded. Bootstrapped phylogenetic trees of trimmed sequences were calculated by the neighbor-joining method (Saitou and Nei, 1987) using Clustal X software (http://www-igbmc.u-strasbg.fr/BioInfo/ClustalX/Top.html) and NJPLOT (http://pbil.univ-lyon1.fr/software/njplot.html) (Perrière and Gouy, 1996).

In Situ Hybridization

Embryos were fixed in (1 mM EGTA, 2 mM $MgSO_4$, 0.1 M MOPS, 0.5 M NaCl, 4% paraformaldehyde) for 1 hour at room temperature before being transferred to 70% EtOH for 1 to 3 hours. Embryos were transferred to fresh 70% EtOH for storage at –20°C. Prior to hybridization, pre-hatch embryos were rehydrated in PBT and manually dechorionated. A whole mount *in situ* hybridization protocol was adapted from Thisse *et al.* (1993) with the following modifications: embryos were hybridized in (50% formamide, 5X SSC, 0.1 mg/ml Heparin, 0.1% Tween-20, 5 mM EDTA, 1 mg/ml tRNA,

1X Denhardt's reagent) at 65°C with probe. Hybridized embryos were washed for 10 min. in (50% hybridization buffer/50% (2XSSC, 0.1% Tween-20)) at 65°C, 10 min. in (2X SSC, 0.1% Tween-20) at 65°C, 30 min. in (0.2X SSC, 0.1% Tween-20) at 65°C, 30 min. in (0.1X SSC, 0.1% Tween-20) at 65°C, 5 min. in (50% (0.1X SSC, 0.1% Tween-20)/50% PBT) at room temperature. Embryos were incubated for at least 60 min. at room temperature in a mixture of PBT, 2.5 mg/ml BSA, 2% goat serum. Then embryos were incubated for 2 hours at room temperature with presorbed alkaline-phophatase-coupled anti-digoxigenin antibody (Boehringer Mannheim) and treated to the standard detection described by Boehringer Mannheim. Embryos were mounted in 40% glycerol with 0.6 µg/ml DAPI. Embryos and hatchlings hybridized with a labeled sense riboprobe showed no staining at any stage (not shown).

RESULTS

Phylogenetic Relationships of OdiT

We cloned a fragment of *T* (*Brachyury*) from *O. dioica* using degenerate primers designed from aligned protein sequences of chordate *T* homologues. Sequencing this fragment provided information to design primers for the amplification of overlapping 3' and 5' fragments of a single gene by RACE PCR. The sequencing of this gene (Fig. 2) showed that it constitutes the 3' UTR and the entire coding region of a *T* homologue we named *O. dioica T* (*OdiT*). In a neighbor-joining (N-J) phylogeny of alignable T-box family protein sequences (173 amino acids), OdiT falls within a strongly supported clade

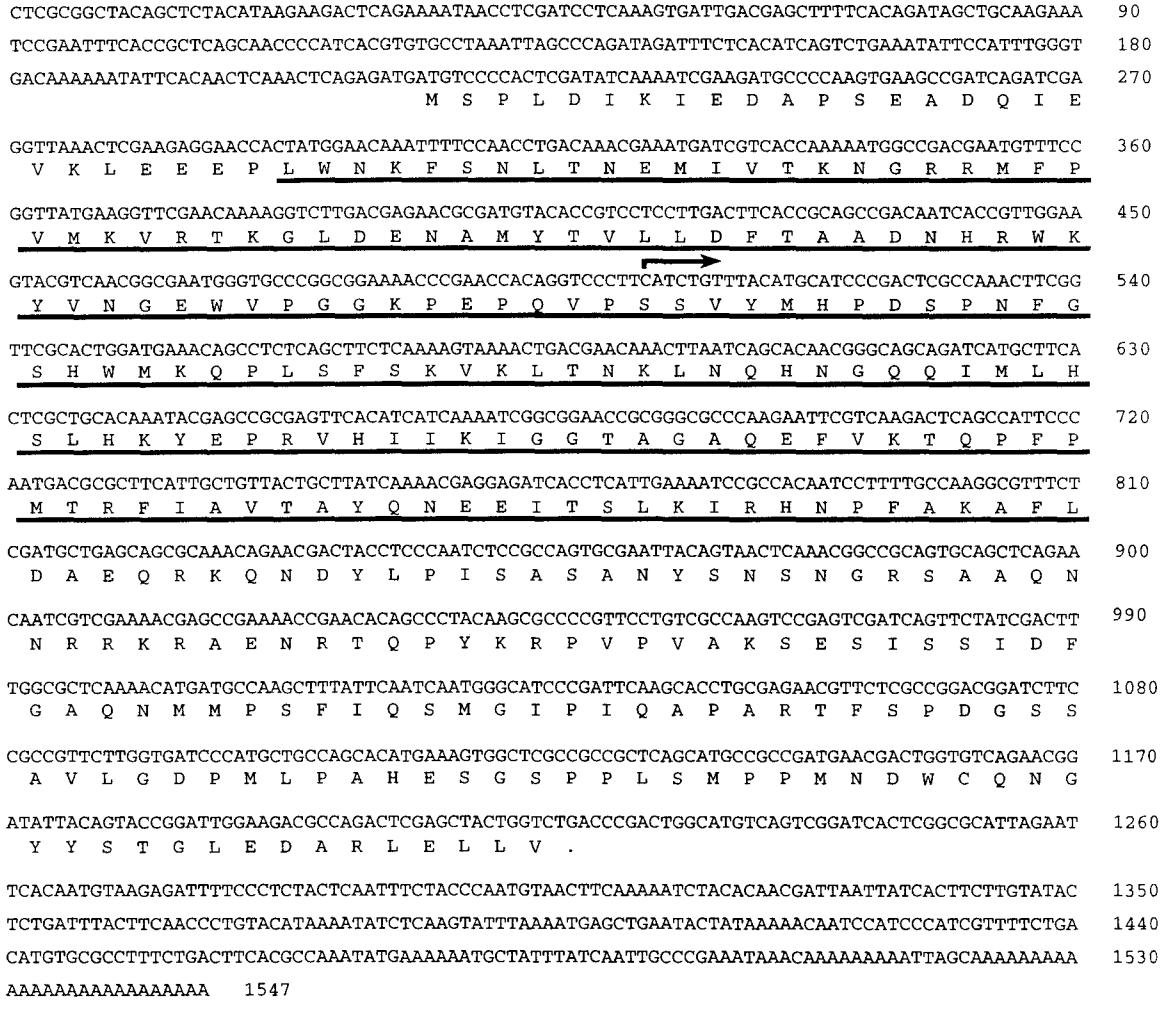


FIG. 2. The nucleotide and predicted amino acid sequences of the entire coding region of *OdiT*. The DNA-binding T-box domain is underlined. An arrow marks the 3' end of the clone used to generate riboprobe for *in situ* hybridizations.

(bootstrap = 999) including T proteins from vertebrates, cephalochordates, urochordates, a hemichordate, echinoderms and a cnidarian (Fig. 3 A). Excluded from this clade were genes representing other T-box subfamilies, as well as a second ascidian T-box gene called *As-T2* which is thought to be orthologous to vertebrate *tbx6* (Mitani *et al.*, 1999). Because the T-domain itself represents nearly all of the alignable sequence and provides a limited number of informative sites, the topology of the gene phylogeny does not conform well to taxonomic relationships. In a N-J phylogeny based on 519 unambiguously alignable nucleotide positions and rooted with sea urchin *T*, the tree topology better reflects expected taxonomic groupings. The *OdiT* groups strongly with its homologues in two ascidians (Fig. 3 B).

Expression of OdiT from Cleaving Embryos to Hatchlings

We studied the expression of *OdiT* in a developmental series of *O. dioica* embryos. The 3' RACE product contains 613 base pairs of predicted coding sequence and 374 bp of 3' UTR; this ~1 kilobase fragment was used to make labeled riboprobes for whole-mount *in situ* hybridization (Fig. 2). Blastomeres in *O. dioica* embryos do not cleave synchronously after the 16-cell stage, so cell numbers were ascertained by counting DAPI-stained nuclei in early embryos. *OdiT* expression was first detected in a bilateral pair of blastomeres in 30-cell embryos (Fig. 4 A, B). With the fifth cleavage (producing 30 cells), some blastomeres have already begun to overgrow their neighbors by epibolic divisions, marking the onset of gastrulation (Delsman, 1910). The two *OdiT*-expressing blastomeres appear at least partly submerged beneath superficial cells, and hence, gastrulation begins in *Oikopleura* before *OdiT* expression commences. This pair of

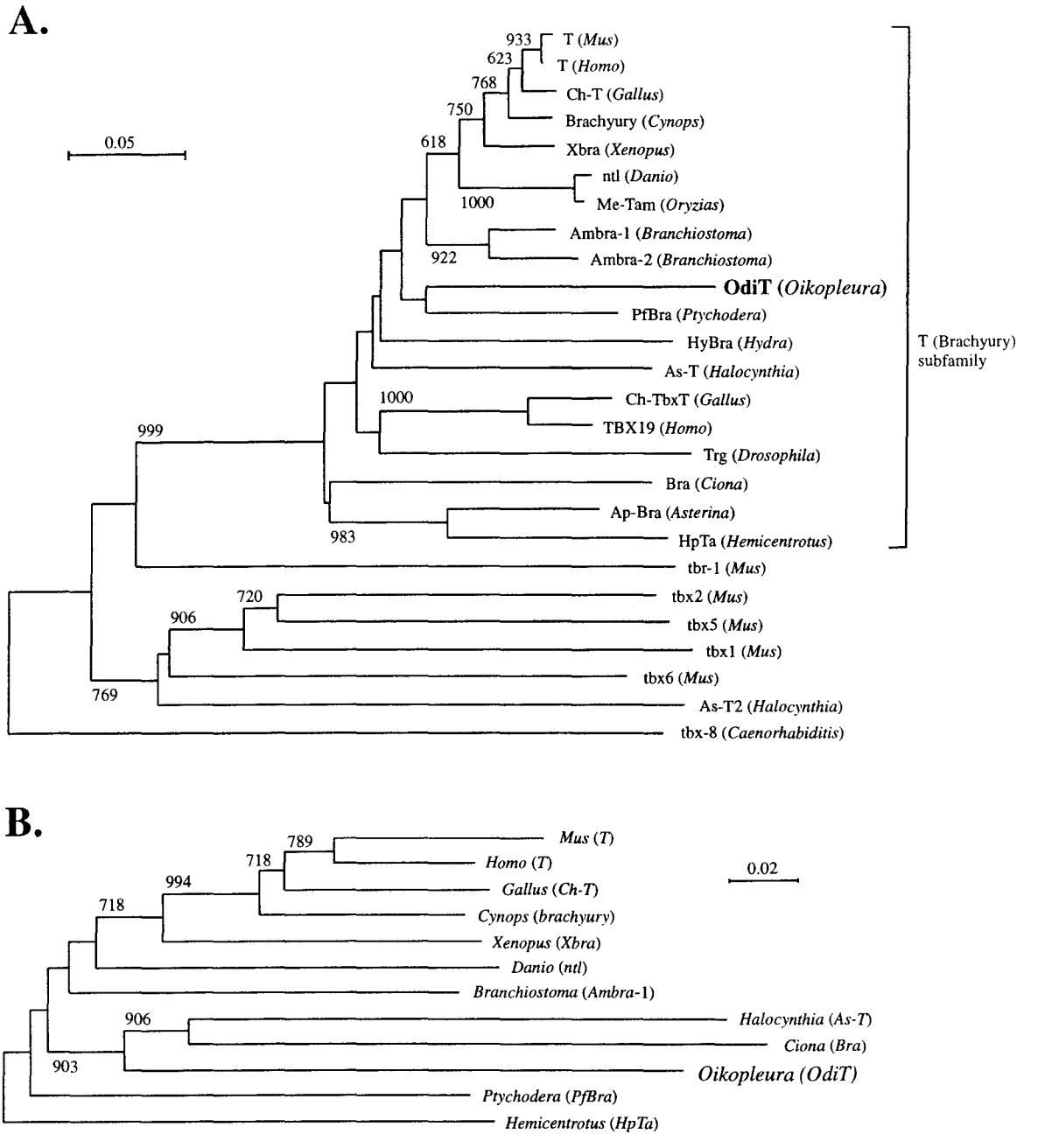


FIG. 3. Gene phylogenies indicate *OdiT* is a *Brachyury* orthologue. (A) Neighbor-joining (N-J) tree of alignable tbx family protein sequence. Included in the tree are *T* homologues from a broad phylogenetic range as well as representative members of the other tbx subfamilies (Wattler *et al.*, 1998). (B) N-J tree of alignable nucleotide sequence from chordate *T* genes, with two deuterostome outgroups, a hemichordate and an echinoderm. Bootstrap values greater than 50% are shown. Genbank accession numbers are X91903 (*Ambra-1*), P80492 (*Ambra-2*), AB018527 (*Ap-Bra*), D16441 (*As-T*), D83265 (*As-T2*), AF123247 (*Ciona Bra*), AB001939 (*Cynops brachyury*), U67086 (*Ch-T*), U67087 (*Ch-TbxT*), D50332 (*HpTa*), AF105065 (*HyBra*), AB001871 (*Me-Tam*), Q07998 (*ntl*), AF204208 (*OdiT*), AB004912 (*PfBra*), NM_003181 (*Homo T*), X51683 (*Mus T*), U57331 (*tbx6*), NM_005149 (*TBX19*), P55965, S74163 (*byn*), M77243 (*Xbra*).

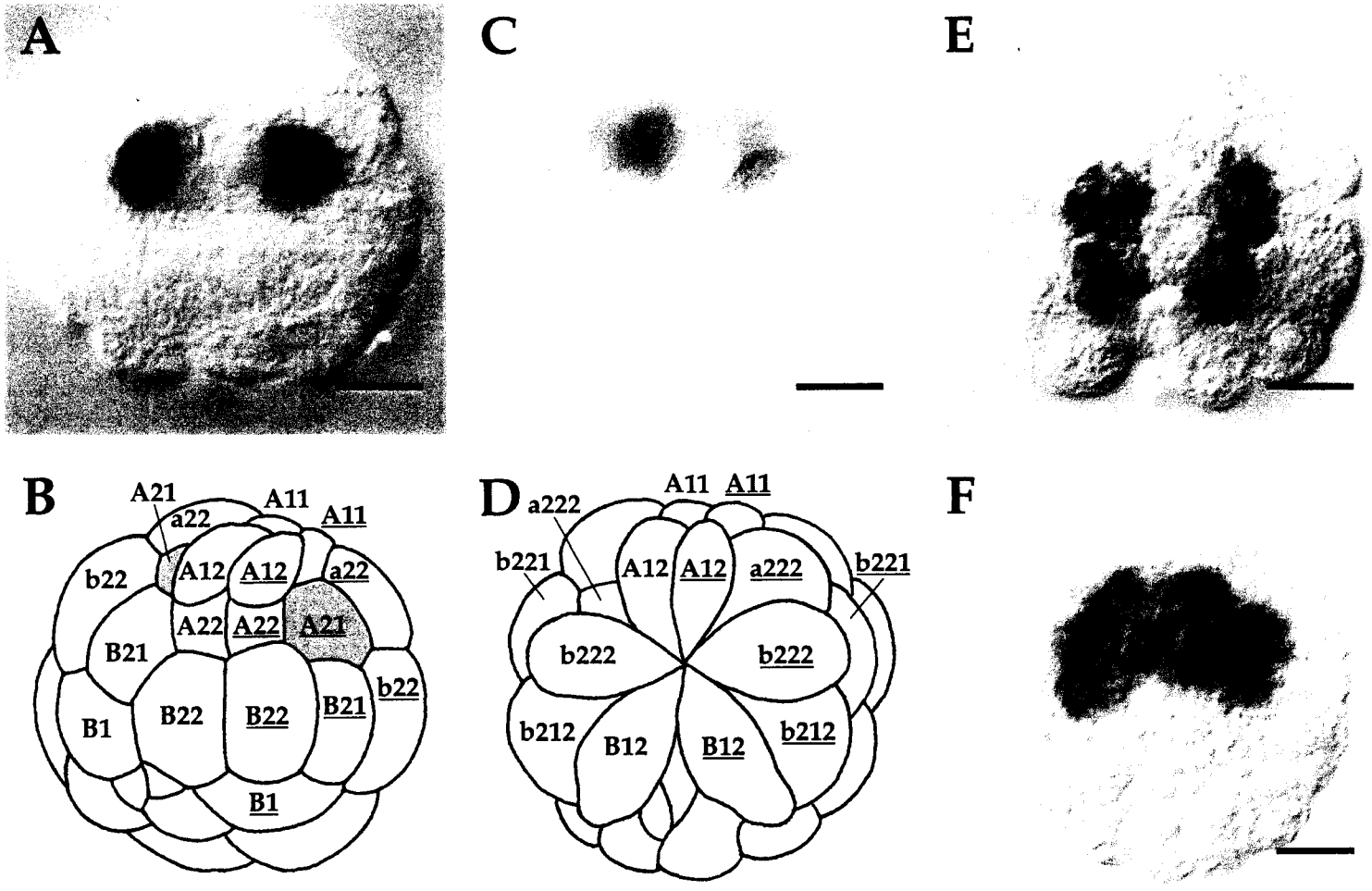


FIG. 4. Expression of *OdiT* in early *O. dioica* embryos (cleavage stage, gastrulae, and neurulae) is detected only in notochord precursors. (A) 30-cell embryo. (B) Diagram of 30-cell embryo, after Delsman (1910). Predicted *OdiT*-expressing cells shaded. (C) 44-cell embryo. Superficial view of blastopore showing 2 internalized *OdiT*-expressing cells. (D) Diagram of 44-cell embryo, vegetal view showing cell arrangement around blastopore. After Delsman (1910). (A-D) Vegetal pole view. (E) Early neurula. Vegetal pole at bottom. (F) Late neurula. Scale bar, 20 μm.

blastomeres are the only cells expressing *OdiT* until after the 44-cell stage, the end of gastrulation; they can be seen in the 44-cell embryo entirely submerged beneath the closed blastopore whose perimeter consists of 6 cells arranged as petals of a flower (Fig. 4 C, D). At neurula stage, when the ball-shaped embryo flattens slightly (Delsman, 1910), two bilateral pairs of blastomeres are labeled by the probe (Fig. 4 E). Up to this stage, expressing blastomeres were separated from one another by two intervening midline cells. Bilateral bands of expressing cells then become contiguous across the midline and form a chevron (Fig. 4 F), beginning the convergence and extension of the notochord.

The tail of *O. dioica* has twenty muscle cells arranged in two rows of ten cells each on either side of the twenty cells of the notochord. Just prior to extension of the notochordal rod, nineteen labeled cells are lined in a staggered row at early tailbud stage (Fig. 5 A). Muscle cell precursors adjacent to notochord precursors, are not labeled by the *OdiT* probe (Fig. 5 B, C). At hatching, nineteen notochordal cells are linearly arranged (Fig. 5 D), and a cell adjacent to the posterior-most cell in the row is also detected by the probe (Fig. 5 E). We interpret this as notochord cell twenty which has not yet intercalated into the notochordal axis.

In larvaceans, the tail is twisted ninety degrees relative to the trunk, dorsal to the left, ventral to the right. In hatchlings whose tails have just straightened, two *T* expressing cells lie adjacent and to the right of the most rostral notochord cells (Fig. 5 F). Their position and shape suggest these are the subchordal cells, products of the caudal endoderm, which migrate back into the tail after retracting into the trunk with rest of the endodermal strand (Delsman, 1912).

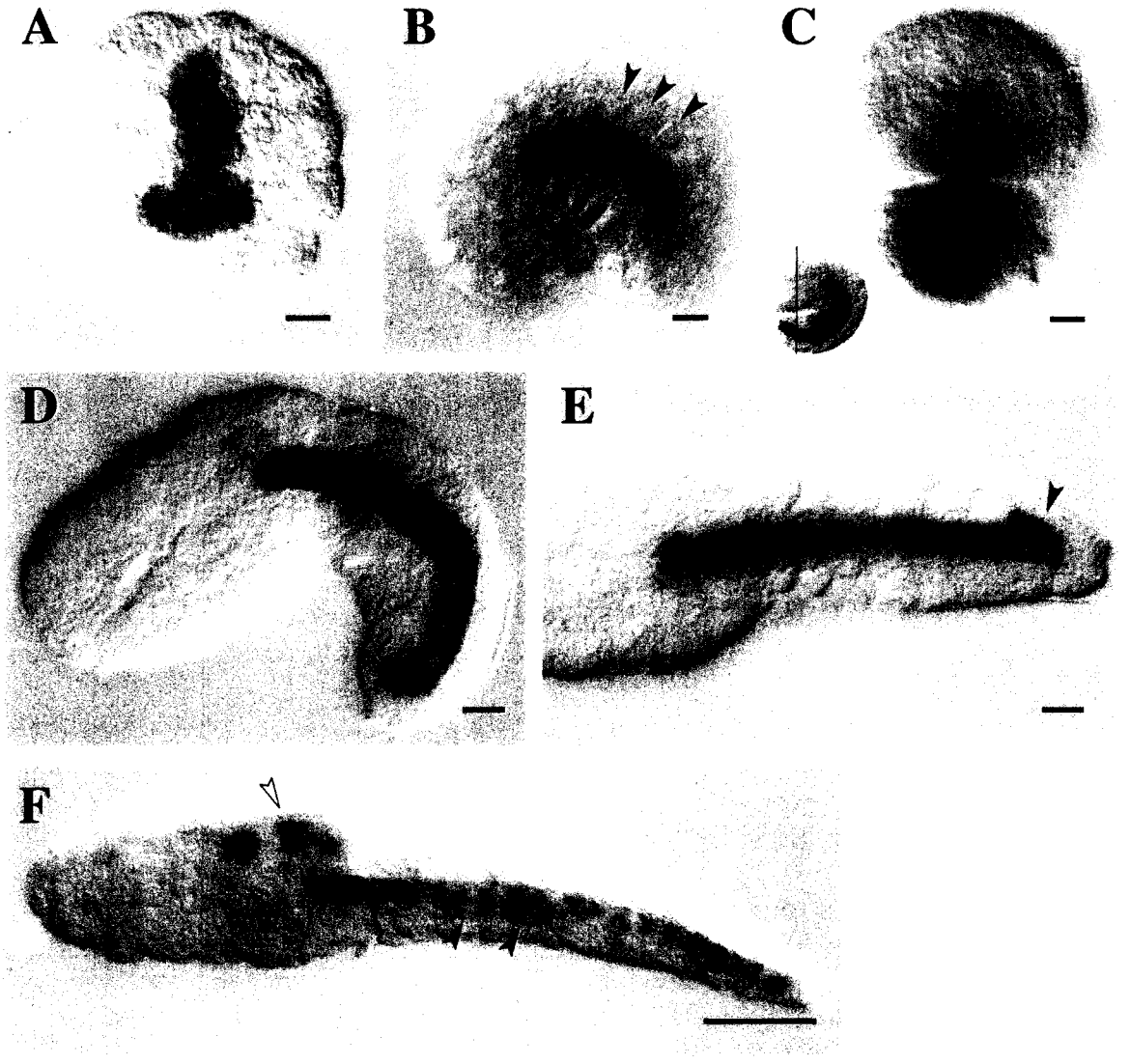


FIG. 5. Expression of in tailbud stage embryos and hatchlings. Transcripts are detected in the notochord, and also in endodermal expression domains. (A) Early tailbud stage, dorsal view, anterior at top. Posterior-most notochord cells lie in lateral positions and have not yet intercalated. (B) Mid-tailbud stage, lateral view. Arrows indicate nuclei of muscle cell precursors. (C) Late tailbud stage. Optical section through trunk and tail. Inset shows plane of optical section. (D) Early hatchling, lateral view. (E) Ventral view of early hatchling. Arrow indicates 20th notochord cell. (F) Larva 30 minutes after hatching. Black arrows indicate two subchordal cell precursors. Open arrow indicates dorsal trunk expression domain. Scale bars A-E, 10 μm; F, 20 μm.

In hatchlings six hours post-fertilization, as organogenesis is being completed in the trunk, *OdiT* signal rapidly decreases in the notochordal cells and new expression domains appear in dorsal and ventral cells of the posterior trunk. Initially, a flattened cell immediately rostral and ventral to the notochord and underlying the ectoderm is labeled by the probe (Fig. 6 A), as are a series of cells at a posterio-dorsal position in the trunk and lying just left of the midline under the epidermis (Fig. 5 F; Fig. 6 A); based on their positions, these cells may be of the developing hindgut and esophagus, respectively. In the latest stage hatchlings we examined ("Stage II"; Galt and Fenaux, 1990) the ventral expression domain was a smooth-bordered oblong lying against the epidermis and surrounded by the expanding body lumen (Fig. 6 B). When fully developed, a sausage-shaped rectum lies in this position.

DISCUSSION

In order to explore the evolutionary origins of developmental innovations in the phylum Chordata, we have begun to investigate the molecular embryology of *Oikopleura dioica.* This organism occupies an important position in chordate phylogeny, and larvacean anatomy suggests that it may facilitate detection of homologies between urochordates and vertebrates on the basis of both morphological and molecular criteria. To investigate the ancestral function of a gene important in the development of a novel chordate character, the notochord, we have cloned the larvacean orthologue (*OdiT*) of the vertebrate *T* gene and analyzed its expression during embryogenesis. We isolated the complete coding region of the *OdiT* message as well as the entire 3' UTR. Gene phylogenies established

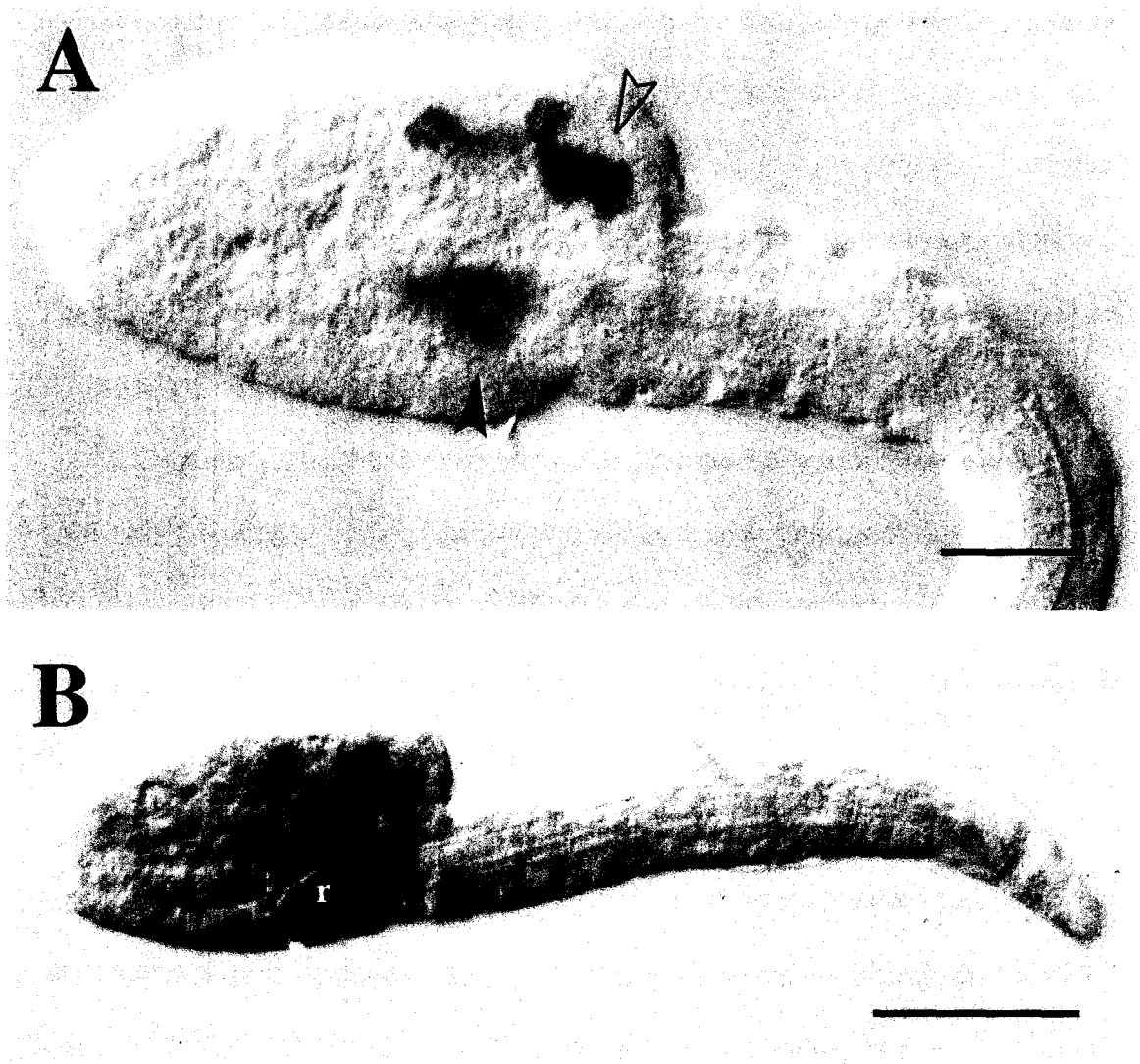


FIG. 6. *OdiT* expression persists into late organogenesis in the trunk, after expression is no longer detected in the notochord. (A) Larva, 1 hour after hatching, lateral view. Black and clear arrows indicate ventral and dorsal trunk expression domains, respectively. (B) Larva 2 hours after hatching. Ventral expression in developing rectum (r). L, trunk lumen. Scale bars A, 20 μm; B, 50 μm.

the orthology of *OdiT* to the *T* (*Brachyury*) subfamily of T-box genes by analysis of both predicted amino acid sequence and nucleotide sequence (Fig. 3). In a phylogeny of T-domain sequences from a broad range of T-box proteins, *OdiT* groups unambiguously within the *T* clade. And in a phylogeny of alignable DNA sequences using a sea urchin *T* as the outgroup, *OdiT* falls within the chordate sequences, and groups strongly with the ascidian *T*s. Orthologues are genes in two species that can trace their ancestry back to a single gene in the last common ancestor of those species (Fitch, 1970). We conclude that *OdiT* is a true *T* orthologue and not a different member of the T-box gene family.

In cephalochordates and vertebrates, *T* is expressed in presomitic mesoderm, in the notochord, and in involuting mesodermal and endodermal cells surrounding the closing blastopore (or equivalent)(e.g., Wilkinson *et al.*, 1990; Schulte-Merker *et al.*, 1992). In previously investigated basal chordates (ascidians), *T* is expressed in cells that have become segregated during cleavage to a purely notochordal lineage, but not in mesoderm lateral to the presumptive notochord or the nearby endoderm (Yasuo and Satoh, 1993, 1994; Corbo *et al.*, 1997). In light of this result in ascidian embryos, it has been proposed that the ancestral role for *T* in chordates was in notochord specification (Yasuo and Satoh, 1994). If the notochord-only hypothesis is true, then the non-notochordal expression domains in cephalochordate and vertebrate embryos represent functions of the *T* gene that were acquired after the evolutionary split between the urochordates and the cephalochordate-vertebrate lineage. Alternatively, the expression pattern in the cephalochordate-vertebrate clade -- in the notochord, surrounding

mesoderm, and endoderm -- could have been the ancestral pattern, and ascidian embryos might have lost *T* function in endoderm and paraxial mesoderm formation.

Peterson *et al.* (1999) favor the notochord-plus-surrounding-cells hypothesis, and propose a parsimonious model for the evolution of *T* function based on gene expression in arthropods and the deuterostome phyla. According to this model, prior to the protostome/deuterostome split, *T* served a function in posterior gut development. Kusch and Reuter (1999) additionally propose that *T* had an ancestral role in mesoderm development in the last common ancestor of protostomes and deuterostomes. The chordate lineage acquired notochord expression after divergence from the echinoderm/hemichordate clade. Although vertebrates and cephalochordates express *T* in posterior endoderm and paraxial mesoderm early in development, neither of these proposed ancestral expression domains has been demonstrated in the normal development of ascidians. In order to evaluate these alternatives, we examined *T* expression in a basally diverging urochordate. Recent molecular phylogenetic analyses have placed the class Appendicularia at a basal node within the urochordate clade (e.g., Wada, 1998). This contradicts the prevailing hypothesis that larvaceans arose by paedomorphosis from the thaliaceans (Garstang, 1928) and are the most derived of the urochordates. Instead, the molecular data support the work of morphologists who argued that the motile body plan of the adult larvacean is a plesiomorphy of Chordata (e.g., Lankester, 1890; Lohmann, 1933). If the notochord-restricted expression pattern of *As-T* is the ancestral condition for Urochordata, the larvacean homologue is predicted to be expressed only in notochord lineage cells. If the ascidian pattern is an apomorphy of the ascidian clade and

larvaceans are basal within the urochordates, the larvacean homologue might share an expression pattern with the outgroup, the cephalochordate-vertebrate clade. In other words, *OdiT* might be expressed in endoderm and non-notochordal mesoderm.

Early Expression of Larvacean T is Restricted to the Developing Notochord

Expression of *OdiT* was detected neither in endodermal cells while these were internalized by epiboly during gastrulation, nor in paraxial mesoderm during tail formation (Figs. 4; 5 A-E). Early *OdiT* expression appeared to be restricted to cells of notochordal fate, as identified by comparison with Delsman's fate map (Delsman, 1910) and as deduced by the continuity of *T* expression in a developmental series that included hatchlings with differentiated notochords. We interpret expression in a single cell adjacent to the end of the nineteen cells lined up in the notochord as a twentieth, late interdigitating notochord cell. This would agree with ascidian development where interdigitation of notochord cells is not synchronous (Cloney, 1964) and notochord elongation is completed in the rostral end earlier than in the caudal end of the extending tail (Miyamoto and Crowther, 1985; Corbo *et al.*, 1997). This cell, however, remains distinctive; Galt (1972) noted a twentieth cell in *O. dioica* that is histologically notochord-like but it remains physically separate from the rest of the notochord.

Our results support some of the observations made by Delsman in *O. dioica* embryos and his predictions of cell fate which he based on comparison with ascidian embryos (Delsman, 1910). There is no modern cell-lineage map for this or any larvacean. The first pair of cells expressing *OdiT*, however, are in the position of A21/A21, left and right side progeny of blastomeres A2/A2 which Delsman (1910) predicted would give

rise to the notochord (Fig. 4 A, B). At early neurula stage, 4 cells are detected by the probe (Fig. 4 E). Two are likely to be the original A21 pair, while the second pair may be progeny of B21/B21. The B21 pair are proposed to correspond approximately to ascidian B6.2/B6.2 (Delsman, 1910; Galt and Fenaux, 1990). In ascidians, progeny of these B-line (posterior, vegetal) blastomeres express *As-T* a cleavage later than do A-line (anterior, vegetal) notochord precursors (Yasuo and Satoh, 1994) and produce the posterior-most notochord cells (Nishida, 1987; Satoh, 1994).

While the fate maps of larvaceans and ascidians are likely to share similarities, the two classes of urochordates differ significantly in the timing and pace of gastrulation. Delsman observed that the gastrulation of *O. dioica* is precocious and succinct; *O. dioica* begins gastrulating with the 5th cleavage (during the transition from 16 to 30 cells) and completes it by the end of the 6th wave of cleavage resulting in 44 cells (Delsman, 1910; Galt and Fenaux, 1990; Fig. 4 A-D). *OdiT* expression is first detected at the 30-cell stage, the early gastrula (Fig. 4 A, B). *As-T* expression, in contrast, is first detected in 64-cell ascidian embryos at least one cleavage before the onset of gastrulation which begins after 110 cells (Yasuo and Satoh, 1994). The onset of *T* expression in these urochordates is apparently independent of the timing of gastrulation. In ascidians, *As-T* is not detected in blastomeres until they become segregated to a purely notochordal fate (Yasuo and Satoh, 1994). Expression of *T* in larvaceans is probably also constrained by the timing of segregation of cell fates with the result that, while it is detected at an earlier cleavage stage than in ascidians, it first appears during, rather than before, gastrulation.

While these results, taken alone, do not contradict the hypothesis that *T*'s role in the ancestral chordate was in notochord specification, rather than in broader mesodermal domains, expression in endoderm cells (discussed below) supports the hypothesis that *T* also performed a function in endoderm development in the chordate ancestor.

Larvacean T is Expressed in Posterior Endoderm

Most oikopleurids have in the tail a characteristic cell type that varies in number between species, *O. dioica* having two such "subchordal cells" (Lohmann, 1933). The subchordal cells are named for their position ventral to the notochord; these cells appear on the animal's right because the tail is twisted ninety degrees relative to the trunk, and thus, caudal right corresponds to ventral. Although their function is not understood, the subchordal cells may have a secretory role (Fredriksson and Olsson, 1991), and Delsman (1912) described them as a product of the caudal endoderm, the endodermal strand. This one cell wide strand runs ventral to the notochord in a position similar to vertebrate hypochord cells (Kimmel *et al.*, 1995) which express many of the same genes expressed in notochord (e.g., Yan *et al.*, 1995). As the caudal endoderm retracts into the trunk during extension of the tail, the presumptive subchordal cells detach themselves from the endodermal strand and migrate back into the tail (Delsman, 1912). Shortly after hatching, two cells to the right of the notochord are detected with the *OdiT* probe (Fig. 5 F). Their position relative to the notochord, their elongate shape, and the fact that they are seen sometimes at the base of the tail or sometimes further caudal suggest they are the two migrating subchordal cells later found in the tailfin of adult *O. dioica*. The putative subchordal cells apparently express *T* as they migrate. What could be the role of *T* in

these cells? In mouse, *T* is necessary for migration of cells through the primitive streak (Wilson *et al.*, 1995), and the authors propose that genes downstream of *T* mediate cell-surface changes necessary for migration of cells away from the midline after ingression. Similarly, in fly embryos mutant for the *T* homologue, *byn* (*brachyenteron*), caudal visceral mesodermal cells are retarded in their ability to migrate, a deficiency the authors attribute to a defect in the cells' adhesive properties (Kusch and Reuter, 1999). The endodermal strand, including subchordal cells, are the only cells described as migratory in larvacean development (Delsman, 1912), and *T* may function to alter adhesion properties for the segregation and migration of these subchordal cells from the endodermal strand.

Expression of *OdiT* after hatching further supports the hypothesis that *T*'s function originally included a role in patterning posterior endoderm. *O. dioica* hatches at 6 hours after fertilization (at 12°C) while organogenesis is still in progress in the trunk and there are no open lumens in the gut or body cavity (Galt and Fenaux, 1990). Shortly after hatching, new expression of *OdiT* appears in the trunk while detectable transcript rapidly disappears from the notochord. The elongate cluster of expressing cells just rostral to the notochord is likely the developing hindgut. In the latest stage of the hatchlings we examined, the borders of this expression domain had smoothened to a sausage shape apposed to the epidermis where Delsman (1912) shows that the anus forms (Fig. 6 B). The identity of an expressing row of cells in the posterio-dorsal trunk is less certain. Given its position in the dorsal trunk slightly left of the midline, we interpret it as the developing esophagus which curves left to feed dorsally into the left half of the two-

lobed stomach. Alternatively, these posterio-dorsal cells may be of the developing left stomach lobe itself which later is distinguished from the right lobe by a row of gland cells extending along the left wall to the esophagus (Fenaux, 1998).

Expression of *OdiT* in the larvacean hindgut provides an interesting parallel to the expression of *T* in an enteropneust hemichordate, *Ptychodera flava*. *PfBra* (a *T* homologue in *P. flava*) is expressed in the anal region of the developing larval hindgut and in the intestine of the metamorphic juvenile (Tagawa *et al.*, 1998; Peterson *et al.*, 1999). The hemichordate expression as well as similar expression of *T* in an embryonic seastar, however, appears to be in structures derived from invaginating ectoderm, rather than in the endodermally derived portion of the gut (Peterson *et al.*, 1999; Shoguchi *et al.*, 1999). The expression of *T* in the hindgut of larvaceans and hemichordates, therefore, cannot be homologized without violating the distinction between two embryonic germ layers. Peterson *et al.* (1999) infer a broader homology of *T* expression in posterior gut between the deuterostome phyla and a protostome outgroup (Arthropoda). In their model for evolution of *T* utilization, the ancestral role for *T* in bilaterians was to help form the posterior gut. In arthropods, the apparent *T* orthologue, *byn*, is expressed in the developing hindgut (Kispert *et al.*, 1994). Kispert *et al.* (1994) note that it is somewhat arbitrary to define this part of the arthropod gut as ectodermal because this label is based on epithelial properties shared with epidermis rather than on the hindgut's internalized position after gastrulation. If one regards hindgut development in these bilaterians as homologous, the larvacean pattern can be reconciled with that described in outgroups to the chordates (hemichordates, seastars, and insects). Therefore, our data support the

model that *T* had a role in posterior gut formation in the protostome/deuterostome ancestor and that this role has been conserved in larvacean urochordates. An apparent absence of *T* expression in vertebrate and cephalochordate hindgut development might represent a shared evolutionary loss of late expression in posterior endoderm in this lineage. The absence of reported *T* expression in the ascidian hindgut might also represent an evolutionary loss. Alternatively, there might have been a heterochronic shift in *T* expression to a later developmental stage than has yet been experimentally examined in ascidians. This explanation predicts that *T* expression will resume at ascidian metamorphosis when the endoderm is patterned to make the juvenile gut. Finally, in the urochordate lineage, absence of early expression of *T* during gastrulation in presumptive endoderm and mesoderm might be a loss correlated with the mosaic nature of urochordate early development or with a change in the mechanics of gastrulation due to small cell number.

This is to our knowledge the first publication on the molecular embryology of any member of the class Appendicularia. Larvaceans have been largely ignored, in part because of the difficulties in culturing them, but also because of the dominance over the last seven decades of Garstang's hypotheses concerning the evolution of the chordates. Garstang (1928) proposed that the cephalochordate-vertebrate lineage and the larvaceans arose by separate events of paedomorphosis; both lineages independently evolved from motile larvae of different sessile chordate ancestors. Ascidians are the most primitive of the urochordates, he argued, while larvaceans are the most derived, having sprung from a larval stage in the thaliacean lineage. Because molecular data suggest that larvaceans

diverged basally in the urochordate clade (e.g., Wada, 1998), and because some aspects of ascidian molecular embryology may not be generalizable to all of Urochordata, investigation of developmental mechanisms in larvacean urochordates could be vital to our understanding of the common ancestor of the chordates.

Bridge to Chapter III

What did the ancestor of the vertebrates look like, and how did vertebrates evolve from their chordate ancestors? In order to approach these evolutionary questions, we first have to agree upon the answers to some wieldier questions: which traits are shared chordate characters and which are uniquely vertebrate characters? These questions are about timing the origin of evolutionary innovations within the chordate phylogeny. Several apparent novelties seem to set vertebrates apart from their chordate relatives, and it has been proposed that such novel traits both permitted and caused the vertebrates to depart radically from the other chordate lineages. Among these innovations are embryonic structures called placodes, which contribute to paired sensory organs of the vertebrate head. There is growing evidence, however, that the origins of some placodes could have predated others, perhaps by as much evolutionary time as to have predated the vertebrates themselves. The discovery in non-vertebrate chordates of sensory organs with morphological and ontogenetic similarity to vertebrate placode derivatives would strongly suggest that some placodes could have been present in the last common ancestor of two or all three of the chordate subphyla. Such a finding would require a reevaluation of the role placodes might have played in driving the origin of the vertebrates.

The following chapter describes an analysis of the development of two organs in the larvacean *O. dioica* that have been hypothesized to be homologues of the vertebrate olfactory organ and anterior pituitary, both placode derivatives. Homologues of several genes that are important for the development of these placodes in vertebrates are expressed in the larvacean ventral organ and the ciliary funnel. These findings are consistent with an origin for placodes at least as ancient as the common ancestor of the three extant chordate subphyla.

CHAPTER III

DEVELOPMENTAL EXPRESSION OF LARVACEAN *EYA*, *SIX*, *PAX* AND *PITX* GENES: IMPLICATIONS FOR THE ORIGIN OF PLACODES AND THE EVOLUTION OF VERTEBRATES

Embryonic placodes contribute to a number of important sensory organs of the vertebrate head, and are widely believed to be a vertebrate innovation that helped propel the evolution of predatory vertebrates from simple, filter-feeding ancestors. The suggestion that placode-like structures might exist in non-vertebrate chordates is controversial, but has recently been rekindled by gene expression studies in cephalochordate and ascidian embryos. The understudied larvacean urochordates may offer important insights into the origin of placodes. Two organs in the larvacean *Oikopleura dioica* were suggested on the basis of topography and morphology to be homologous to placode-derived vertebrate organs. I have analyzed the embryonic expression of larvacean *Eya*, *Six*, *Pax* and *Pitx* genes in *O. dioica*, and found that several genes from these placode-marking gene families are expressed in the primordia of the ventral organ and the ciliary funnel. These findings are consistent with hypotheses set forth by earlier authors who proposed that these larvacean organs are homologous to the vertebrate olfactory organ and anterior pituitary, both products of embryonic placodes.

INTRODUCTION

A popular evolutionary scenario describes the rise of vertebrates from filter feeding ancestors to aggressive raptors, and attributes this change largely to the innovation of two embryonic tissues: neurogenic placodes and neural crest (Gans and Northcutt, 1983; Northcutt and Gans, 1983). Placodes and crest broadly share several characteristics: they form from columnar epithelia, undergo epithelial-to-mesenchyme transformation, produce migratory cells by delamination (though not true of all placodes), and generate some overlapping cell types (e.g., neuroendocrine cells, sensory neurons, glia). They differ in significant ways, however, and may not share a common evolutionary origin. Placodes originate only in the head ectoderm, while crest cells migrate from the margin between the neural plate and non-neural epidermis in both the trunk and head. Additionally, placodes and crest produce only partly overlapping classes of cell types (reviewed in Baker and Bronner-Fraser, 1997). Placodes contribute to a number of paired sensory organs that would have facilitated the evolution of an active, predatory life-style in an ancient vertebrate; these placode-derived structures include the olfactory organ, otic vesicles of the inner ear, the lenses of the eyes, cranial sensory ganglia such as those involved in gustation, and electrosensory lateral line organs.

Because no placode-like structures or their derivatives have been uncontroversially recognized in non-vertebrate chordates, biologists generally assume that placodes rapidly arose together within the early vertebrates. In support of this

hypothesis, placodes not only share the characters cited above, but they also express a number of related genes, suggesting their evolutionary and developmental affinity. While none of the following genes is exclusively expressed in placodes, they do show regional specificity for placodal domains. Paralogues of *eyes absent* (*eya*), a gene family important in development of the eye in both flies and vertebrates, are also expressed in the adenohypophyseal (anterior pituitary) placode and all of the neurogenic placodes (Xu *et al.*, 1997; David *et al.*, 2001). *Pitx* genes mark the most anterior placodes, the olfactory and adenohypophyseal (Gage and Camper, 1997; Lanctôt *et al.*, 1997). Several *Pax* family genes are important for the development of placodes: *Pax6* is expressed in the olfactory, lens, and adenohypophyseal placodes (Walther and Gruss, 1991; Kioussi *et al.*, 1999; reviewed in Callaerts *et al.,* 1997), and the closely related genes *Pax2* and *Pax8* are expressed in the otic placodes (Heller and Brändli, 1999). *Six* family genes are important for eye development in both flies and vertebrates, but vertebrate *Six* paralogues are also expressed in the olfactory, adenohypophyseal, and otic placodes (Oliver *et al.*, 1995; reviewed in Kawakami *et al.*, 2000). Significantly, *Six* genes are important in the developing forebrain, including those regions that will become physically and functionally connected to the olfactory and adenohypophyseal organs: the olfactory bulbs and the neurohypophysis or posterior pituitary (e.g., Jean *et al.*, 1999; Ghanbari *et al.*, 2001). Despite examples of shared gene expression in placodes, Begbie and Graham (2001) have recently argued that the diversity of the organs derived from placodes, and the diversity of ways in which they are induced to form indicate that placodes cannot be considered closely related embryonic structures. In addition, the placodes, which

variously contribute to eyes, ears, olfactory and lateral line organs, the pituitary, and sensory ganglia, might not all share a common evolutionary origin within Vertebrata.

Are Placodes a Vertebrate Invention?

The unique acquisition of placodes might have been very important for the changes in life history and body plan experienced by early vertebrates. If placodes predated the vertebrates, however, we will have to qualify the role placodes might have played in vertebrate evolution. Potential placode homologues might be identifiable in non-vertebrate chordates on the basis of developmental, topographical, ultrastructural, and molecular characters. Some evidence in favor of a more ancient origin for placodes has come from analysis of cephalochordate and ascidian molecular embryology. An amphioxus *Msx* gene marks a pair of regions in the anterior epidermis of the larva, and because vertebrate *Msx* genes are placode markers, the authors suggest these patches could represent ectodermal placodes (Sharman *et al.*, 1999). This *Msx* domain may mark the primordium of sensory structures called the corpuscles of de Quatrefages, but the authors don't offer a direct correspondence of these to a particular vertebrate placode derivative. The single ascidian homologue of vertebrate *Pax2*, *Pax5* and *Pax8* is expressed in paired primordia of the atrium (exhalent siphon) which the authors propose to be homologous to the otic placodes (Wada *et al.*, 1998). As in vertebrates, *Pitx* is expressed in the developing Hatschek's pit, the putative cephalochordate homologue of the vertebrate adenohypophysis (Yasui *et al.*, 2000). Other investigative tools and simple observation have also been used to argue the presence of placodes or their derivatives in non-vertebrate chordates. Hatschek's pit showed immunoreactivity for several vertebrate

pituitary hormones (Nozaki and Gorbman, 1992), and a homogenate of Hatschek's pit tissue could induce spermiation in young toads, suggesting that it has gonadotropic hormones (Fang and Wang, 1984). Burighel *et al.* (1998) and Manni *et al.* (1999) argue that the ascidian neurohypophyseal duct is a candidate homologue for both neural crest and for the combined olfactory and adenohypophyseal placodes, citing its neurogenicity and the migratory behavior of the cells it spawns.

The Adenohypophysis and Olfactory Placodes May be More Ancient than Vertebrates

Various authors have argued that the adenohypophysis predated the vertebrates and was anciently chemosensory (Olsson, 1990; Nozaki and Gorbman, 1992; Gorbman, 1995; Manni *et al.*, 1999). These arguments are based in part on the location of the placode or its presumed homologues in non-vertebrate chordates in the roof of the buccal cavity where, in an aquatic organism, it would be in constant communication with waterborne chemicals from the environment. Because of the importance of synchronizing sexual maturation and reproductive behavior among conspecifics, the systems for collecting olfactory information, such as sex pheromones, and for eliciting maturational and behavioral changes are probably anciently linked (Muske, 1993).

Ontogenetic, molecular, and comparative morphology data collected from the range of extant vertebrates suggest the olfactory and pituitary organs were functionally and perhaps physically connected in the last common ancestor of the vertebrates. In modern agnathans, the two organs originate from a common nasohypophyseal placode, and a connection between them persists in a nasohypophyseal duct until metamorphosis in the lamprey ammocoete larva. The primordia of the separate pituitary and olfactory

placodes of gnathostomes are contiguous with each other and with the hypothalamus primordium in earlier, neural plate stage embryos. Transplantation fate mapping in amphibians and birds shows that both the adenohypophyseal and olfactory placodes develop from the anterior neural ridge (ANR), which forms a crescent around the rostral neural plate (Couly and Le Douarin, 1985; Kawamura and Kikuyama, 1992). Kawamura and Kikuyama (1992) found that the amphibian medial ANR gave rise to the hypophyseal cone (equivalent to Rathke's pouch in mammals), and that this tissue, after separating from the neural tube, divided rostro-caudally to form the adenohypophysis but also a part of the preoptic hypothalamus. Therefore, some ANR neuronal cells migrated into the rostral brain after neural tube closure was complete. The olfactory placodes exhibit a similar capability; they produce not only the cells of the olfactory epithelium, but also cells expressing gonadotropin releasing hormone (GnRH) that migrate away from the placode and into the brain (Murakami *et al.*, 1992). Some of these GnRH-producing neurons lying along the olfactory nerve may modulate the responses of ciliary receptor neurons to environmental chemical cues that are pertinent to sexual maturation or activity (Wirsig-Wiechmann, 2001; Muske, 1993). Other cells migrate further into the brain to populate a region in the hypothalamus; these cells send processes to the neurohypophysis where, by GnRH secretion, they will direct hormone release from the anterior pituitary (reviewed in Tobet *et al.*, 1997). In addition to these ontogenetic and functional connections between the olfactory and pituitary organs, molecular markers expressed by both tissues further support their affinity. Antigenicity for pituitary hormones has been found in the developing olfactory organ: somatostatin

immunoreactivity in the developing chick olfactory organ, and ACTH-ir in the olfactory organ and nasohypophyseal duct of ammocoetes (reviewed in Gorbman, 1995).

An Analysis of Placode Candidates in a Larvacean Urochordate

Could the evolution of placodes have predated the evolution of the vertebrates, and are there homologues of specific vertebrate placodes in other chordates? To address these questions, I have studied the molecular embryology of a larvacean urochordate, *Oikopleura dioica*. While other urochordates, including ascidians and thaliaceans, undergo a metamorphosis that radically rearranges or destroys their chordate characters, larvaceans retain an easily recognized chordate body plan into adulthood. *O. dioica* has other traits that facilitate study of its anatomy and embryology: it is transparent at all stages, can produce clutches of up to 500 eggs, has very rapid development, and has a short generation time (less than two weeks) with seasonally-independent spawning. The larvacean brain consists of fewer than 70 cells (Martini, 1909), not all of which are neurons (Georges *et al.*, 1988). Despite this extreme miniaturization and apparent disparity from the vertebrate brain, two organs of the larvacean nervous system, the ventral organ and the ciliary funnel, have been proposed as homologues of vertebrate placode derivatives. Bollner *et al.* (1986) studied the fine structure and innervation of the ventral organ, a comb of ciliary cells in an epidermal pit near the lower lip, and suggested its homology to the vertebrate olfactory epithelium. Earlier authors had also referred to the ventral organ as an olfactory organ, but at other times as a taste bud (Martini, 1909; Lohmann, 1933). The ciliary funnel is a cone of tissue extending ventrally from the right side of the brain and ending in a perforation into the roof of the pharynx. Because of its

resemblance to early stages of the developing anterior pituitary (Rathke's pouch), the ciliary funnel has been homologized with the adenohypophysis. Cilia lining the funnel and directed inward beat constantly, creating a current from the buccal cavity into the funnel (Lohmann, 1933). Holmberg (1982), who analyzed the ultrastructure of the ciliary funnel, suggested the distinct ventral and dorsal regions of the funnel correspond to the adenohypophysis and the neurohypophysis, respectively. Although the capacity for chemosensation in *O. dioica* needs to be experimentally confirmed, individuals may be responsive to the chemical presence of conspecifics; randomly chosen pairs of animals, when cultured in close quarters, mature into a female and a male seemingly more often than would be expected if sex determination were random or genetically predetermined (Bassham, unpublished observation). The presence of olfactory- and pituitary-like organs in a larvacean urochordate would support an ancient origin for at least some of the placodes inherited by modern vertebrates.

In this study, I show that the homologues of several genes important for the early development of the olfactory and pituitary placodes are expressed in the primordia of the larvacean ventral organ and ciliary funnel. These findings are consistent with aspects of the homology assignments given these organs by earlier authors. There are, however, differences in the expression of these genes relative to expression of their vertebrate homologues. Here I present an analysis of the nucleotide sequences and developmental expression of *Eya*, *Pax1/9*, *Pax2/5/8*, *Pax6*, *Pitx*, and three *Six* genes from *Oikopleura dioica* in the context of the origin of placodes.

MATERIALS AND METHODS

Animal Husbandry

O. dioica were collected off the coast of Oregon and Vancouver Island, B.C.. Animals were cultured in natural seawater and their diet was supplemented with microalgae: *Micromonas pusilla* (strain DW-8), *Nanochloropsis* sp., and/or cryopreserved *Nanochloropsis* sp., *Isochrysis* sp. *(T-iso)*, and *Tetraselmis chui (*strain *PLY429)* (glycerol cryopastes from Brineshrimp Direct). Synchronously fertilized embryo clutches were cultured at 12°C in tissue culture coated Petri dishes.

Cloning and Sequence Analysis

Poly-A mRNA was purified (Micro-FastTrack 2.0 mRNA isolation kit, Invitrogen) from approximately 1600 hatchlings ranging from just hatching to tailshift stage, just before the building of the first house. 3' and 5' RACE cDNA pools were constructed using the SMART RACE cDNA Amplification Kit (Clontech).

Gene fragments were cloned by degenerate PCR. Primers were designed against conserved regions of aligned Genbank sequences from other organisms, and weighted slightly for *O. dioica* codon bias to reduce degeneracy at the 5' ends. Codon bias was cumulatively estimated from several cloned larvacean genes. The following primers were used:

Eya (forward 5' CGAGGCGGAGTNGA(CT)TGGATG 3'; reverse 5' GGTGGCGGAGTA(AGT)AT(AG)TT(CT)TC(AGT)AT 3'); *Pax19*, *Pax37* (forward 5'

GCACGGCTG(C/T)GTN(A/T)(A/G)NAA(A/G)AT 3'); *Pax258*, *Pax6* (forward 5' GCAATTGG(I/C)GG(I/C)AG(T/C)AA(A/G)CC 3'); all *Pax* genes (reverse 5' GTCTGTGTC(C/T)CK KAT (C/T)(C/T)C CCA); *Pitx* (forward 5' GAAGACAACG(AGC)ACNCA(CT)TT(CT)AC 3'; reverse 5' CTTGCTTCAGTNA(AG)(AG)TTNGTCCA 3'); *Six1/2*, *Six3/6a, Six3/6b* (primers a generous gift from J. Willoughby).

Sequence alignments were performed using MegAlign (DNASTAR) and ClustalX software (http://www-igbmc.u-strasbg.fr/BioInfo/ClustalX/Top.html). Gene phylogenies were calculated using ClustalX software and trees were graphically drawn in NJPLOT (http://pbil.univ-lyon1.fr/software/njplot.html) (Perrière and Gouy, 1996). Genbank accession numbers for protein sequences used to generate the phylogenetic trees in Figures 1 B, 2, 4, and 6 are as follows: Eya1 NP_571268 (Dre); Eya Q05201 (Dme), Eya2 AAC98479 (Gga); Eya1 P97767, Eya2 AAB48018, Eya3 P97480, Eya4 Q9Z191 (Mmu); Eya-1ß AAK31355 (Xla); Optix AAF59147, Six4 AAF51640, So Q27350 (Dme); Six2.1 NP_571858, Six3 BAA31752, Six4.1 BAB18513, Six4.2 NP_571793, Six6 AAC27449, Six7 NP_571429, Six8 AAC78389 (Dre); Six3 CAA75380, Six9 CAA09774 (Gga); Six2 BAA11825, Six3a Q62233, Six4 Q61321, Six5 BAA11824, Six9 CAA09775 (Mmu); Six2 CAC86663 (Pdu); Six1 AAK11607 (Rra); Six1 AAF91422, Six3 AAF63242 (Xla); Pitx AAF03901 (Bbe); Ptx AAF57099 (Dme); Pitx2a NP_571050 (Dre); Ptx AAC23684, Pitx2 AAC27322 (Gga); Cart-1 AAB08960 (Hsa); Alx4 O35137, Ptx1 P70314, Pitx2a Q9W5Z2, Pitx3 AAB87380 (Mmu); X-Pitx-1 Q9W751, Pitx2b CAA06697, Pitx3 AAF97592 (Xla); Pax-2 alpha AAC12733, Pax-2 beta AAC12734, Pax37 AAF89581 (Bfl); Pax1 CAB42656 (Bla); Pax19 BAA74829, Pax6 BAB85207 (Cin); Pax6 BAA24025 (Cpy); DjPax-6 BAA75672 (Dja); ey

NP_524628, Gsb AAF47315, Pax2 NP_524633, Pox meso NP_524276, Pox neuro A38153, Prd AAF53160 (Dme); Pax2b NP_571715, Pax3 AAC41253, Pax-6.2 AAC96095, Pax7C AAC41255, Pax9 NP_571373 (Dre); Pax-2/5/8 BAA36346 (Efi); Pax5 AAC34300, Pax8 AAC31810 (Fru); Pax6 BAA23004, Pax-1 AAA65058, Pax-2 BAA88987, Pax3 BAB85652, Pax-5 BAA76951, Pax7 BAA23005, Pax9 P55166 (Gga); Pax19 BAA74831, HrPax258 BAA28833, Pax-37 BAA12289 (Hro); Pax1 NP_006183, Pax-3 P23760, Pax6 A41644, Pax7 CAA65521, Pax9 NP_006185 (Hsa); Pax6 BAB62531, LjPax2/5/8 BAB85112, Pax9 BAB12396 (Lja); Pax-6 AAB40616 (Lop); Pax2 CAA39302, Pax5 NP_032808, Pax-8 CAA67904 (Mmu); Pax19 BAA78380 (Pfl); Pax2/5/8 AAB70245, Pax-6 AAA75363 (Pli); Pax258 CAB96396, Pax-6 CAA71094 (Pma); Pax7 AAL04156 (Pmr). Full species names are given in the figure captions.

In Situ Hybridization

Embryos for *in situ* analysis were treated as in Bassham and Postlethwait (2000), or with the following modification to the protocol: probe hybridization time was shortened from over night to 2.5 to 4 hours.

Microscopy and Photography

Live adults were anaesthetized with MESAB and photographed on a Leica DMLB compound microscope with a SPOT RT Color digital camera (Diagnostic Instruments, Inc.). Fixed embryos mounted in 10% glycerol and live embryos mounted in seawater on Cel-Line teflon-gasketed slides (Erie Scientific Co.) were photographed using the above microscope and camera.

RESULTS

Cloning and Characterization of Molecular Sequences

All gene sequences reported here were obtained by degenerate PCR amplification of fragments using primers that were based on alignment of homologous genes, and by RACE PCR (rapid amplification of cDNA ends). Post-hatching developmental stages were pooled to generate the RACE PCR template cDNA. One *Eya*, one *Pitx*, three *Six*, and four *Pax* genes were identified from this cloning effort. Though the RNA source for making the cDNA pools was purified poly-A mRNA, cloned RACE fragments for some of the genes contained unprocessed introns. Sequence data from these introns, from genomic fragments, and from clones representing possible alternately spliced messages, provided genomic structure information for several of the genes.

Larvacean Eya

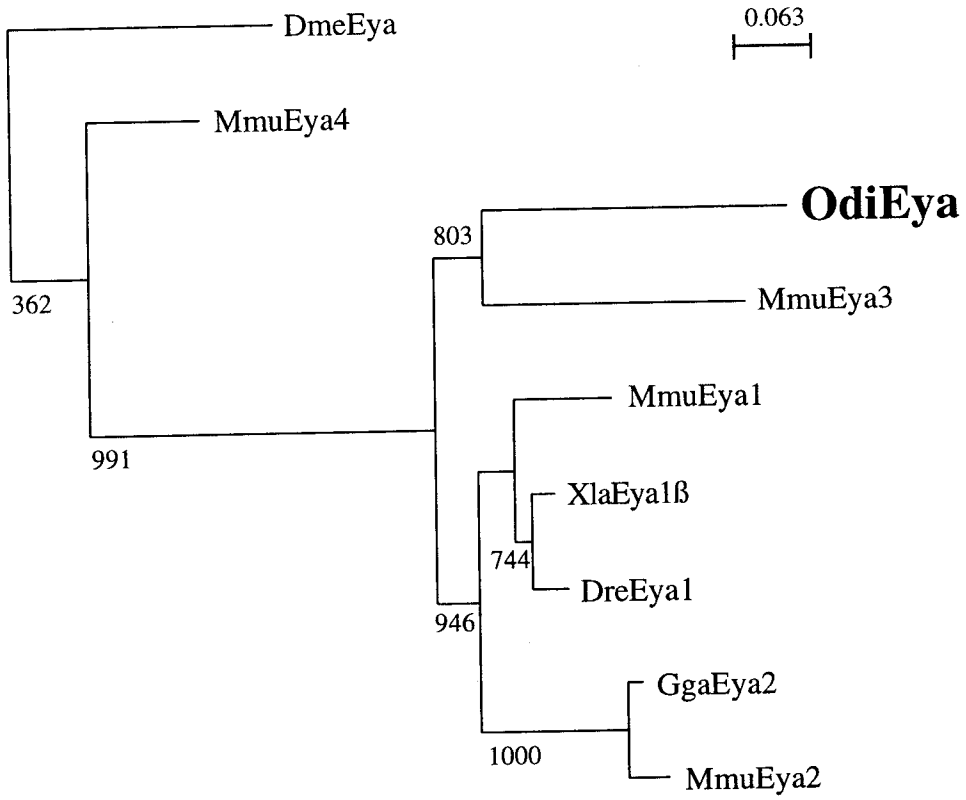
Eya proteins contain a conserved domain, the Eya-domain that is essential for protein-protein interactions, including interaction with Six genes (Pignoni *et al.*, 1997; Ohto *et al.*, 1999). Eya has no recognized DNA-binding ability, and is believed to act as a co-activator in a transcription factor protein complex (reviewed in Hanson, 2001). Although the predicted aminoterminus is probably incomplete in the *O. dioica Eya* cloned in this study, the full length of the conserved Eya-domain is present (Fig. 1, A) allowing a neighbor-joining gene tree to be calculated from aligned Eya-domains of other proteins (Fig. 1, B).

A.

```
ACGCGGGGATATTTTTTCGTATTCTATTTCTGAAAAATTATTTACGCAAATAATAAATGAATCTGTTTTTGTAACAAAAAGAGAAGAGAAAATTATTAGA  100

  R  G  D  I  F  S  Y  S  I  S  E  K  L  F  T  Q  I  I  N  E  S  V  F  V  T  K  R  E  E  K  I  I  R

AGATGGCCTATCCAGGCATTTGGGCGAATTCATATGGCCAGATGTATCAAGGTGCACCTGGAAACAATAACAACTACGGCTCAGACAAGCCAGTATCCTC  200

  R  W  P  I  Q  A  F  G  R  I  H  M  A  R  C  I  K  V  H  L  E  T  I  T  T  T  A  Q  T  S  Q  Y  P

CAAGCGCGCAAACAAGCCAACAAGCTCCGATTCCTGCAAACGGCCAGGCCGAGTATGGAGCGCAATATTACGCTACTCAGGGGCACTATCCAAACTATCA  300

  P  S  A  Q  T  S  Q  Q  A  P  I  P  A  N  G  Q  A  E  Y  G  A  Q  Y  Y  A  T  Q  G  H  Y  P  N  Y  H

CAACAACTATTATGTCAGCAGTACTCAACCTGCTTTAACTGCAACTTCCGGAAGTCAAGTTACGTATCAGCTACAGGAACCGCCAAATGAAGATATTGGA  400

  N  N  Y  Y  V  S  S  T  Q  P  A  L  T  A  T  S  G  S  Q  V  T  Y  Q  L  Q  E  P  P  N  E  D  I  G

GAGATTAAATATACAATTGATGATACTGACAGGCGTGGCGGCAAAAAATCAAAGAAGCGCGACACTGAAAACGGCGACAGCGAGATCGAGCGCGTGTTCA  500

  E  I  K  Y  T  I  D  D  T  D  R  R  G  G  K  K  S  K  K  R  D  T  E  N  G  D  S  E  I  E  R  V  F

TCTGGGACCTGGACGAGACGATCATCATCTTCCACTCGCTGCTCACGTCGTCTTATGCCCAGCGCTACGGAAAGGATACAAACGCGACGTTGCAACTTGG  600

  I  W  D  L  D  E  T  I  I  I  F  H  S  L  L  T  S  S  Y  A  Q  R  Y  G  K  D  T  N  A  T  L  Q  L  G

GCTCAGAATGGAGGAGCTCATCTTCAACTTGGCAGACAACCACTTGTTCTTCAACGACCTGGAAGAGTGCGACCAGGTTCATTGCGACGATGTCGCAGCT  700

  L  R  M  E  E  L  I  F  N  L  A  D  N  H  L  F  F  N  D  L  E  E  C  D  Q  V  H  C  D  D  V  A  A

GATGATAACGGACAAGATCTTGCAAACTATAATTTTCAAACTGATGGATTCAGGACAGCAAATGCTGCAGGCGCGATTTGCCTTGGCCCCGGAGTTCGCG  800

  D  D  N  G  Q  D  L  A  N  Y  N  F  Q  T  D  G  F  R  T  A  N  A  A  G  A  I  C  L  G  P  G  V  R

GAGGAGTTGACTGGATGAGAAAACTCGCTTTTCGATACAGAAGAATCAAAGAGCTCTACGATACTTTCAAAGACAACATTCCAGAGCTGCTGTCAAGTCC  900

  G  G  V  D  W  M  R  K  L  A  F  R  Y  R  R  I  K  E  L  Y  D  T  F  K  D  N  I  P  E  L  L  S  S  P

AGTCGGGGAAGCGTGGAGGCAGCTGAGGACGGAAATCGACCAAAACACGGATAACTGGTGCGCACTCGCGCTCAAGGCGCTTGGGCTCATTCACAATCGA  1000

  V  G  E  A  W  R  Q  L  R  T  E  I  D  Q  N  T  D  N  W  C  A  L  A  L  K  A  L  G  L  I  H  N  R

CCTGGCTGCGTAAATATAATGGTAACAACAACGCAGCTCGTACCAGCAATTGCCAAATGCATGCTCTATGAACTTGGTCCAGTTTTCAACATCGACAATG  1100

  P  G  C  V  N  I  M  V  T  T  T  Q  L  V  P  A  I  A  K  C  M  L  Y  E  L  G  P  V  F  N  I  D  N

TTTACTCCGCGACCAAGATTGGAAAAGAAAGTTGTTTCGAAAGAATTATTCAGCGCTTTGGTAAAAATGTCACATACGTTTGCGTCGGAGATGGAAGAGA  1200

  V  Y  S  A  T  K  I  G  K  E  S  C  F  E  R  I  I  Q  R  F  G  K  N  V  T  Y  V  C  V  G  D  G  R  E

GGAAGAAATTGCTGCGAAAACGCACAATGTCCCATTCTGGCCAATTTCGACACACAACGATCTGATGGCGCTCCACCAAGCTCTTGAACTTGAATATCTG  1300

  E  E  I  A  A  K  T  H  N  V  P  F  W  P  I  S  T  H  N  D  L  M  A  L  H  Q  A  L  E  L  E  Y  L

TAAAAAAAGAAATCTGCACAAAAAGTCTAAAATACCTGCCGAACAGAATGAAAATACAAAATTACAAAAAAAAAAAAAAAAAAAAAAAAAA  1391
```

FIG. 1. Sequence and gene tree of *Eya*. (A) *Eya* nucleotide and predicted protein sequences. The protein predicted from the cloned sequences is probably incomplete at the 5' end. There are no in-frame 5' stop codons or aminoterminal methionines in a good context by start site consensus standards (Kozak, 1996), and this protein sequence is shorter than Eya proteins reported from other animals. The Eya domain is shown boxed. (B) The conserved Eya domain was used to generate this neighbor-joining tree. *O. dioica* Eya groups within the vertebrate clade of Eya proteins, rather than with the fly orthologue, but the larvacean gene lies on a long branch. (Dre *Danio rerio*; Dme *Drosophila melanogaster*; Gga *Gallus gallus*; Mmu *Mus musculus*; Odi *Oikopleura dioica*; Xla *Xenopus laevis*. Only bootstrap values greater than 500 are shown.

B.


DmeEya
0.063
MmuEya4
OdiEya
803
362
MmuEya3
MmuEya1
991
XlaEya1ß
744
DreEya1
946
GgaEya2
1000
MmuEya2

Larvacean Six Genes

Drosophila and vertebrate *Six* genes fall into three clades: *Sine Oculis*, *Six1* and *Six2* form one group, *D-Six3*, *Six3* and *Six6* form the second, and *D-Six4*, *Six4* and *Six5* form the third (Seo *et al.*, 1999). All Six proteins have a Six-domain that mediates protein-protein interactions and a homeodomain that mediates DNA-binding (Pignoni *et al.*, 1997). Six1/Six2 and Six4/Six5 class proteins have been shown to interact with Eya proteins and cause the translocation of the Eya/Six complex into the nucleus (Ohto *et al.*, 1999). Six3/Six6 class proteins, however, may interact with proteins other than Eya. Three larvacean Six genes were cloned in this study. In a gene phylogeny, one larvacean Six (Six1/2) groups with high bootstrap support with Six1 and Six2 class proteins, while the two other larvacean Sixes group strongly with the Six3/Six6 clade (Fig. 2). Within the Six3/Six6 clade *O. dioica* Six3/6a and Six3/6b group most strongly with each other, and their protein sequences share a number of unique changes, implying that they are not singly orthologous to particular *Six3* or *Six6* genes in vertebrates. They likely represent an independent duplication in the larvacean or urochordate lineage.

Figure 3 (A-C) shows complete coding sequences for three larvacean *Six* genes. While in Six1/2 a homeodomain tetrapeptide diagnostic for Six1/Six2 class proteins is conserved (Fig. 3 A), in both the Six3/6 proteins one residue of the motif is altered such that QKTH becomes, in the larvacean sequences, QKSH. Variation at this and other conserved sites within the homeodomain and Six-domain make the larvacean Six3/6 proteins more divergent than any other reported Six3/Six6 class proteins. By comparison,

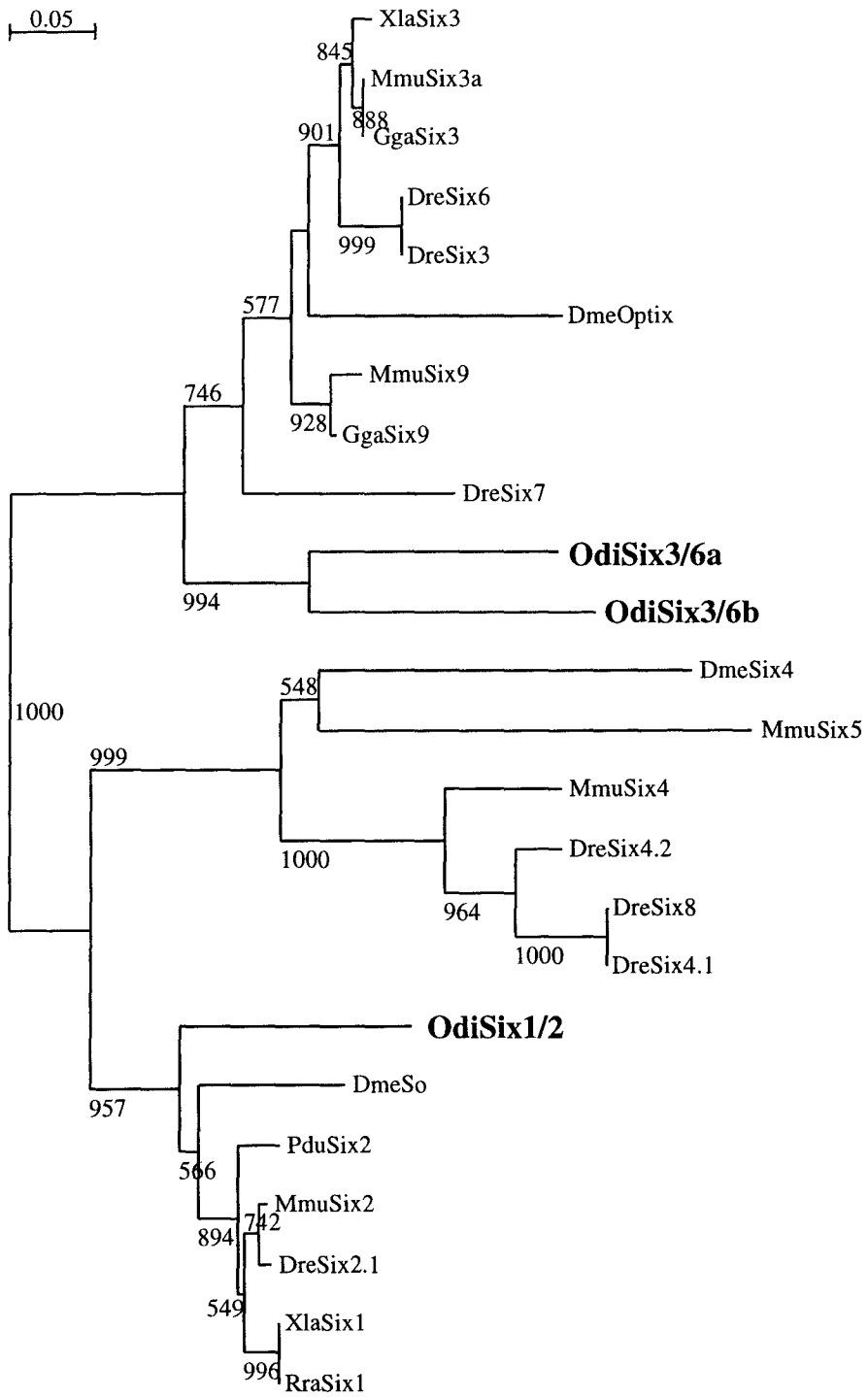
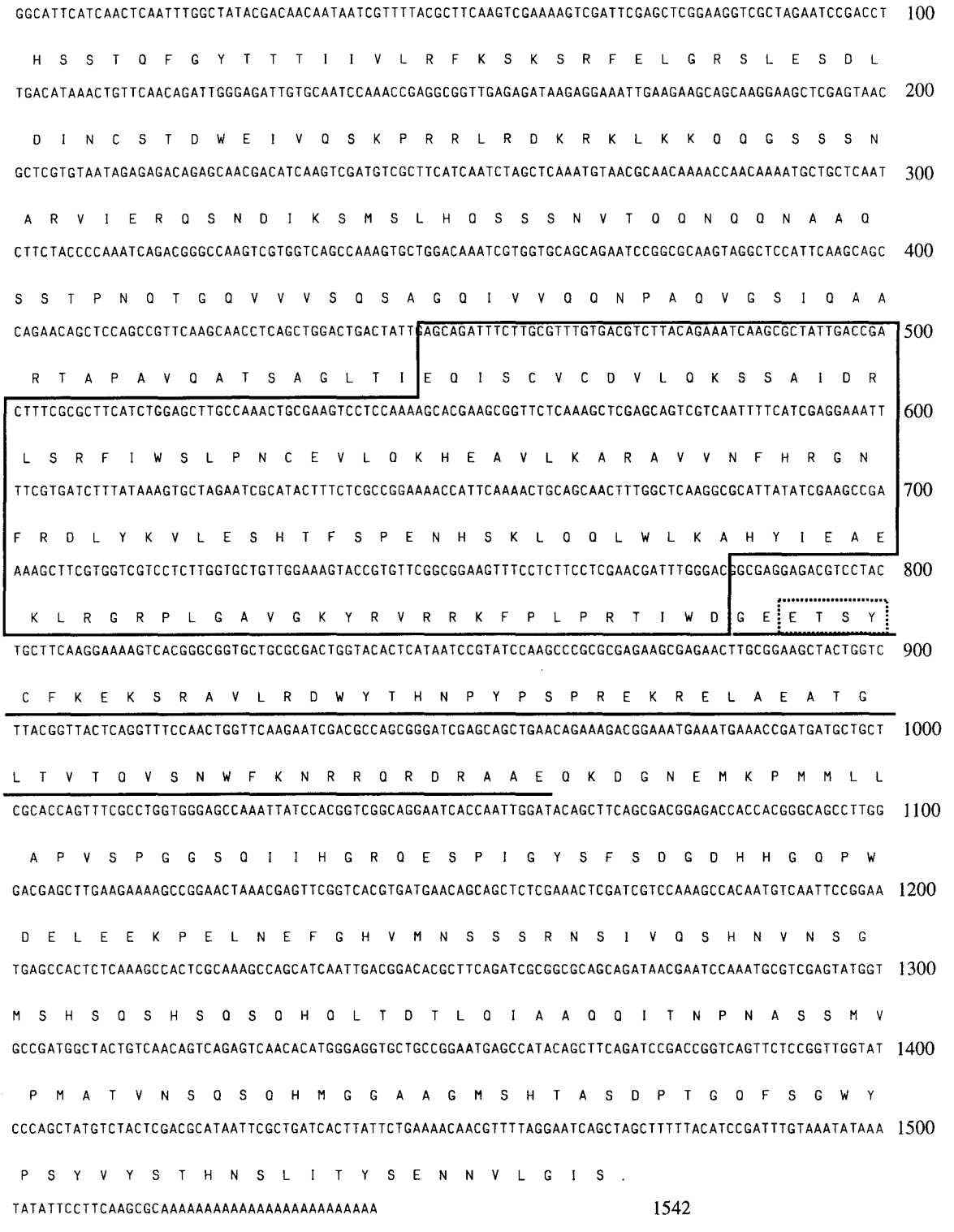


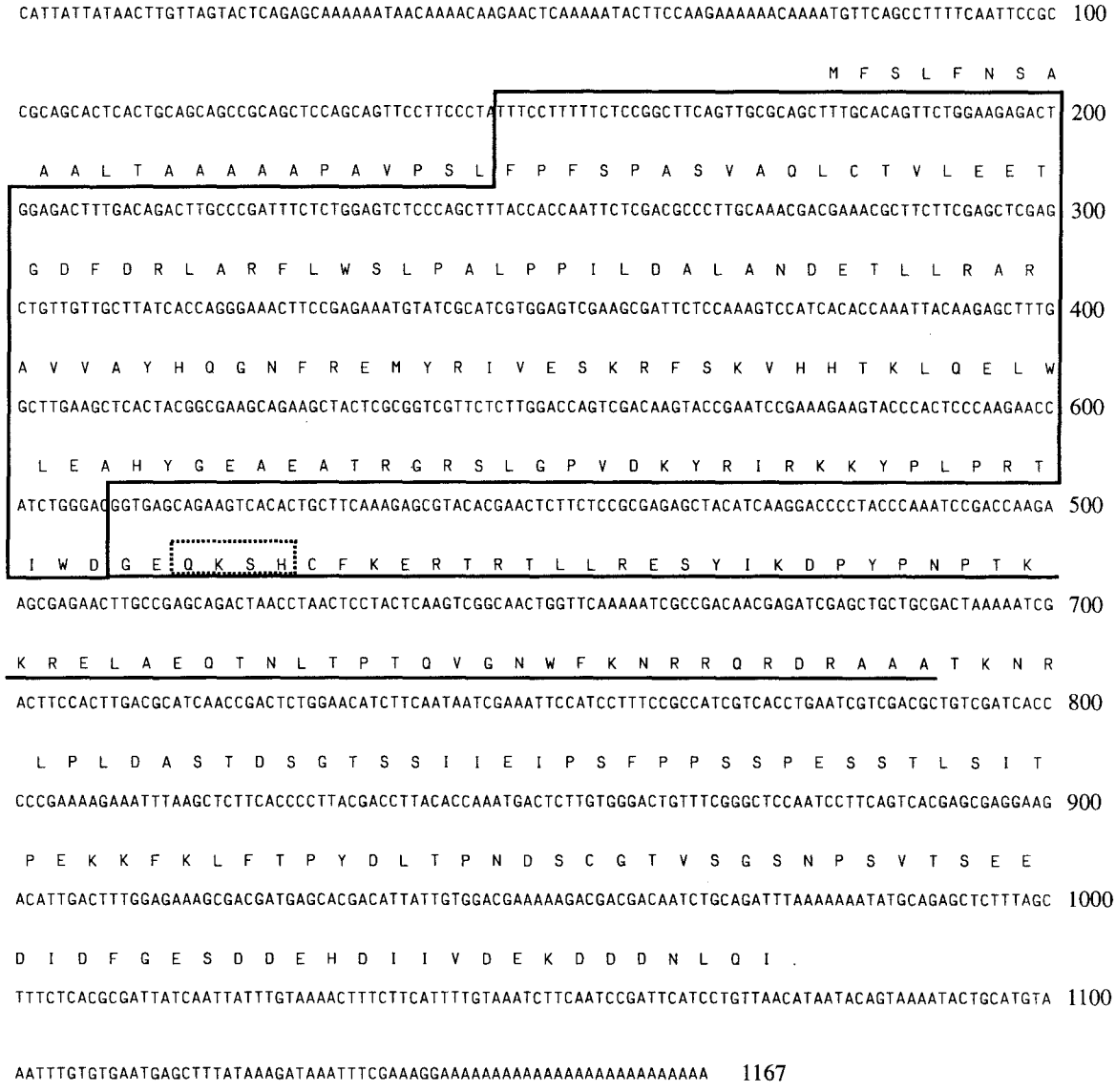
FIG. 2. Gene tree of Six proteins. The conserved Six domain and homeodomain were aligned to generate this tree. A Six1/Six2 clade including *O. dioica* Six1/2, but excluding Six4/Six5 proteins, is supported with a high bootstrap value. *O. dioica* Six3/6a and Six3/6b group with other Six3 and Six6 class proteins, but they group most strongly with each other, implying that they are not singly orthologous to particular *Six3* or *Six6* genes. They likely represent an independent duplication in the larvacean or urochordate lineage. Only bootstrap values greater than 500 are shown. (Dre *Danio rerio*; Dme *Drosophila melanogaster*; Gga *Gallus gallus*; Mmu *Mus musculus*; Pdu *Platynereis dumerilii*; Odi *Oikopleura dioica*; Rra *Rattus rattus*; Xla *Xenopus laevis*).

FIG. 3. DNA and predicted protein sequence of *Six1/2*, *Six3/6a*, and *Six3/6b*. In A-C, the Six domain is boxed and the homeodomain is underlined. Four residues that are diagnostic for each Six subfamily are boxed with a dotted line. (A) *Six1/2*. Because the only 5' ATG upstream of the Six-domain is not in a particularly strong Kozak context and because there are no upstream stop codons, the protein sequence is displayed as incomplete at the 5' end. However, if the starting methionine is not within the sequence shown here, this larvacean Six1/2 protein would have an unusually long aminoterminus compared to other Six1 and Six2 proteins. (B) *Six3/6a*. The predicted sequence is probably complete; there are two upstream, in-frame stop codons. (C) *Six3/6b*. Although there are no upstream stop codons in the short 5' UTR, the protein is probably complete as the predicted start codon is not in a weak Kozak consensus and the Six3/6 proteins of other organisms often have even shorter aminotermini than the larvacean sequence. Arrows mark two intron positions. The first intron was 73bp. Two alleles were discovered in alignments of genomic and cDNA sequences. The allele shown in C had a longer intron (170 versus 164bp) and a three-codon insertion (shaded box) just outside the homeodomain.

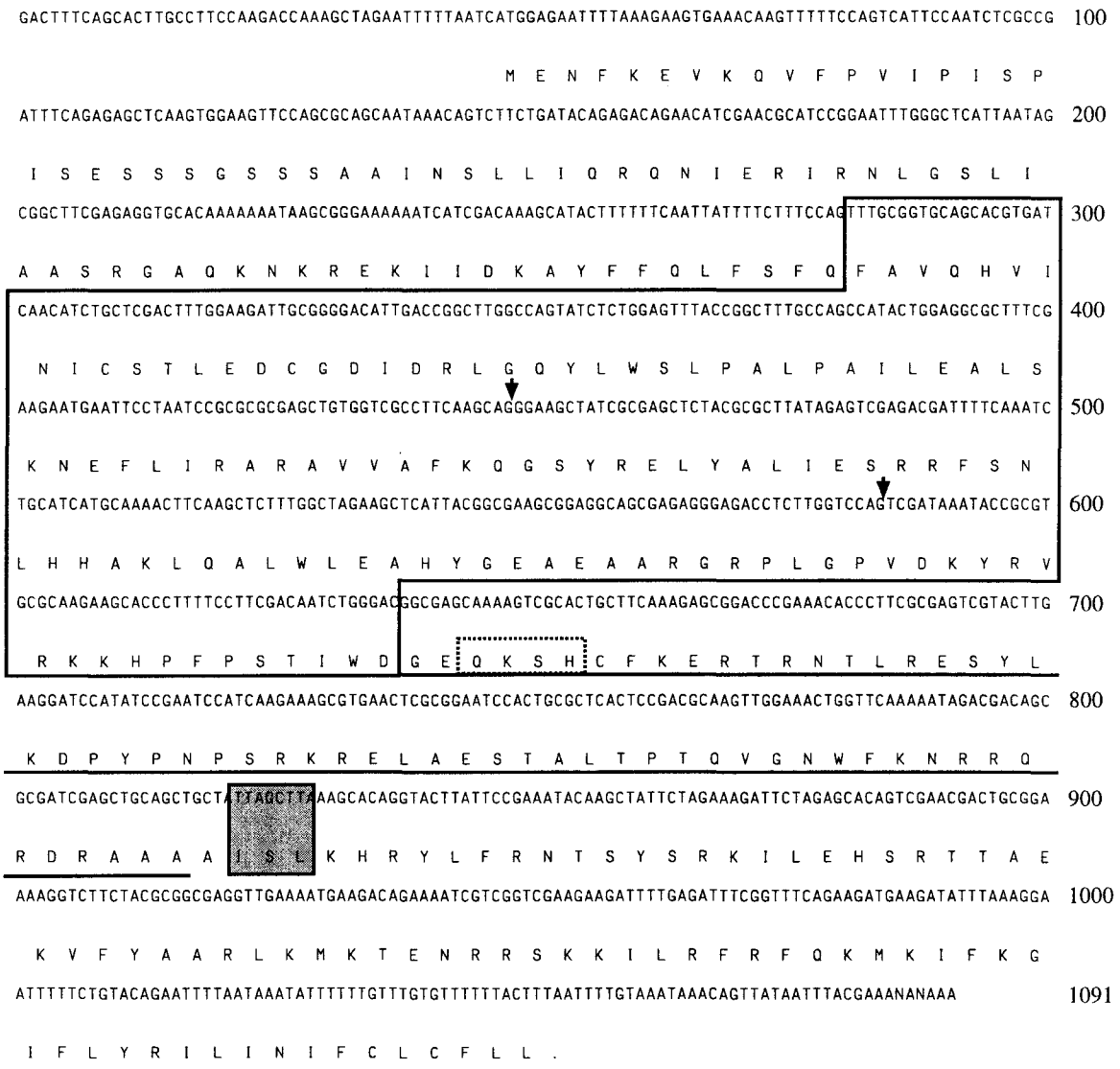
A.


GGCATTCATCAACTCAATTTGGCTATACGACAACAATAATCGTTTTACGCTTCAAGTCGAAAAGTCGATTCGAGCTCGGAAGGTCGCTAGAATCCGACCT 100
H S S T Q F G Y T T T I I V L R F K S K S R F E L G R S L E S D L
TGACATAAACTGTTCAACAGATTGGGAGATTGTGCAATCCAAACCGAGGCGGTTGAGAGATAAGAGGAAATTGAAGAAGCAGCAAGGAAGCTCGAGTAAC 200
D I N C S T D W E I V Q S K P R R L R D K R K L K K Q Q G S S S N
GCTCGTGTAATAGAGAGACAGAGCAACGACATCAAGTCGATGTCGCTTCATCAATCTAGCTCAAATGTAACGCAACAAAACCAACAAAATGCTGCTCAAT 300
A R V I E R Q S N D I K S M S L H Q S S S N V T Q Q N Q Q N A A Q
CTTCTACCCCAAATCAGACGGGCCAAGTCGTGGTCAGCCAAAGTGCTGGACAAATCGTGGTGCAGCAGAATCCGGCGCAAGTAGGCTCCATTCAAGCAGC 400
S S T P N Q T G Q V V V S Q S A G Q I V V Q Q N P A Q V G S I Q A A
CAGAACAGCTCCAGCCGTTCAAGCAACCTCAGCTGGACTGACTATTCAGCAGATTTCTTGCGTTTGTGACGTCTTACAGAAATCAAGCGCTATTGACCGA 500
R T A P A V Q A T S A G L T I E Q I S C V C D V L Q K S S A I D R
CTTTCGCGCTTCATCTGGAGCTTGCCAAACTGCGAAGTCCTCCAAAAGCACGAAGCGGTTCTCAAAGCTCGAGCAGTCGTCAATTTTCATCGAGGAAATT 600
L S R F I W S L P N C E V L Q K H E A V L K A R A V V N F H R G N
TTCGTGATCTTTATAAAGTGCTAGAATCGCATACTTTCTCGCCGGAAAACCATTCAAAACTGCAGCAACTTTGGCTCAAGGCGCATTATATCGAAGCCGA 700
F R D L Y K V L E S H T F S P E N H S K L Q Q L W L K A H Y I E A E
AAAGCTTCGTGGTCGTCCTCTTGGTGCTGTTGGAAAGTACCGTGTTCGGCGGAAGTTTCCTCTTCCTCGAACGATTTGGGACGGCGAGGAGACGTCCTAC 800
K L R G R P L G A V G K Y R V R R K F P L P R T I W D G E E T S Y
TGCTTCAAGGAAAAGTCACGGGCGGTGCTGCGCGACTGGTACACTCATAATCCGTATCCAAGCCCGCGCGAGAAGCGAGAACTTGCGGAAGCTACTGGTC 900
C F K E K S R A V L R D W Y T H N P Y P S P R E K R E L A E A T G
TTACGGTTACTCAGGTTTCCAACTGGTTCAAGAATCGACGCCAGCGGGATCGAGCAGCTGAACAGAAAGACGGAAATGAAATGAAACCGATGATGCTGCT 1000
L T V T Q V S N W F K N R R Q R D R A A E Q K D G N E M K P M M L L
CGCACCAGTTTCGCCTGGTGGGAGCCAAATTATCCACGGTCGGCAGGAATCACCAATTGGATACAGCTTCAGCGACGGAGACCACCACGGGCAGCCTTGG 1100
A P V S P G G S Q I I H G R Q E S P I G Y S F S D G D H H G Q P W
GACGAGCTTGAAGAAAAGCCGGAACTAAACGAGTTCGGTCACGTGATGAACAGCAGCTCTCGAAACTCGATCGTCCAAAGCCACAATGTCAATTCCGGAA 1200
D E L E E K P E L N E F G H V M N S S S R N S I V Q S H N V N S G
TGAGCCACTCTCAAAGCCACTCGCAAAGCCAGCATCAATTGACGGACACGCTTCAGATCGCGGCGCAGCAGATAACGAATCCAAATGCGTCGAGTATGGT 1300
M S H S Q S H S Q S Q H Q L T D T L Q I A A Q Q I T N P N A S S M V
GCCGATGGCTACTGTCAACAGTCAGAGTCAACACATGGGAGGTGCTGCCGGAATGAGCCATACAGCTTCAGATCCGACCGGTCAGTTCTCCGGTTGGTAT 1400
P M A T V N S Q S Q H M G G A A G M S H T A S D P T G Q F S G W Y
CCCAGCTATGTCTACTCGACGCATAATTCGCTGATCACTTATTCTGAAAACAACGTTTTAGGAATCAGCTAGCTTTTTACATCCGATTTGTAAATATAAA 1500
P S Y V Y S T H N S L I T Y S E N N V L G I S .
TATATTCCTTCAAGCGCAAAAAAAAAAAAAAAAAAAAAAAA 1542

B.


CATTATTATAACTTGTTAGTACTCAGAGCAAAAAATAACAAAACAAGAACTCAAAAATACTTCCAAGAAAAAACAAAATGTTCAGCCTTTTCAATTCCGC 100
M F S L F N S A
CGCAGCACTCACTGCAGCAGCCGCAGCTCCAGCAGTTCCTTCCCTATTTCCTTTTTCTCCGGCTTCAGTTGCGCAGCTTTGCACAGTTCTGGAAGAGACT 200
A A L T A A A A A A P A V P S L F P F S P A S V A Q L C T V L E E T
GGAGACTTTGACAGACTTGCCCGATTTCTCTGGAGTCTCCCAGCTTTACCACCAATTCTCGACGCCCTTGCAAACGACGAAACGCTTCTTCGAGCTCGAG 300
G D F D R L A R F L W S L P A L P P I L D A L A N D E T L L R A R
CTGTTGTTGCTTATCACCAGGGAAACTTCCGAGAAATGTATCGCATCGTGGAGTCGAAGCGATTCTCCAAAGTCCATCACACCAAATTACAAGAGCTTTG 400
A V V A Y H Q G N F R E M Y R I V E S K R F S K V H H T K L Q E L W
GCTTGAAGCTCACTACGGCGAAGCAGAAGCTACTCGCGGTCGTTCTCTTGGACCAGTCGACAAGTACCGAATCCGAAAGAAGTACCCACTCCCAAGAACC 600
L E A H Y G E A E A T R G R S L G P V D K Y R I R K K Y P L P R T
ATCTGGGACGGTGAGCAGAAGTCACACTGCTTCAAAGAGCGTACACGAACTCTTCTCCGCGAGAGCTACATCAAGGACCCCTACCCAAATCCGACCAAGA 500
I W D G E Q K S H C F K E R T R T L L R E S Y I K D P Y P N P T K
AGCGAGAACTTGCCGAGCAGACTAACCTAACTCCTACTCAAGTCGGCAACTGGTTCAAAAATCGCCGACAACGAGATCGAGCTGCTGCGACTAAAAATCG 700
K R E L A E Q T N L T P T Q V G N W F K N R R Q R D R A A A T K N R
ACTTCCACTTGACGCATCAACCGACTCTGGAACATCTTCAATAATCGAAATTCCATCCTTTCCGCCATCGTCACCTGAATCGTCGACGCTGTCGATCACC 800
L P L D A S T D S G T S S I I E I P S F P P S S P E S S T L S I T
CCCGAAAAGAAATTTAAGCTCTTCACCCCTTACGACCTTACACCAAATGACTCTTGTGGGACTGTTTCGGGCTCCAATCCTTCAGTCACGAGCGAGGAAG 900
P E K K F K L F T P Y D L T P N D S C G T V S G S N P S V T S E E
ACATTGACTTTGGAGAAAGCGACGATGAGCACGACATTATTGTGGACGAAAAAGACGACGACAATCTGCAGATTTAAAAAAATATGCAGAGCTCTTTAGC 1000
D I D F G E S D D E H D I I V D E K D D D N L Q I .
TTTCTCACGCGATTATCAATTATTTGTAAAACTTTCTTCATTTTGTAAATCTTCAATCCGATTCATCCTGTTAACATAATACAGTAAAATACTGCATGTA 1100
AATTTGTGTGAATGAGCTTTATAAAGATAAATTTCGAAAGGAAAAAAAAAAAAAAAAAAAAAAAA 1167

C.


GACTTTCAGCACTTGCCTTCCAAGACCAAAGCTAGAATTTTTAATCATGGAGAATTTTAAAGAAGTGAAACAAGTTTTTCCAGTCATTCCAATCTCGCCG 100
M E N F K E V K Q V F P V I P I S P
ATTTCAGAGAGCTCAAGTGGAAGTTCCAGCGCAGCAATAAACAGTCTTCTGATACAGAGACAGAACATCGAACGCATCCGGAATTTGGGCTCATTAATAG 200
I S E S S S G S S S A A I N S L L I Q R Q N I E R I R N L G S L I
CGGCTTCGAGAGGTGCACAAAAAAAATAAGCGGGAAAAAATCATCGACAAAGCATACTTTTTTCAATTATTTTCTTTCCAGTTTGCGGTGCAGCACGTGAT 300
A A S R G A Q K N K R E K I I D K A Y F F Q L F S F Q F A V Q H V I
CAACATCTGCTCGACTTTGGAAGATTGCGGGGACATTGACCGGCTTGGCCAGTATCTCTGGAGTTTACCGGCTTTGCCAGCCATACTGGAGGCGCTTTCG 400
N I C S T L E D C G D I D R L G Q Y L W S L P A L P A I L E A L S
AAGAATGAATTCCTAATCCGCGCGCGAGCTGTGGTCGCCTTCAAGCAGGGAAGCTATCGCGAGCTCTACGCGCTTATAGAGTCGAGACGATTTTCAAATC 500
K N E F L I R A R A V V A F K Q G S Y R E L Y A L I E S R R F S N
TGCATCATGCAAAACTTCAAGCTCTTTGGCTAGAAGCTCATTACGGCGAAGCGGAGGCAGCGAGAGGGAGACCTCTTGGTCCAGTCGATAAATACCGCGT 600
L H H A K L Q A L W L E A H Y G E A E A A R G R P L G P V D K Y R V
GCGCAAGAAGCACCCTTTTCCTTCGACAATCTGGGACGGCGAGCAAAAGTCGCACTGCTTCAAAGAGCGGACCCGAAACACCCTTCGCGAGTCGTACTTG 700
R K K H P F P S T I W D G E Q K S H C F K E R T R N T L R E S Y L
AAGGATCCATATCCGAATCCATCAAGAAAGCGTGAACTCGCGGAATCCACTGCGCTCACTCCGACGCAAGTTGGAAACTGGTTCAAAAATAGACGACAGC 800
K D P Y P N P S R K R E L A E S T A L T P T Q V G N W F K N R R Q
GCGATCGAGCTGCAGCTGCTATTAGCTTAAAGCACAGGTACTTATTCCGAAATACAAGCTATTCTAGAAAGATTCTAGAGCACAGTCGAACGACTGCGGA 900
R D R A A A A I S L K H R Y L F R N T S Y S R K I L E H S R T T A E
AAAGGTCTTCTACGCGGCGAGGTTGAAAATGAAGACAGAAAATCGTCGGTCGAAGAAGATTTTGAGATTTCGGTTTCAGAAGATGAAGATATTTAAAGGA 1000
K V F Y A A R L K M K T E N R R S K K I L R F R F Q K M K I F K G
ATTTTTCTGTACAGAATTTTAATAAATATTTTTTGTTTGTGTTTTTTACTTTAATTTTGTAAATAAACAGTTATAATTTACGAAANANAAA 1091
I F L Y R I L I N I F C L C F L L .

the homeodomain of larvacean Six1/2 is much better conserved, differing only in residues that are variable among the other Six1/Six2 class proteins.

Larvacean Pitx

The *Pitx* genes belong to the *Aristaless*-related subfamily within the *Paired*-class homeobox gene superfamily. In addition to the DNA-binding homeodomain, Aristaless-related proteins share a carboxyterminal Aristaless domain (reviewed in Meijlink, 1999). *O. dioica* Pitx clearly groups with other Pitx sequences in a neighbor-joining tree of aligned homeodomains from Pitx and other Aristaless-related proteins (Fig. 4). However, no discernable Aristaless domain could be found in the *O. dioica* Pitx cloned in this study. Evidence of alternative splicing in this larvacean gene opens the possibility that other isoforms may exist that encode an Aristaless-domain (Fig. 5).

Pitx clones were highly variable. Not only were unprocessed introns present in some clones, but also potential splice variants and polymorphisms affecting protein length and sequence. In addition, it is possible that the larvacean *Pitx* gene uses two different promoters (Fig. 5, B), and at one splice site is a coding sequence length difference that could represent a polymorphism rather than alternate splicing (Fig. 5, A).

Larvacean Pax Genes

Pax proteins fall into four subfamilies. All Pax proteins have a DNA-binding Paired domain. The presence or absence of 3 kinds of other domains or motifs characterizes the proteins of each subfamily. Pax1/Pax9 class proteins have a characteristic octapeptide sequence of unknown function, Pax2/Pax5/Pax8 class proteins

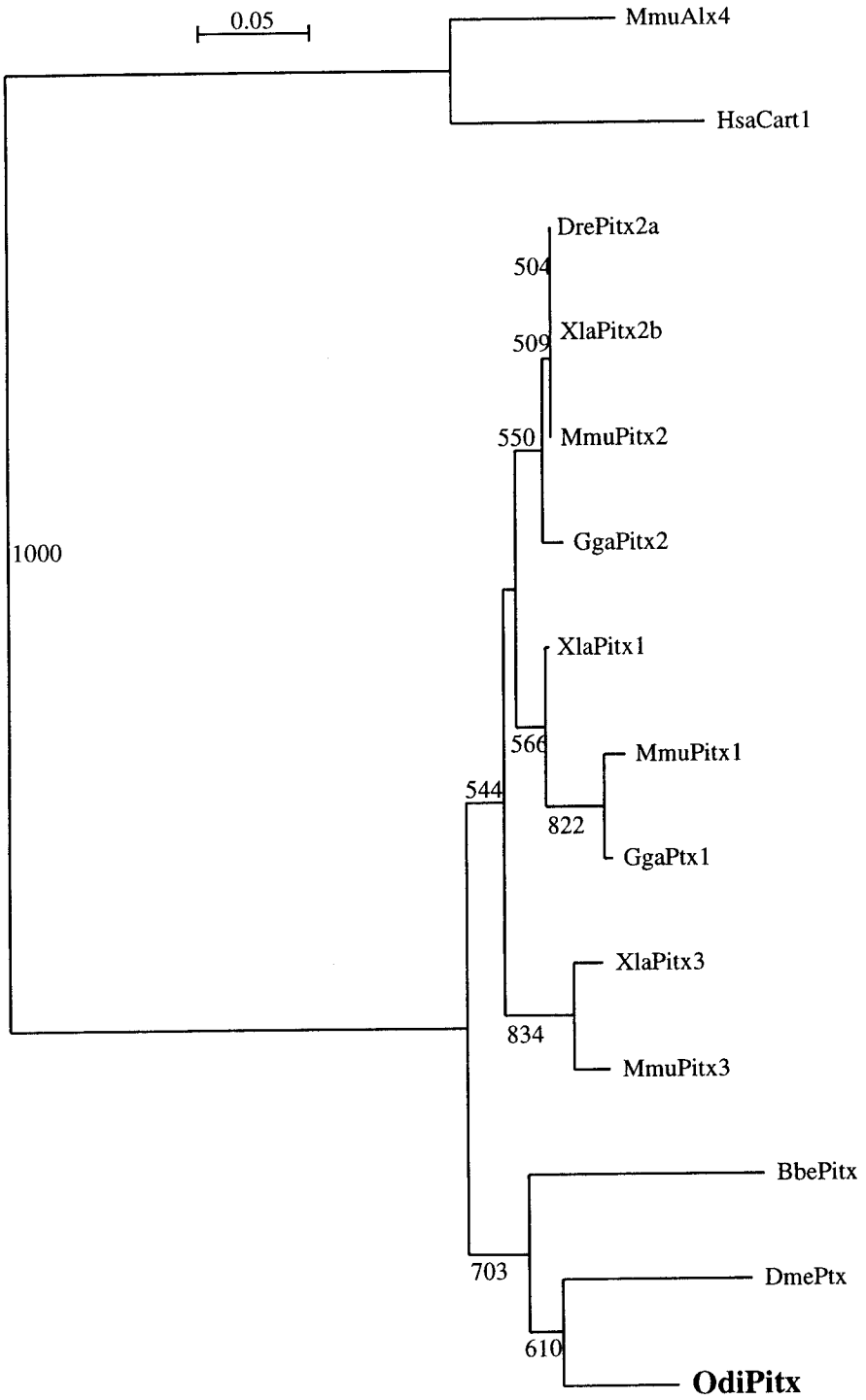
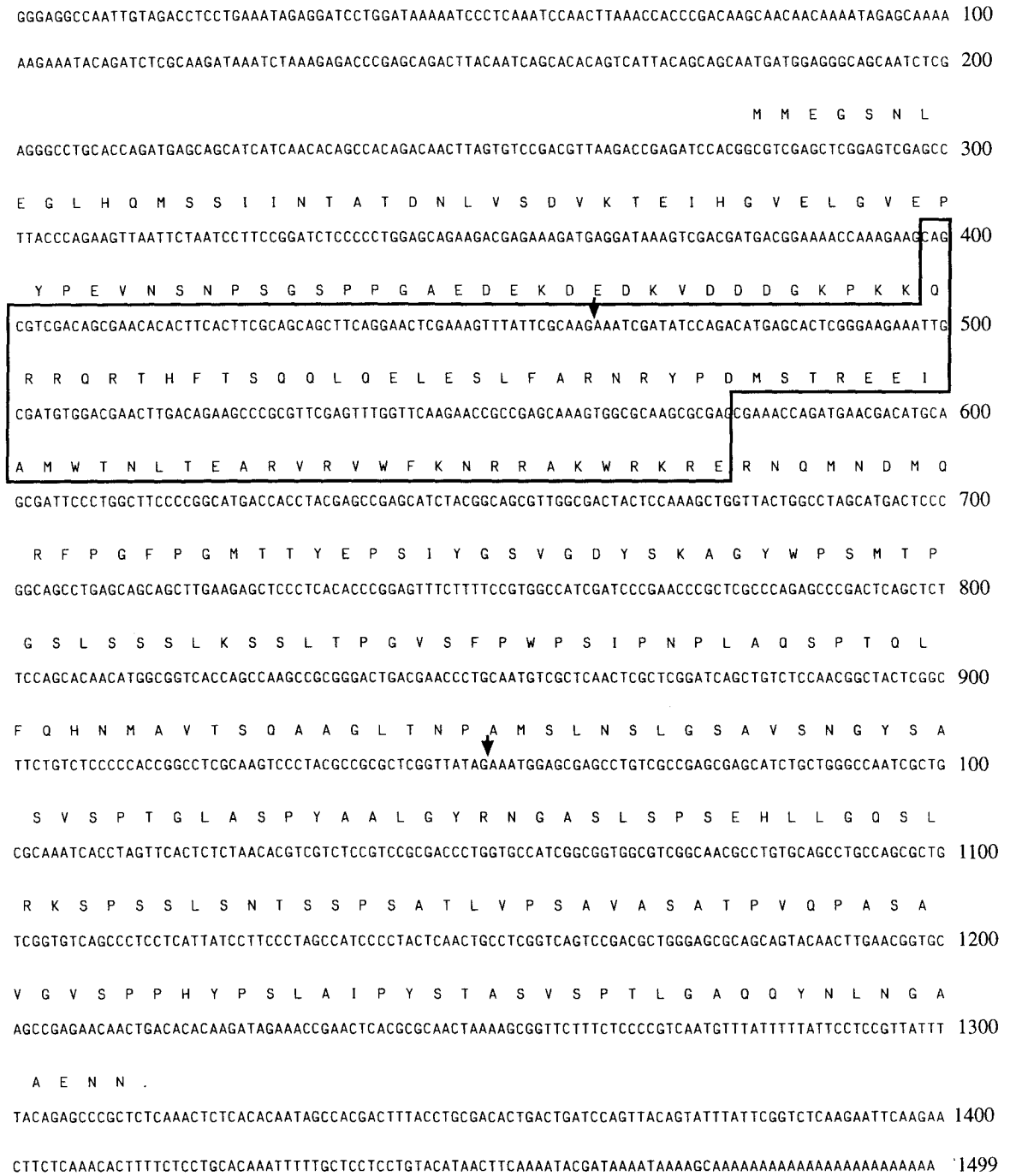


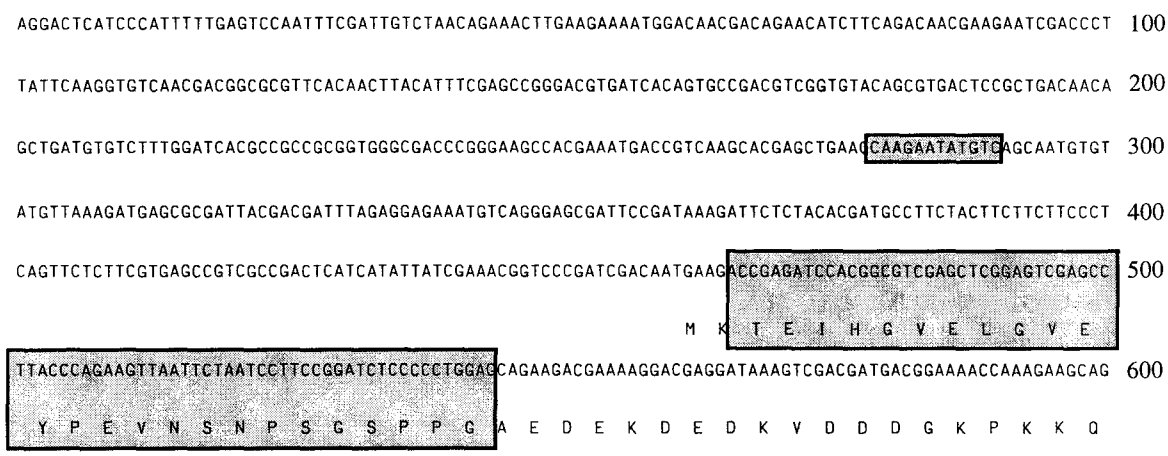
FIG. 4. Gene tree of Pitx proteins. The homeodomains and a short region upstream to it were aligned to generate this neighbor-joining tree. *O. dioica* Pitx falls within the Pitx clade to the exclusion of Aristaless-related proteins (Alx4 and Cart1) that are not part of the Pitx subfamily. Within the Pitx clade, the *O. dioica* protein groups with cephalochordate (Bbe) and fly (Dme) homologues. Only bootstrap values over 500 are shown. (Bbe *Branchiostoma belcheri*; Dre *Danio rerio*; Dme *Drosophila melanogaster*; Hsa *Homo sapiens*; Gga *Gallus gallus*; Mmu *Mus musculus*; Odi *Oikopleura dioica*; Xla *Xenopus laevis*).

FIG. 5. DNA and predicted protein sequence of putative *Pitx* isoforms. Two different kinds of 5' ends were cloned from RACE products; these may result from alternate splicing or from use of an alternate promoter. It is not known which potential 5' and 3' characteristics occur together in the same transcript. The cDNA and protein shown here is a composite of overlapping PCR products, and only the longest putative product is shown in A. (A) A putative Pitx cDNA and protein product. There are two in-frame, upstream stop codons. Two intron positions are marked with arrows and correspond to introns of 134 and perhaps 92bp. This second intron site is unusual in that some clones contained an additional 7 nucleotides of coding sequence at the predicted splice junction. This generates a frame-shift and a stop codon almost immediately after the splice site. Such variation in splice site position could be a polymorphism. The homeodomain is boxed. (B) An alternate 5' end present in several *Pitx* RACE clones. A short and long indel (shaded boxes) were seen in a single clone, generating the predicted 5' protein difference seen in C. (C) A comparison of two predicted 5' ends. 1) is as in B. 2) is created by the deletions boxed in B.

A.


GGGAGGCCAATTGTAGACCTCCTGAAATAGAGGATCCTGGATAAAAATCCCTCAAATCCAACTTAAACCACCCGACAAGCAACAACAAAATAGAGCAAAA 100
AAGAAATACAGATCTCGCAAGATAAATCTAAAGAGACCCGAGCAGACTTACAATCAGCACACAGTCATTACAGCAGCAATGATGGAGGGCAGCAATCTCG 200
M M E G S N L
AGGGCCTGCACCAGATGAGCAGCATCATCAACACAGCCACAGACAACTTAGTGTCCGACGTTAAGACCGAGATCCACGGCGTCGAGCTCGGAGTCGAGCC 300
E G L H Q M S S I I N T A T D N L V S D V K T E I H G V E L G V E P
TTACCCAGAAGTTAATTCTAATCCTTCCGGATCTCCCCCTGGAGCAGAAGACGAGAAAGATGAGGATAAAGTCGACGATGACGGAAAACCAAAGAAGCCAG 400
Y P E V N S N P S G S P P G A E D E K D E D K V D D D G K P K K Q
CGTCGACAGCGAACACACTTCACTTCGCAGCAGCTTCAGGAACTCGAAAGTTTATTCGCAAGAAATCGATATCCAGACATGAGCACTCGGGAAGAAATTG 500
R R Q R T H F T S Q Q L Q E L E S L F A R N R Y P D M S T R E E I
CGATGTGGACGAACTTGACAGAAGCCCGCGTTCGAGTTTGGTTCAAGAACCGCCGAGCAAAGTGGCGCAAGCGCGAGCGAAACCAGATGAACGACATGCA 600
A M W T N L T E A R V R V W F K N R R A K W R K R E R N Q M N D M Q
GCGATTCCCTGGCTTCCCCGGCATGACCACCTACGAGCCGAGCATCTACGGCAGCGTTGGCGACTACTCCAAAGCTGGTTACTGGCCTAGCATGACTCCC 700
R F P G F P G M T T Y E P S I Y G S V G D Y S K A G Y W P S M T P
GGCAGCCTGAGCAGCAGCTTGAAGAGCTCCCTCACACCCGGAGTTTCTTTTCCGTGGCCATCGATCCCGAACCCGCTCGCCCAGAGCCCGACTCAGCTCT 800
G S L S S S L K S S L T P G V S F P W P S I P N P L A Q S P T Q L
TCCAGCACAACATGGCGGTCACCAGCCAAGCCGCGGGACTGACGAACCCTGCAATGTCGCTCAACTCGCTCGGATCAGCTGTCTCCAACGGCTACTCGGC 900
F Q H N M A V T S Q A A G L T N P A M S L N S L G S A V S N G Y S A
TTCTGTCTCCCCCACCGGCCTCGCAAGTCCCTACGCCGCGCTCGGTTATAGAAATGGAGCGAGCCTGTCGCCGAGCGAGCATCTGCTGGGCCAATCGCTG 100
S V S P T G L A S P Y A A L G Y R N G A S L S P S E H L L G Q S L
CGCAAATCACCTAGTTCACTCTCTAACACGTCGTCTCCGTCCGCGACCCTGGTGCCATCGGCGGTGGCGTCGGCAACGCCTGTGCAGCCTGCCAGCGCTG 1100
R K S P S S L S N T S S P S A T L V P S A V A S A T P V Q P A S A
TCGGTGTCAGCCCTCCTCATTATCCTTCCCTAGCCATCCCCTACTCAACTGCCTCGGTCAGTCCGACGCTGGGAGCGCAGCAGTACAACTTGAACGGTGC 1200
V G V S P P H Y P S L A I P Y S T A S V S P T L G A Q Q Y N L N G A
AGCCGAGAACAACTGACACACAAGATAGAAACCGAACTCACGCGCAACTAAAAGCGGTTCTTTCTCCCCGTCAATGTTTATTTTTATTCCTCCGTTATTT 1300
A E N N .
TACAGAGCCCGCTCTCAAACTCTCACACAATAGCCACGACTTTACCTGCGACACTGACTGATCCAGTTACAGTATTTATTCGGTCTCAAGAATTCAAGAA 1400
CTTCTCAAACACTTTTCTCCTGCACAAATTTTTGCTCCTCCTGTACATAACTTCAAAATACGATAAAATAAAAGCAAAAAAAAAAAAAAAAAAAAAAAA 1499

B.



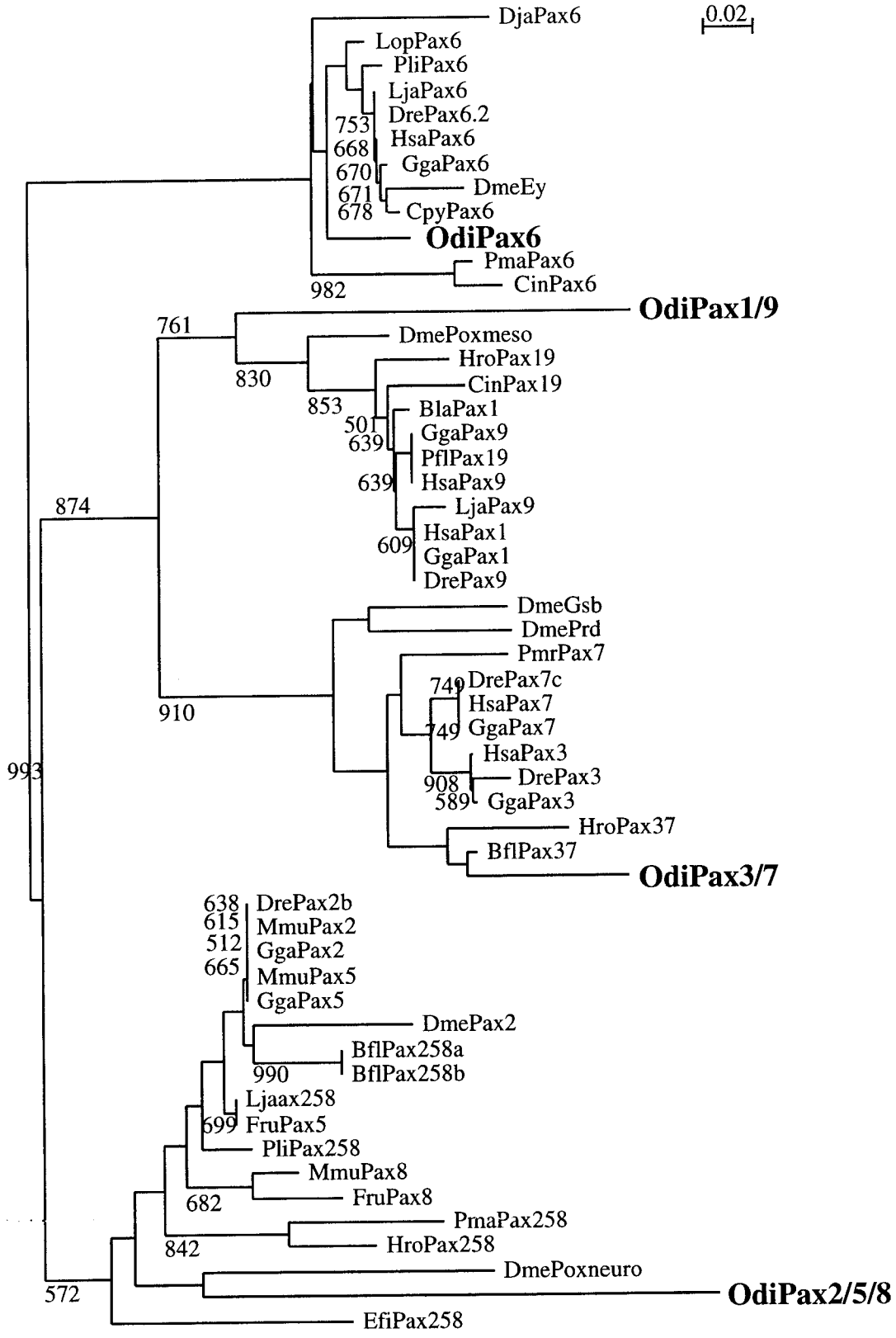
C.

1. MKTEIHGVELGVEPYPEVNSNPSGSPPGAEDEKDED
2. MPSTSSSLSSLREPSPTHHIIETVPIDNEAEDEKDED

have the octapeptide and a partial homeodomain, Pax3/Pax7 class proteins have both the octapeptide and a complete homeodomain, and Pax4/Pax6 class proteins have a complete homeodomain but lack the octapeptide. Four larvacean *Pax* genes were cloned in this study and each one groups with one of the four *Pax* subfamilies in a neighbor-joining tree based on aligned paired domains (Fig. 6).

The proteins deduced from *O. dioica Pax1/9* and *Pax2/5/8* clones (Fig. 7; Fig. 8) are divergent relative to orthologous vertebrate proteins (Fig. 6). Both proteins appear unalignable with orthologous sequences outside the paired domain itself; no indication of an octapeptide was found in either larvacean protein, and the Pax2/5/8 appears also to lack the partial homeodomain. Predicted Pax2/5/8 proteins from ascidians *Halocynthia roretzi* and *Phallusia mammilata* have an octapeptide sequence although it is relatively divergent in *P. mammilata*. Like *O. dioica* Pax2/5/8, however, the ascidian sequences align very poorly with the partial homeodomain found in orthologues, so the degradation of these regions may be a common trend in Urochordates. Nonetheless, Wada *et al.* (1998) tentatively identify what they describe as an eroded remnant of a homeodomain. In vertebrates, multiple splice variants have been found for *Pax2* and *Pax8*. There are at least nine isoforms of a *Xenopus Pax2* gene (Heller and Brändli, 1997), generated by alternative splicing. These include isoforms that delete exons containing the octapeptide, the partial homeodomain, or both. One *O. dioica Pax2/5/8* clone contains a 42bp insertion that could be an alternately spliced exon; the insertion changes the reading frame, creates a predicted protein with a different carboxyterminal sequence and shortens the protein by 78 amino acids (Fig. 8, B). In addition, the *O. dioica* consensus is

FIG. 6. Gene tree of Pax proteins. Paired domain protein sequence was aligned to generate this neighbor-joining tree. Larvacean proteins group with each of four Pax subfamily clades. Larvacean Pax1/9 and Pax2/5/8 paired boxes are very divergent relative to orthologous genes and lie on long branches. Only bootstrap values over 500 are shown. (larvacean: Odi *Oikopleura dioica*; ascidians: Cin *Ciona intestinalis*; Hro *Halocynthia roretzi*; Pma *Phallusia mammilata*; cephalochordates: Bfl *Branchiostoma floridae*; Bla *Branchiostoma lanciolatum*; hemichordates and echinoderms: Pfl *Ptychodera flava*; Pli *Paracentrotus lividus*; sponge: Efl *Ephydatia fluviatilis*; protostomes: Dga *Dugesia japonica*; Dme *Drosophila melanogaster*; Lop *Loligo opalescens*; vertebrates: Cpy *Cynops pyrrhogaster*; Dre *Danio rerio*; Fru *Fugu rubripes*; Gga *Gallus gallus*; Hsa *Homo sapiens*; Lja *Lethenteron japonicum*; Mmu *Mus musculus*; Pmr *Petromyzon marinus*; Xla *Xenopus laevis*).


DjaPax6
0.02
LopPax6
PliPax6
LjaPax6
DrePax6.2
753
HsaPax6
668
GgaPax6
670
DmeEy
671
678
CpyPax6
OdiPax6
PmaPax6
CinPax6
982
OdiPax1/9
761
DmePoxmeso
HroPax19
830
CinPax19
853
BlaPax1
501
GgaPax9
639
PflPax19
639
HsaPax9
874
LjaPax9
HsaPax1
609
GgaPax1
DrePax9
DmeGsb
DmePrd
PmrPax7
749
DrePax7c
HsaPax7
910
749
GgaPax7
HsaPax3
993
908
DrePax3
589
GgaPax3
HroPax37
BflPax37
OdiPax3/7
638
DrePax2b
615
MmuPax2
512
GgaPax2
665
MmuPax5
GgaPax5
DmePax2
BflPax258a
990
BflPax258b
Ljaax258
699
FruPax5
PliPax258
MmuPax8
682
FruPax8
PmaPax258
842
HroPax258
DmePoxneuro
572
OdiPax2/5/8
EfiPax258

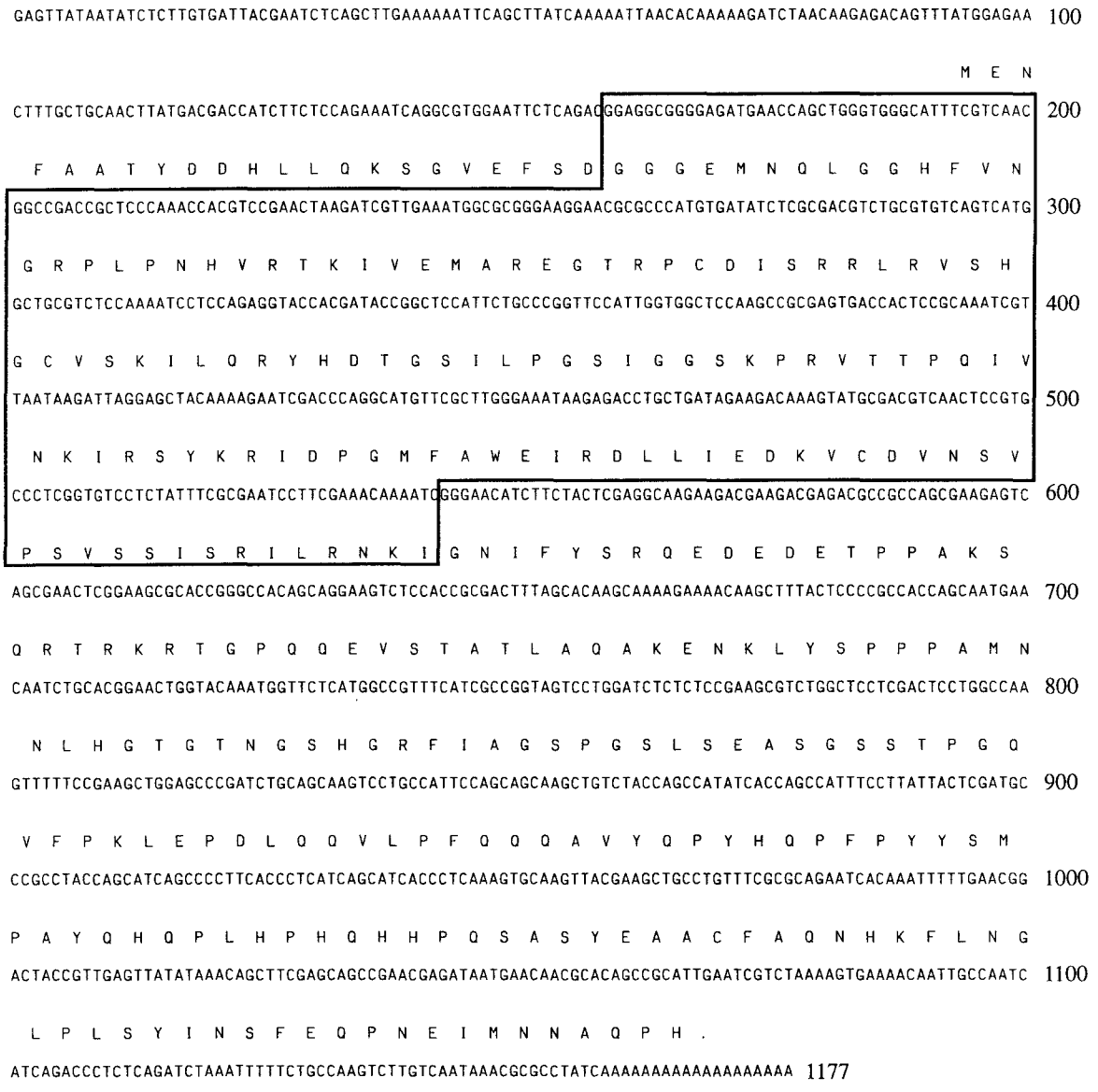
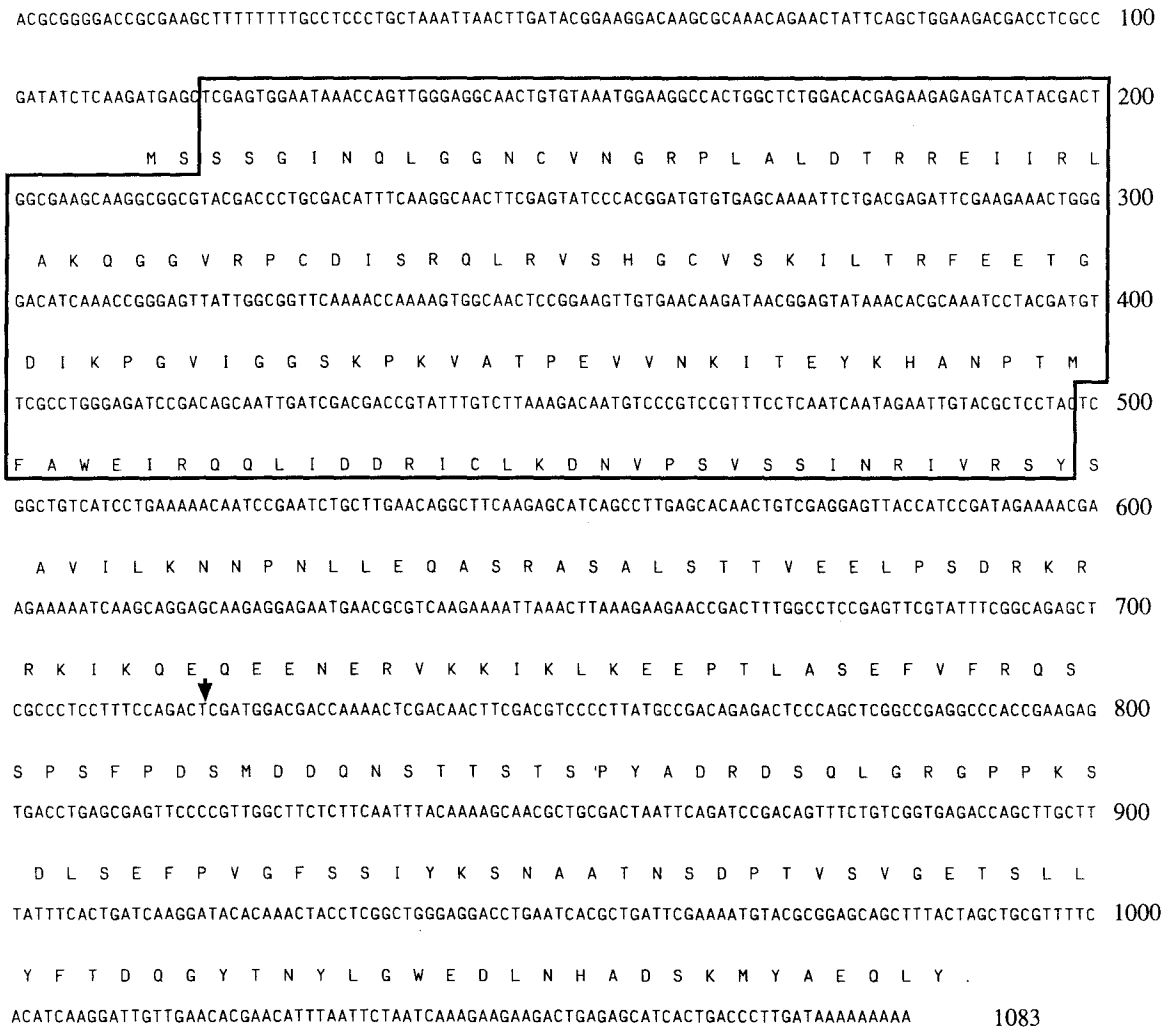


FIG. 7. DNA and predicted protein sequence of P*ax1/9*. The predicted protein is probably complete; though there is no in-frame, upstream stop codon, the predicted starting ATG is in a strong Kozak context, and most vertebrate Pax1 and Pax9 proteins have a shorter aminoterminus than the larvacean protein shown here. The paired domain is boxed. No octapeptide, present in other Pax1 and Pax9 class proteins, could be identified.

A.



B.

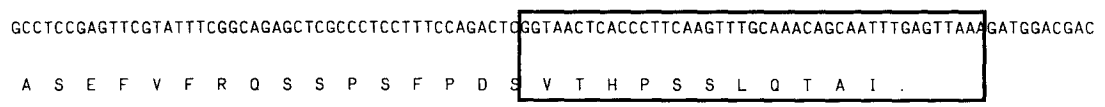


FIG. 8. DNA and predicted protein sequence of *Pax2/5/8*. (A) There is at least one stop codon in each reading frame upstream of the predicted start site, and the 5' end of the predicted protein is probably complete. The paired domain is boxed. No octapeptide or partial homeodomain, present in other Pax2, Pax5, and Pax8 class proteins, could be identified with confidence. An arrow marks the location of a 43bp insertion present in one 3' RACE clone. This may represent an alternate splice form that changes the sequence and truncates the carboxyterminus. (B) Alternate carboxyterminus created by a possible 43bp insertion.

significantly shorter than the average of other Pax2/5/8 family proteins carboxyterminal to the paired domain, despite having a complete UTR and poly-A tail. It is possible that, as in vertebrate orthologues, alternate splicing occurs in the transcripts of *O. dioica Pax2/5/8*, and isoforms may exist that contain conserved domains missing from the sequences cloned in this study.

The paired domain of *O. dioica* Pax6 is relatively much better conserved than the paired domains of larvacean Pax1/9 and Pax2/5/8 (Fig. 6). Like the ascidian Pax6 orthologue, the larvacean protein shows no evidence of a conserved linker region between the paired and homeodomains and is lacking all but a few residues of a conserved carboxyterminal motif (Fig. 9) present in the Pax6 proteins of non-urochordates.

The sequence and expression of *O. dioica Pax3/7* will be discussed in a separate report.

Structure of the Larvacean Brain and Sense Organs

The brain of *O. dioica* is small, consisting of approximately 70 cells, not all of which are neurons (Martini, 1909; Georges *et al.*, 1988). Three sets of paired nerves emanate from the brain (Olsson *et al.*, 1990): the first, most rostral pair (Fig. 10, 11) form a connection to axons from a structure called the "ventral organ" (Fig. 12, C), the second pair bifurcate to innervate presumed sensory cells of the upper and lower lips and ventral pharynx, and the third paired nerves terminate on cells of the ciliary ring in the spiracle or gill and nearby trunk epidermis. Close to the mouth and housed in a slit-like pocket of epidermis running perpendicular to the body axis, the ventral organ is a row of approximately 30

FIG. 9. DNA and predicted protein sequence of *Pax6*. The predicted protein is probably complete; there are two in-frame, upstream stop codons, and the starting ATG is in a strong Kozak context. An arrow marks a 120bp intron. An intron at this location is not present in the ascidian orthologue (Glardon *et al.*, 1997). The paired domain is boxed and the homeodomain is underlined (solid line). A short remnant of a carboxyterminal motif conserved in several other vertebrate and invertebrate Pax6 proteins is underlined with a dashed line. This motif is eroded also in the ascidian orthologue (Glardon *et al.*, 1997).

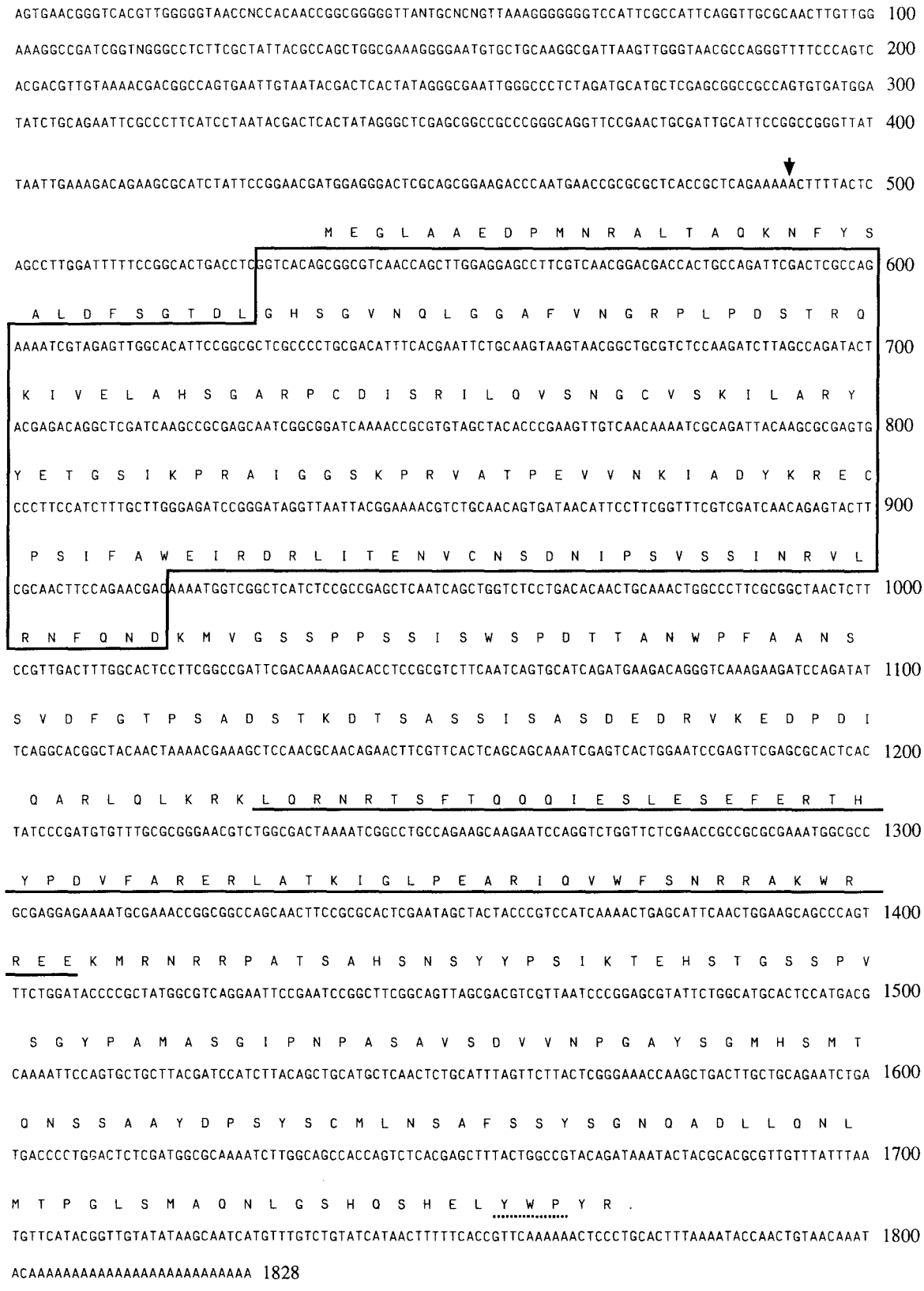
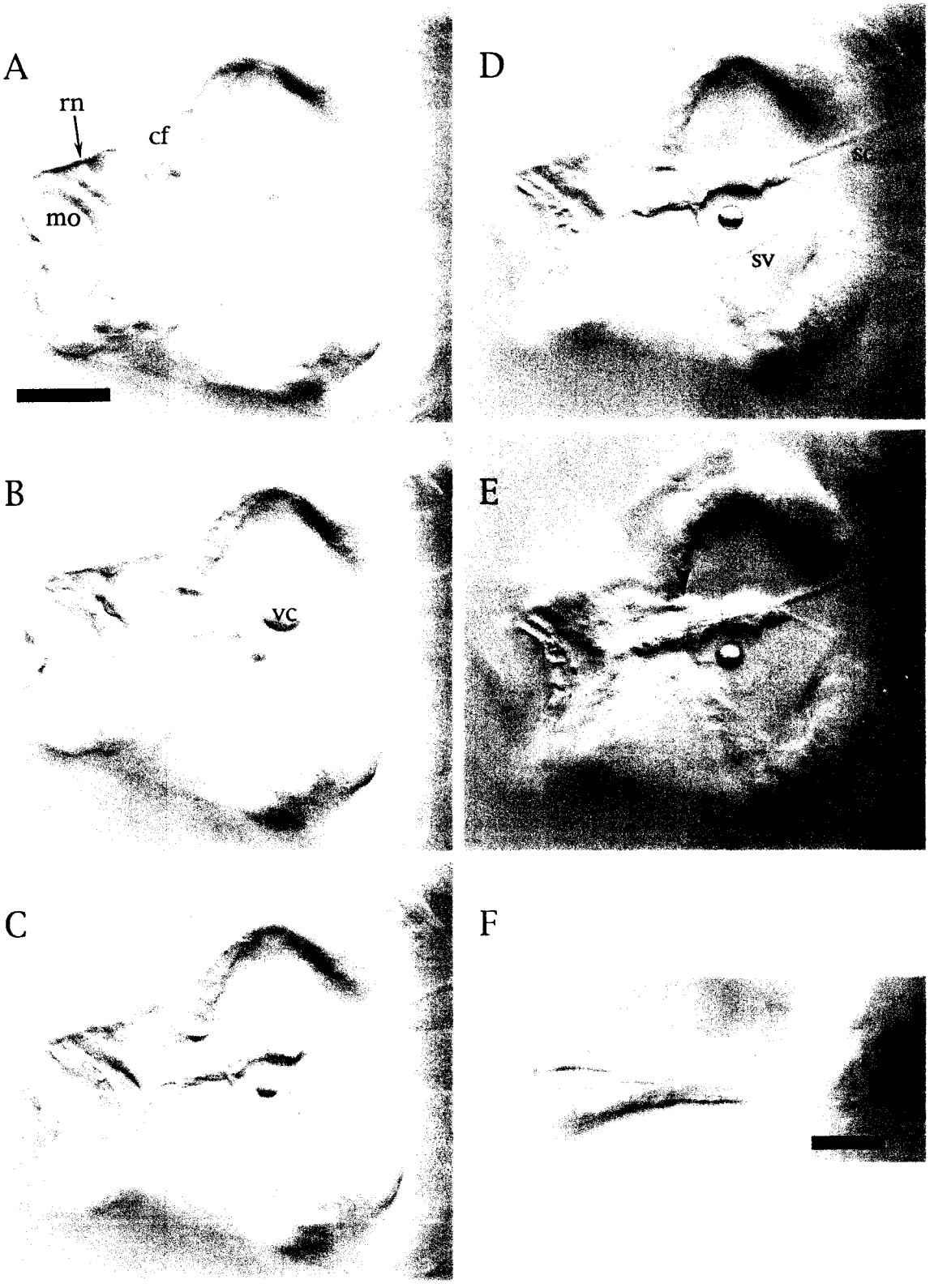

AGTGAACGGGTCACGTTGGGGGTAACCNCCACAACCGGCGGGGGTTANTGCNCNGTTAAAGGGGGGGTCCATTCGCCATTCAGGTTGCGCAACTTGTTGG 100
AAAGGCCGATCGGTNGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGAATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTC 200
ACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGA 300
TATCTGCAGAATTCGCCCTTCATCCTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGTTCCGAACTGCGATTGCATTCCGGCCGGGTTAT 400
TAATTGAAAGACAGAAGCGCATCTATTCCGGAACGATGGAGGGACTCGCAGCGGAAGACCCAATGAACCGCGCGCTCACCGCTCAGAAAAACTTTTACTC 500
M E G L A A E D P M N R A L T A Q K N F Y S
AGCCTTGGATTTTTCCGGCACTGACCTCGGTCACAGCGGCGTCAACCAGCTTGGAGGAGCCTTCGTCAACGGACGACCACTGCCAGATTCGACTCGCCAG 600
A L D F S G T D L G H S G V N Q L G G A F V N G R P L P D S T R Q
AAAATCGTAGAGTTGGCACATTCCGGCGCTCGCCCCTGCGACATTTCACGAATTCTGCAAGTAAGTAACGGCTGCGTCTCCAAGATCTTAGCCAGATACT 700
K I V E L A H S G A R P C D I S R I L Q V S N G C V S K I L A R Y
ACGAGACAGGCTCGATCAAGCCGCGAGCAATCGGCGGATCAAAACCGCGTGTAGCTACACCCGAAGTTGTCAACAAAATCGCAGATTACAAGCGCGAGTG 800
Y E T G S I K P R A I G G S K P R V A T P E V V N K I A D Y K R E C
CCCTTCCATCTTTGCTTGGGAGATCCGGGATAGGTTAATTACGGAAAACGTCTGCAACAGTGATAACATTCCTTCGGTTTCGTCGATCAACAGAGTACTT 900
P S I F A W E I R D R L I T E N V C N S D N I P S V S S I N R V L
CGCAACTTCCAGAACGACAAAATGGTCGGCTCATCTCCGCCGAGCTCAATCAGCTGGTCTCCTGACACAACTGCAAACTGGCCCTTCGCGGCTAACTCTT 1000
R N F Q N D K M V G S S P P S S I S W S P D T T A N W P F A A N S
CCGTTGACTTTGGCACTCCTTCGGCCGATTCGACAAAAGACACCTCCGCGTCTTCAATCAGTGCATCAGATGAAGACAGGGTCAAAGAAGATCCAGATAT 1100
S V D F G T P S A D S T K D T S A S S I S A S D E D R V K E D P D I
TCAGGCACGGCTACAACTAAAACGAAAGCTCCAACGCAACAGAACTTCGTTCACTCAGCAGCAAATCGAGTCACTGGAATCCGAGTTCGAGCGCACTCAC 1200
Q A R L Q L K R K L Q R N R T S F T Q Q Q I E S L E S E F E R T H
TATCCCGATGTGTTTGCGCGGGAACGTCTGGCGACTAAAATCGGCCTGCCAGAAGCAAGAATCCAGGTCTGGTTCTCGAACCGCCGCGCGAAATGGCGCC 1300
Y P D V F A R E R L A T K I G L P E A R I Q V W F S N R R A K W R
GCGAGGAGAAAATGCGAAACCGGCGGCCAGCAACTTCCGCGCACTCGAATAGCTACTACCCGTCCATCAAAACTGAGCATTCAACTGGAAGCAGCCCAGT 1400
R E E K M R N R R P A T S A H S N S Y Y P S I K T E H S T G S S P V
TTCTGGATACCCCGCTATGGCGTCAGGAATTCCGAATCCGGCTTCGGCAGTTAGCGACGTCGTTAATCCCGGAGCGTATTCTGGCATGCACTCCATGACG 1500
S G Y P A M A S G I P N P A S A V S D V V N P G A Y S G M H S M T
CAAAATTCCAGTGCTGCTTACGATCCATCTTACAGCTGCATGCTCAACTCTGCATTTAGTTCTTACTCGGGAAACCAAGCTGACTTGCTGCAGAATCTGA 1600
Q N S S A A Y D P S Y S C M L N S A F S S Y S G N Q A D L L Q N L
TGACCCCTGGACTCTCGATGGCGCAAAATCTTGGCAGCCACCAGTCTCACGAGCTTTACTGGCCGTACAGATAAATACTACGCACGCGTTGTTTATTTAA 1700
M T P G L S M A Q N L G S H Q S H E L Y W P Y R .
TGTTCATACGGTTGTATATAAGCAATCATGTTTGTCTGTATCATAACTTTTTCACCGTTCAAAAAACTCCCTGCACTTTAAAATACCAACTGTAACAAAT 1800
ACAAAAAAAAAAAAAAAAAAAAAAAAAA 1828

FIG.10. Optical sections of the anterior trunk and brain of a live adult. A, most ventral; E most dorsal. (A) The bottom of the ciliary funnel (cf) and the most ventral cells of the brain bulbs that line the rostral nerve (rn, rostral nerve). (B) One of the large, non-neural "ventral cells" (vc) described by Georges *et al.* (1998), obvious in the ventral brain. (C) A view of the medial connection of the rostral nerve trunks to the brain. (D) A section through the brain and sensory vesicle (sv) at the level of the statolith and spinal cord (sc). (E) An arrowhead marks where the ciliary funnel connects to the brain. (F) A longitudinal view of the ciliary funnel (ventral to the left). (mo, mouth). Scale bars A-E, 50μm; F, 10μm.


A
rn
cf
mo
D
sv
B
vc
E
C
F

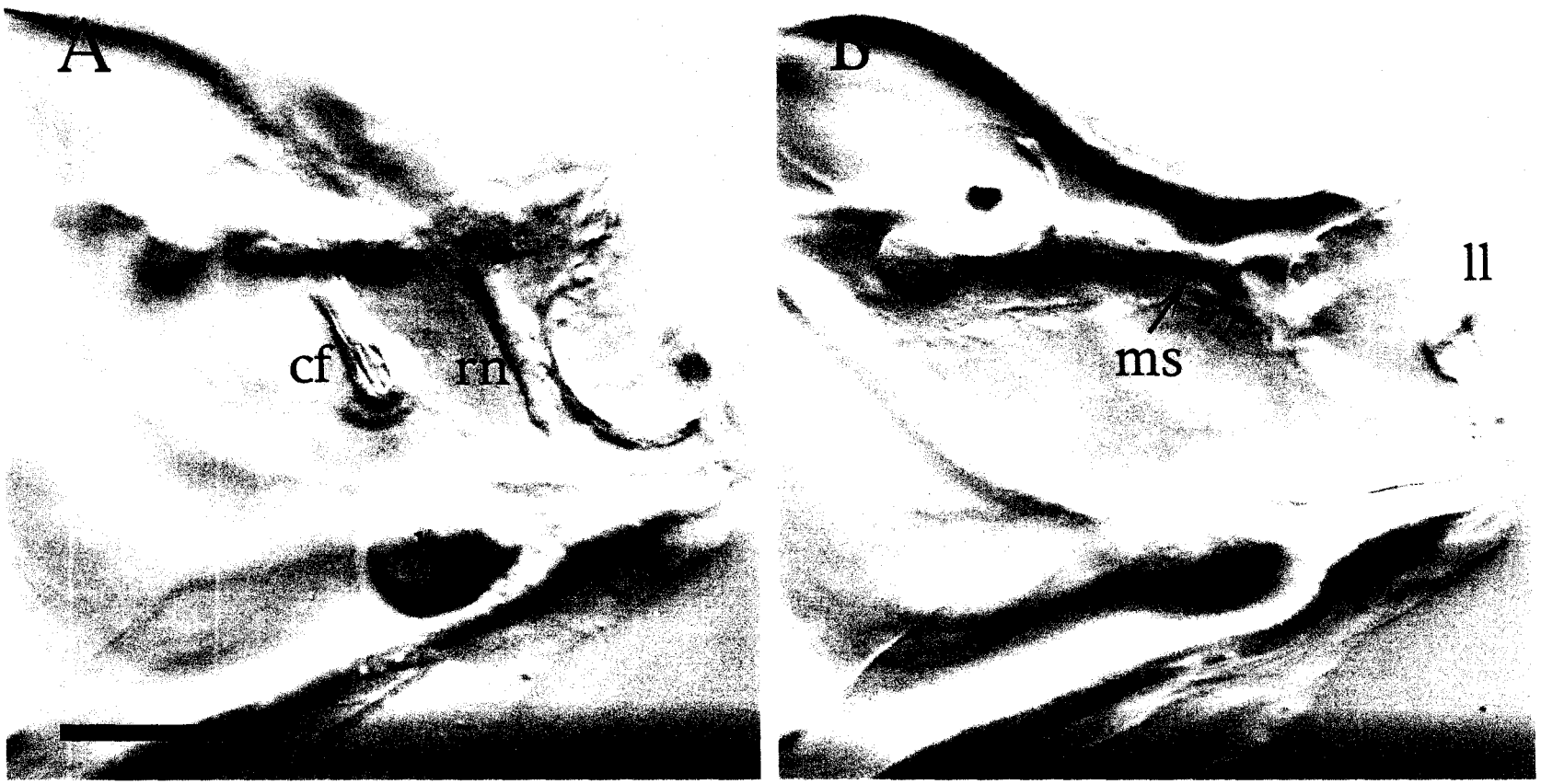


FIG. 11. Right side views of the adult brain. In subsequent figures, animals will appear with the anterior oriented to the left, unless right side structures need to be shown, such as in these views of the adult brain. (A) Right of the midline, showing a longitudinal section of the lower half of the ciliary funnel (cf), the length of the right rostral nerve (rn), and the right buccal gland (bg). (B) Midline showing where the right and left rostral nerves meet the medial stalk (ms) from the brain. (ll, lower lip). Scale bar, 50μm.

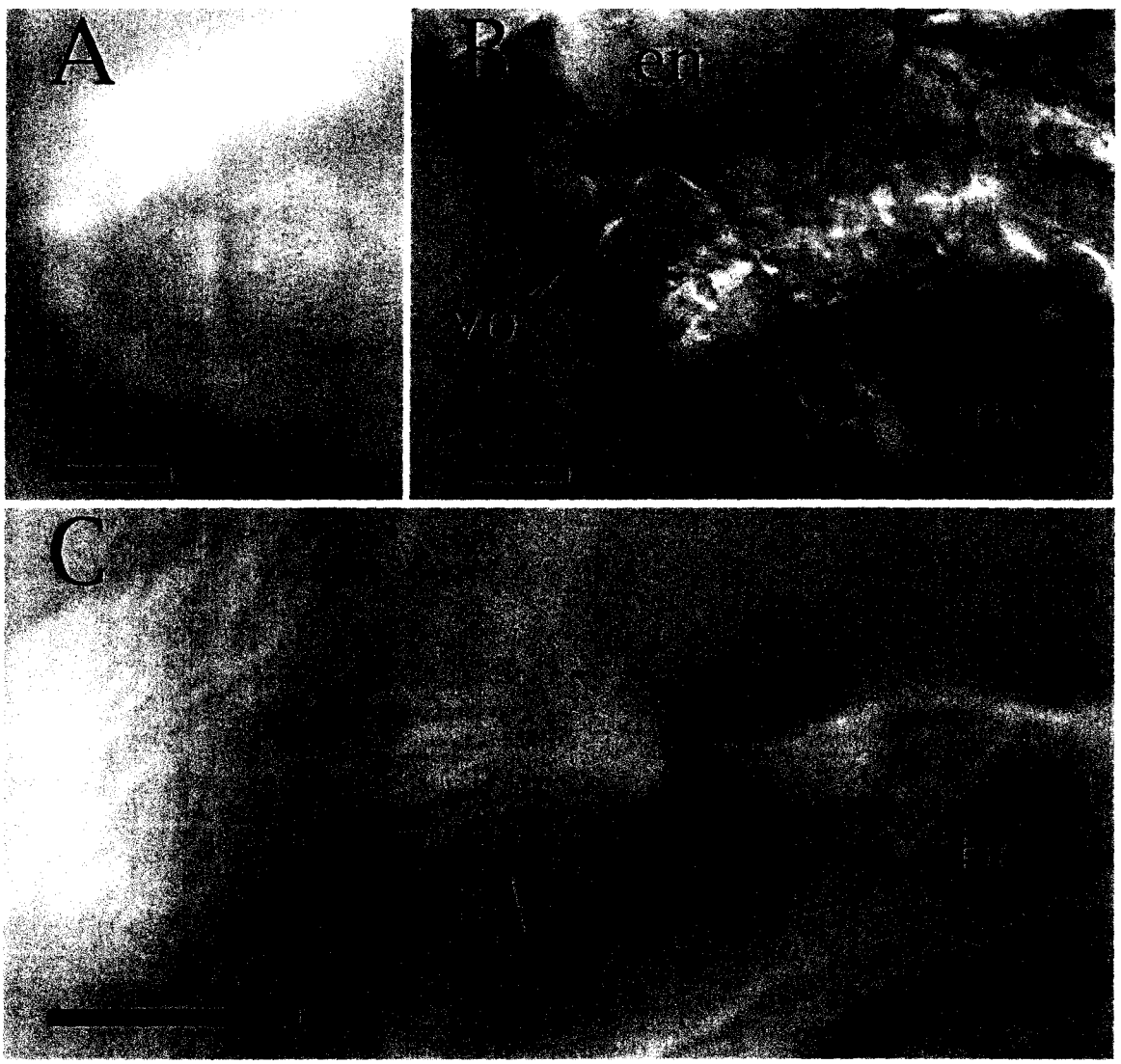


FIG. 12. Anatomy of the ventral organ. (A) Tailshift stage, superficial view of ventral side of the anterior trunk showing the comb of cilia from the ventral organ. Anterior to the left. (B) Lateral view of an adult ventral organ (vo). The animal had to be slightly crushed to see this optical section. (C) Adult rostral nerve, showing the last bulb cell (bc) and the separate axons (arrowhead) from the ciliary cells of the ventral organ just before they become bundled into a cross-sectionally round nerve. (en, endostyle; mo, mouth). Scale bars A, C 10μm; B 20μm.

cells, each bearing a modified cilium on the apical surface and sending an axon from the basal surface toward the brain (Fig. 12, A, B) (Bollner *et al.*, 1986). These cells were suggested by Bollner *et al.* (1986) to be primary sensory cells, and by various authors to be olfactory or otherwise chemosensory (Martini, 1909; Lohmann, 1933; Bollner *et al.*, 1986). Four cells on each of the rostral nerve tracks line the junction of axons from the brain and axons from the ciliary cells of the ventral organ; these fall into three cell types, based on differences in their ultrastructure. One cell appears to ensheath the axon bundle. This cell and two others send processes into the brain. The fourth cell has no processes, but all four cells exhibit synaptic connections with either brain axons, sensory axons or both (Bollner *et al.*, 1986). On the basis of these observations and assuming the cells to be an outgrowth of the brain, Bollner *et al.* (1986) called them collectively the "rostral brain bulbs" and tentatively suggested their homology to the vertebrate olfactory bulbs. However, Delsman (1912) described the development of these bulbs from the rostral epidermis, not from the brain: on either side of the midline, rows of epidermal cells protrude inward, become connected to the brain by a medial stalk, and ultimately separate from the epidermis in a medial to lateral progression. My observations on live and fixed embryos agree with Delman's description. The medial brain stalk is likely to be the combined left and right bundles of axons that come from the brain to meet the ventral organ axons at the level of the bulbs, as described by Bollner *et al.* (1986).

On the left side of the brain is a large vesicle containing a crystalline statolith (Fig. 10, D), sensory cells whose cilia contact the lith (Holmberg, 1984), and putatively pressure-sensitive coronet cells (Olsson, 1975). Opposite this sensory vesicle is the

ciliary funnel, a cone which slopes rostrally and ventrally from the right side of the brain and whose base forms an opening into the dextro-dorsal wall of the pharynx (Fig. 11, A). The funnel is divided into two parts, based on ultrastructural characteristics (Holmberg, 1982). In the ventral part, the funnel walls, formed of a monolayer of cells, bear many long cilia that constantly beat toward the dorsal end of the funnel (Fig. 10, F). The dorsal part of the funnel is unciliated, and its vertex slightly extends above the funnel's contact with the brain. A central cavity periodically forms between the funnel "tip" cells, at least one of which is innervated by a process from the brain. A granular and filamentous nuage in the cavity appears to be released through interstices between the tip cells into the haemocoel (Holmberg, 1982). Olsson (1969) and Holmberg (1982) have argued that the ciliary funnel is homologous to both the adeno- and neurohypophyses. This hypothesis mirrors ideas in the classical literature in which Hatschek's pit in cephalochordates (e.g., Goodrich, 1917) and the neurohypophyseal duct or neural gland complex of ascidians (e.g., Willey, 1893) have long been debated as adenohypophysis homologues.

As an anatomical and developmental reference for these unfamiliar animals, two figures of optical sections through the trunks of live hatchlings are included. Figure 13 shows the trunk organization of a 12 hour old hatchling. Figure 14 shows the anatomy of a tailshift stage hatchling; this stage marks the end of embryonic development, and the beginning of the juvenile stage.

The description of results below is divided roughly along anatomical lines: expression of genes in the developing ventral organ, ciliary funnel, and gill.

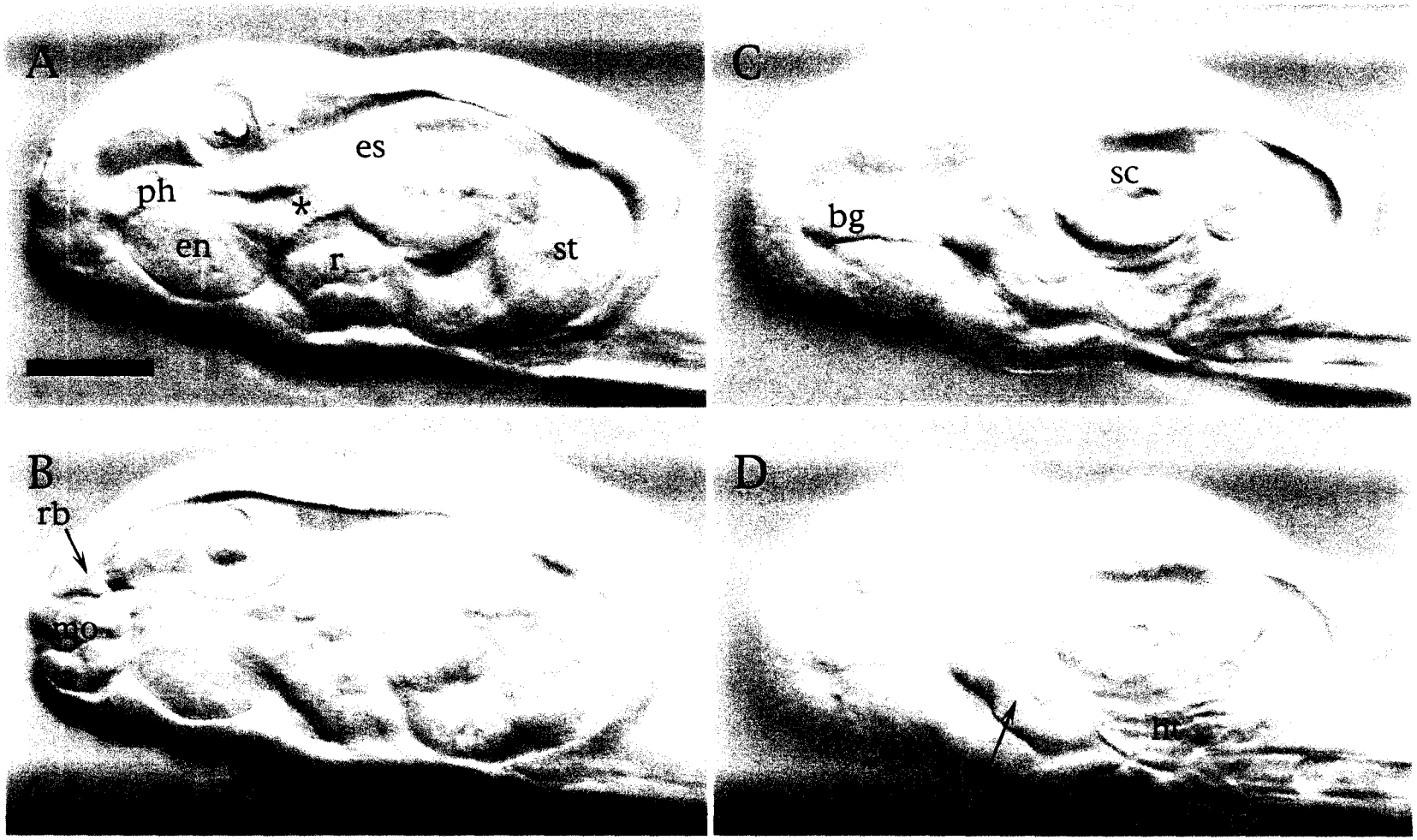


FIG. 13. Optical sections through a live 12 hour old hatchling. A, most medial to D, most lateral. (A) view of pharynx in the saddle region (*) between the two lateral spiracle pouches. B) Rudiment of the rostral brain bulbs (rb) protruding inward from the epidermal epithelium. (C) The buccal gland tissue (bg) is still migrating forward. (D) Spiracles (sp) have not yet penetrated the epidermis. (en, endostyle; es, esophagus; ht, heart; mo, mouth; ph, rostral pharynx; sc, canal that connects the right and left lobes of the stomach; st, stomach; r, rectum). Scale bar, 20μm.

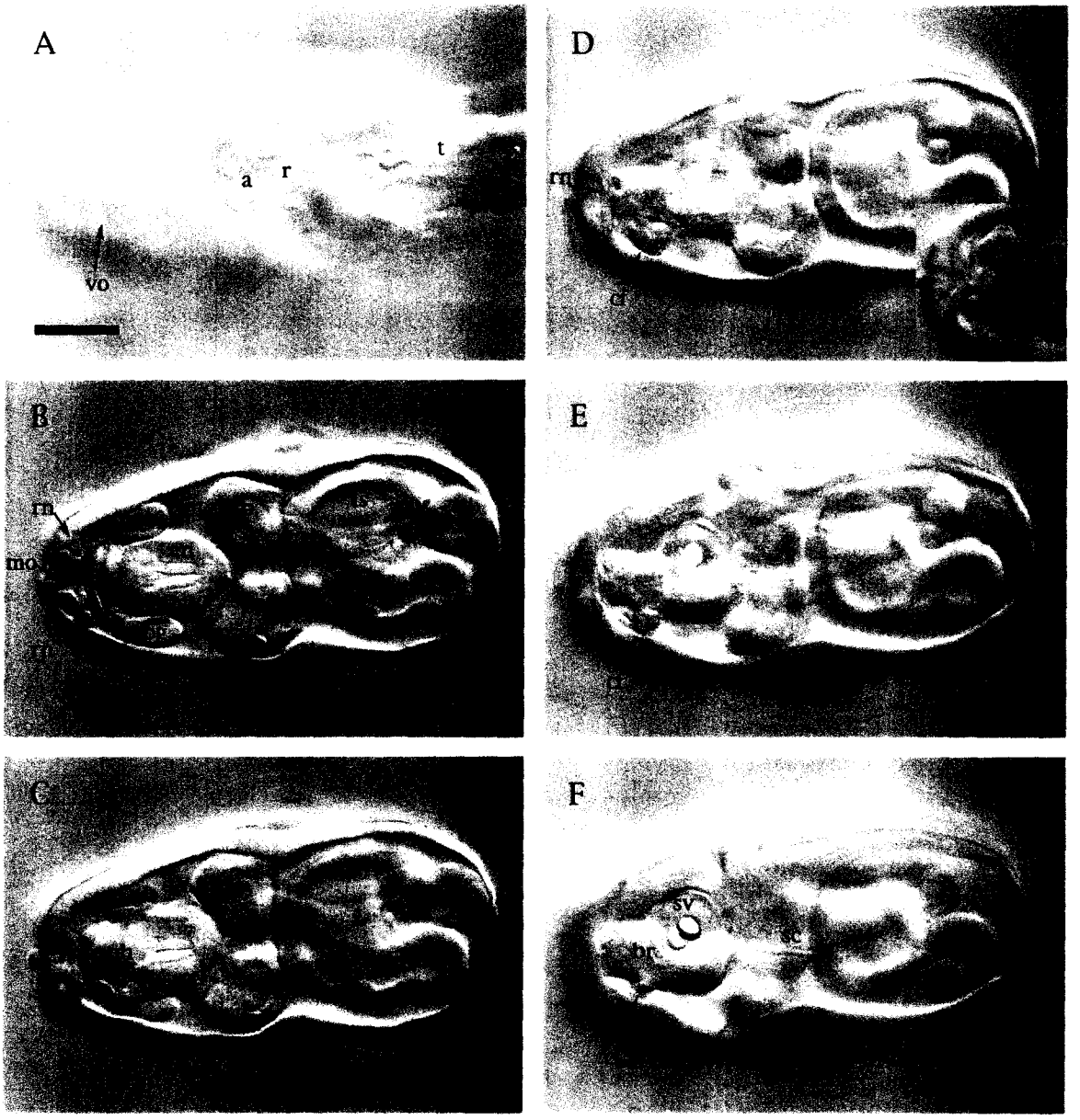


FIG. 14. Optical section series through a tailshift stage juvenile seen in ventral view. A, most ventral to F, most dorsal. (A) Superficial view of the ventral surface, showing the comb of cilia and groove of the ventral organ (vo). (B) Several important features are visible in this section: the mouth opening (mo), the most ventral of the bulb cells of the rostral nerve (rn), buccal glands (bg) on either side of the endostyle (en), the spiracles (sp), the right and left stomach lobes (rs, ls), and one of the paired Langerhans mechanoreceptors (mr). (C) The pharynx (ph) bends up and over the endostyle. (D) The rostral nerves (rn) issue from a medial brain stalk. The ciliary funnel and the rostral nerves can be seen more clearly in a dorsal view (inset). (E) The ciliary funnel narrows more dorsally. (F) This section cuts through the sensory vesicle (sv) and the top of the ciliary funnel. (a, anus; r, rectum; t, tail). Scale bar, 20μm.

Eya, Six1/2, and Pitx Expression in the Developing Ventral Organ

Larvacean *Six1/2* and *Eya* have expression patterns that overlap in the rostral epidermis. In early hatchlings, *Six1/2* and *Eya* are expressed in epidermal cells surrounding the presumptive mouth. On the ventral surface near the mouth, two patches of *Six1/2* expression meet at the midline (Fig. 15, B). Similar patterns are observed for both *Eya* and *Pitx* (Fig. 15, A, C). It is thought that the ventral organ, which forms in this region, develops from epidermal cells that sink into a slit-shaped pit perpendicular to the body axis (Bollner *et al.*, 1986). The columnar epithelium of the trunk epidermis of young hatchlings conforms closely to the underlying organs and can be seen to indent and appear thickened in various places, including at the margins of the endostyle (Fig. 16, D). In lateral view, *Six1/2* expression is just rostral to the thickened epidermis at the anterior border of the underlying endostyle; if the expressing cells are presumptive ciliary receptor cells, then this represents a stage before their involution (Fig. 16, E).

Vertebrate *Eya1* and *Eya2* require the expression of *Pax6* for their expression in the olfactory placode (Xu *et al*, 1997), and expression of the *Drosophila* homologue of *Six1/2* (*so*) is directly activated by *Pax6* in the eye (Niimi *et al.*, 1999). In addition some Eya and Six proteins can interact and probably act cooperatively to activate downstream genes (Ohto *et al.*, 1999). Despite the overlapping expression of larvacean *Eya* and *Six1/2* in the rostral epidermis, however, the larvacean *Pax6* homologue cloned in this study is

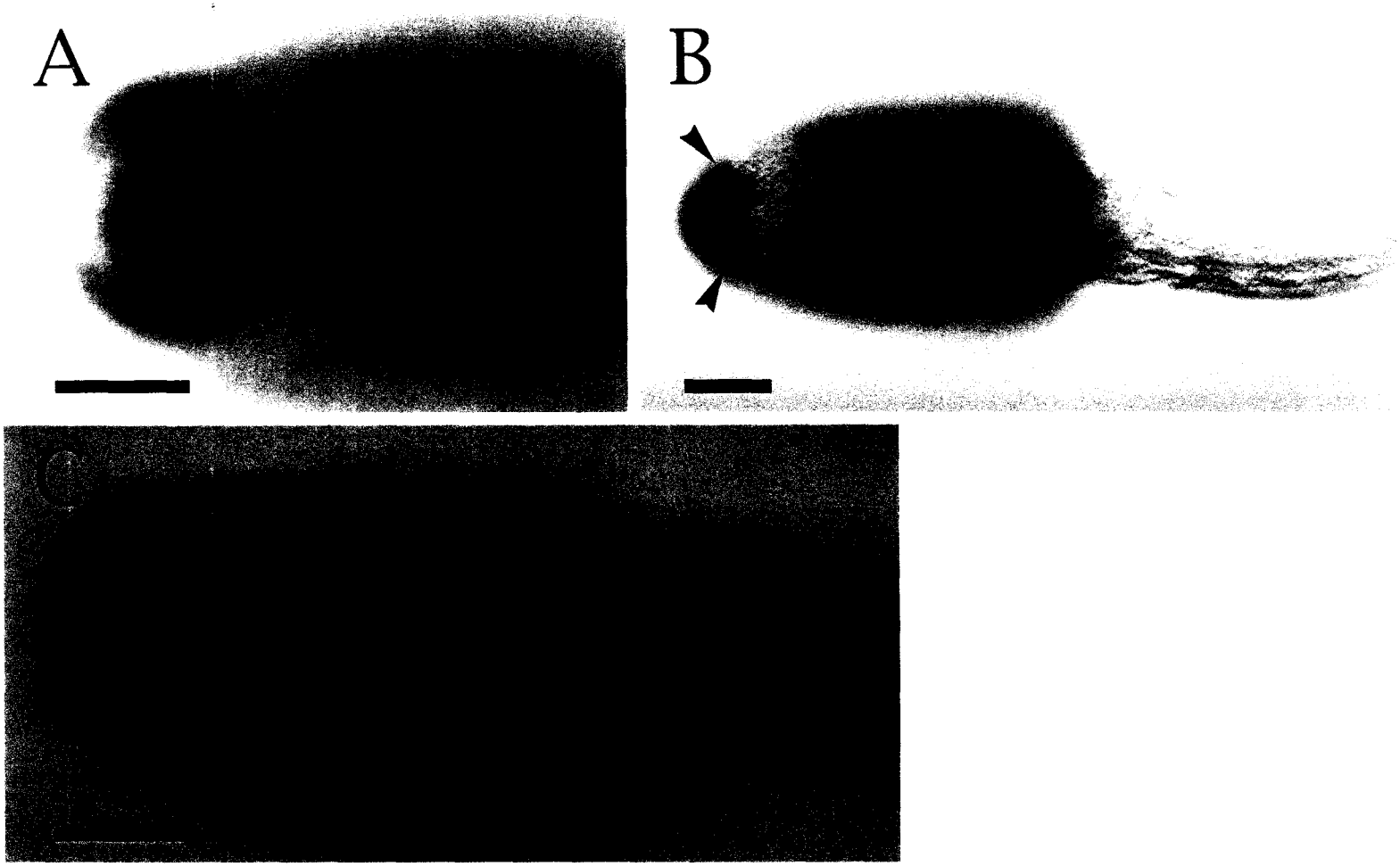


FIG. 15. Expression of *Eya*, *Six1/2* and *Pitx* in the ventral organ primordium in 8.5 hour hatchlings. (A) Expression of *Eya*, (B) *Six1/2*, and (C) *Pitx* in the ventral epidermis. Views of *Eya* and *Pitx* are slightly oblique, so more of the dorsal expression is visible than in B. *Six1/2* expression appears as two patches (arrowheads) that just touch at the midline. It is not clear whether the staining in most of the posterior trunk epidermis (but not the tail) is specific or is a background artifact of this particular probe. Scale bars A, C 10μm; B, 20μm.

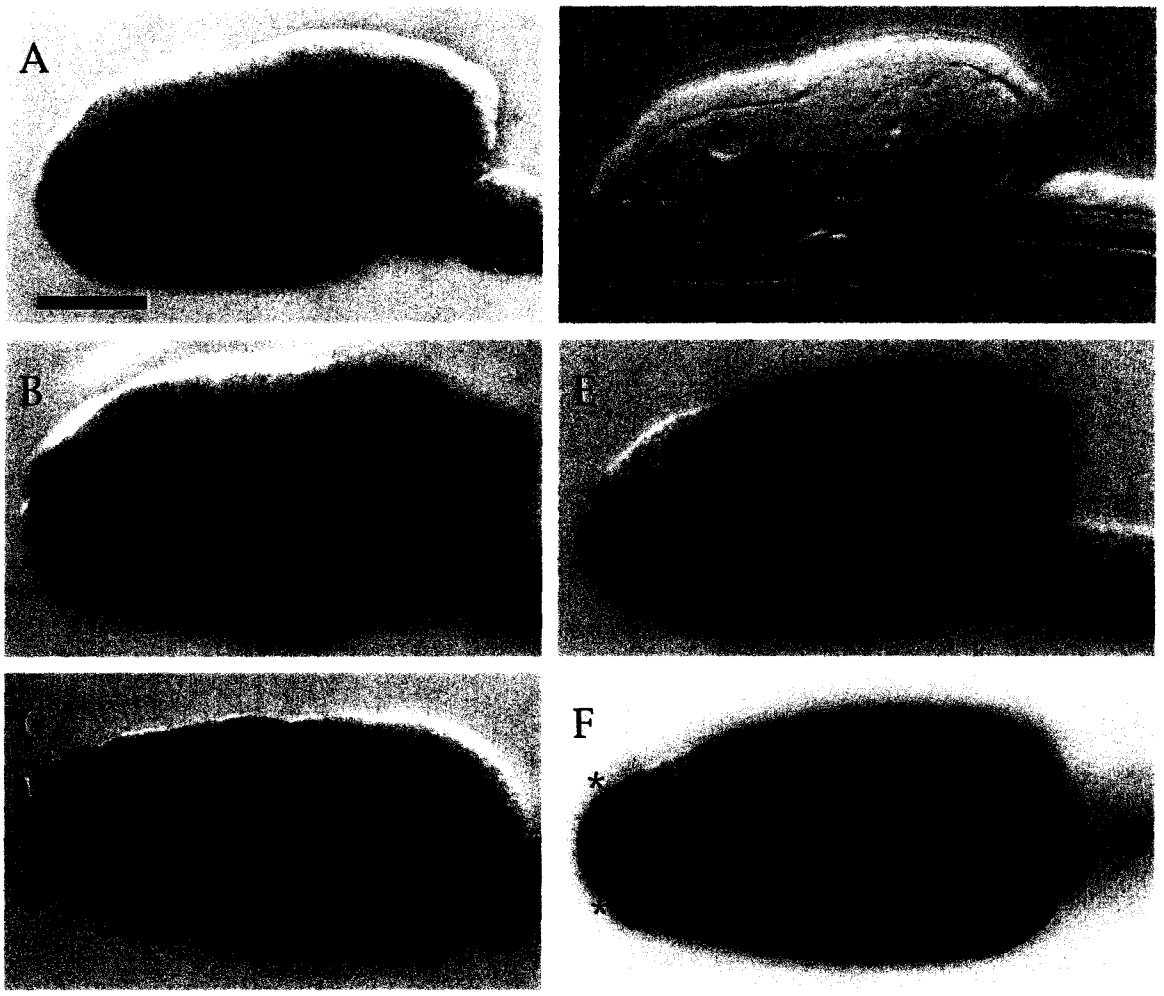


FIG. 16. *Eya* and *Six1/2* expression domains overlap in the rostral epidermis. (A) *Eya* expression in an 8.5 hour old hatchling. (B, C) Lateral and dorsal views of *Eya* expression in a 12 hour hatchling. (D) 8.5 hour live embryo showing the thick epidermis of the trunk, especially thickened at the margins of the endostyle (en). (E, F) Lateral and dorsal views of *Six1/2* expression in an 8.5 hour embryo. In dorsal views, asterisks indicate the row of epidermal cells expressing *Eya* and *Six1/2*.

not expressed in an overlapping domain at any time from early tailbud stage forward (Fig. 17). *Pax6* is expressed only in the forebrain, anterior to the sensory vesicle, and in a small spot to the right of the midline in the posterior brain (Fig. 17, C, D).

Pitx expression is comparatively broader than both *Six1/2* and *Eya* in the rostral epidermis (Fig. 15, C; Fig. 18, B, C; Fig. 19, A). A band of cells around the presumptive mouth strongly expresses *Pitx* until hatchlings are 13 hours old. Excluded from this epidermal expression are at least the upper lips cells, which in later stages can be easily recognized by sensory bristles. *Pitx* expression expands into the rostral pharynx, but its posterior border does not reach the pharyngeal gill endoderm (Fig. 19, A; Fig. 20, A). The ciliary funnel expresses *Pitx* at the level of the pharynx (Fig. 20 A, B). At late stages, while expression in the epidermis has nearly disappeared, *Pitx* is expressed in the rostral nerves (Fig. 20, A).

Other Sites of Pitx and Eya Expression

Pitx expression is complex and dynamic. Expression is detectable at least as early as neurula stage (not shown), and in early tailbud stage embryos, *Pitx* is expressed in an anterior domain and in the tail muscle cells (Fig. 18, A). This muscle expression disappears and only the domain in the rostrum persists through hatching (Fig. 18, B). New sites of expression appear shortly after hatching: a cell that disrupts the contiguity of the last notochord cell with the notochordal rod (Fig. 18, C), and epidermal cells at the junction of the tail and trunk (Fig. 18, B, C). The latter expression persists until late stage hatchlings (Fig 19, A). The buccal glands, which come to lie lateral to the endostyle, will secrete bioluminescent material into the house rudiment. The developing buccal gland

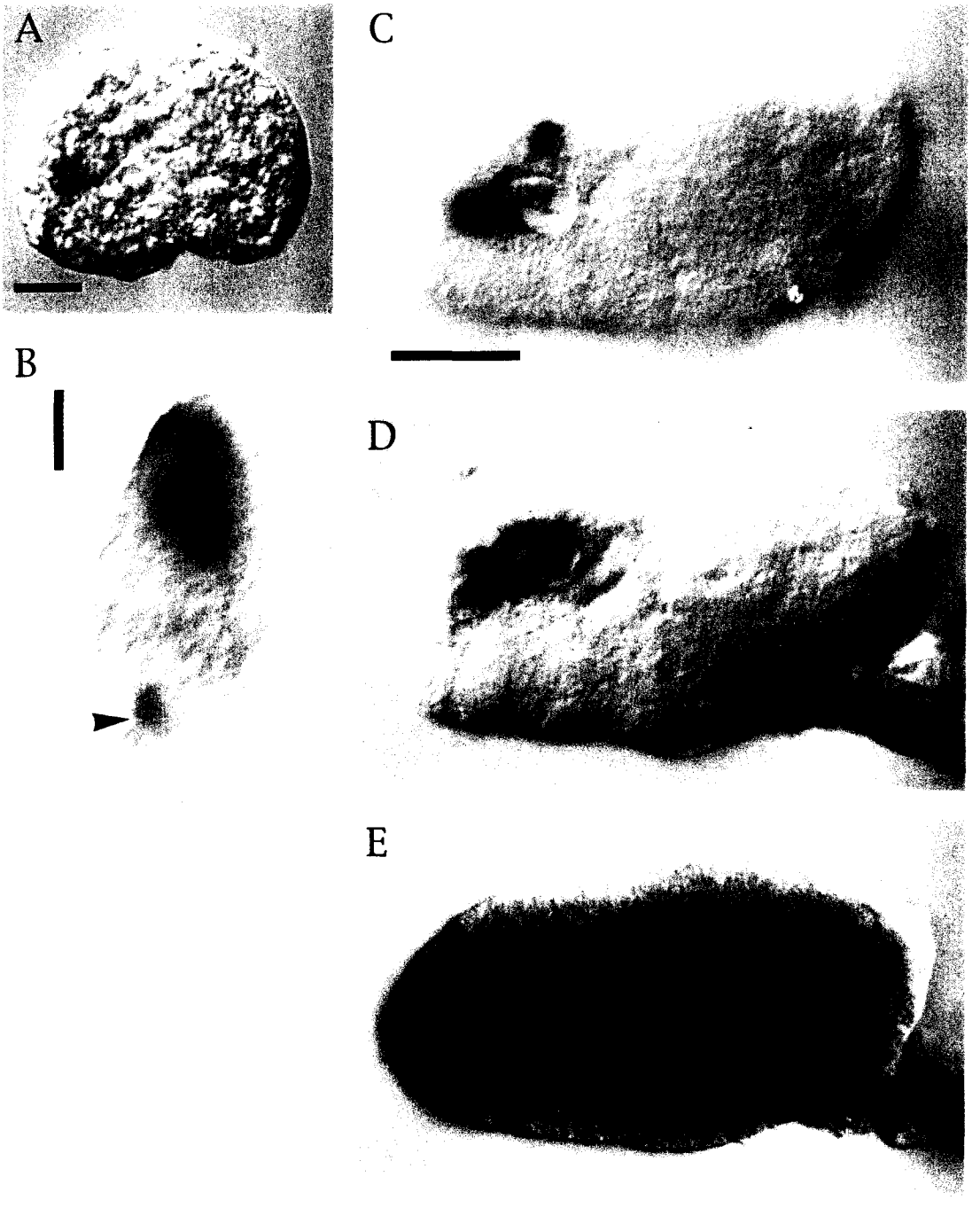


FIG. 17. Expression of *Pax6* only in the CNS and not in the rostral epidermis. (A) An early tailbud stage embryo (4 hours old). (B) A dorsal view, just after hatching (6 hours). *Pax6* is expressed asymmetrically in the brain. A small spot of expression at the base of the tail on the left side could be in the developing caudal ganglion (arrowhead), a group of mostly neural cells that make up a swelling in the spinal cord. The tail is rotated 90° in larvaceans so that the spinal cord lies to the left of the notochord. (C, D) Dorsal and ventral views of a 10.5 hour old hatchling. Expression is restricted to the brain, rostral to the sensory vesicle, and a more posterior spot on the right side of the brain. (E) *Eya* expression for comparison. The *Pax6* and *Eya* domains do not overlap at any stage. Scale bars A, B, 20μm; C-E, 20μm.

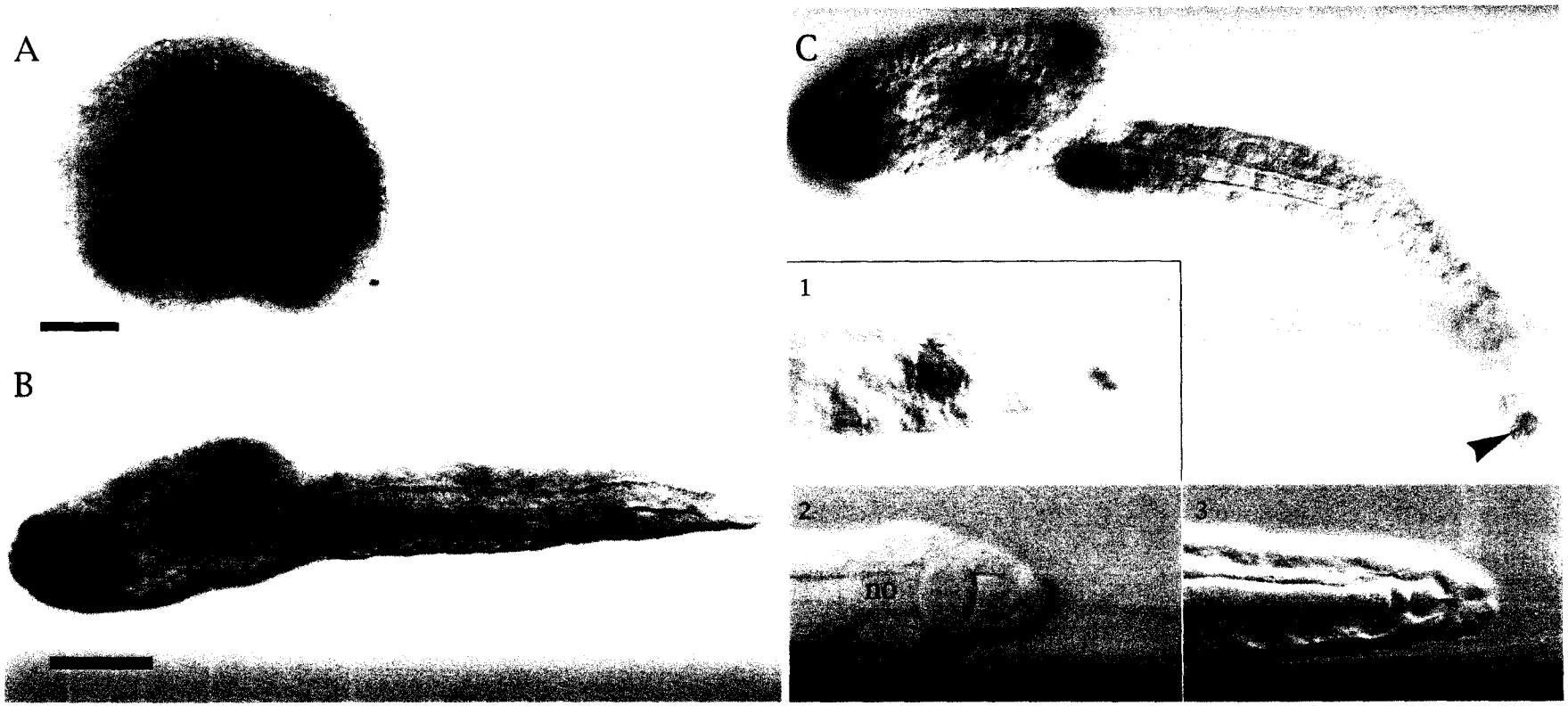


FIG. 18. Early expression domains of *Pitx*. (A) Tailbud stage embryo (4.5 hours old). An oblique lateral view shows *Pitx* expression in the two bands of tail muscle cells and in a rostral domain. (B) A 6.5 hour old hatchling, 30 minutes after hatching. There is a new domain at the base the tail. Expression in muscle cells has ceased. (C) A 7 hour old hatchling. *Pitx* is transiently expressed in a single cell (arrowhead) at the tip of the tail. This cell separates the last notochord cell from the rest of the notochord. Insets show an enlargement of the tail tip (1) in the same hatchling as in C, (2) in a live 8.5 hour hatchling, and (3) in a 10.5 hour hatchling showing the transformation of this cell over time. It is not known what the function of this cell is, although classical anatomists proposed the tail tip of hatchling larvaceans carries a sensory structure (Lohmann, 1933). (no, notochord) Scale bars A, 20μm; B, C, 30μm.

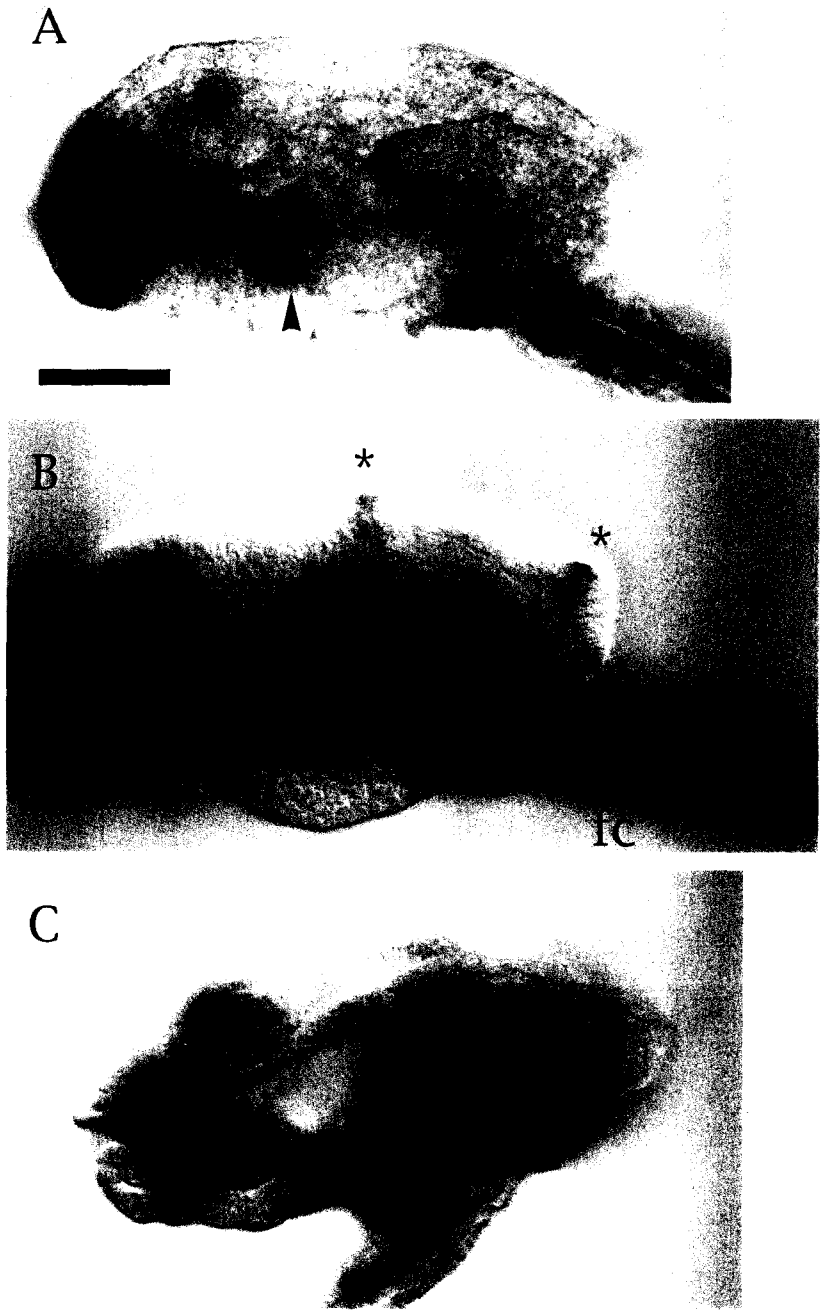
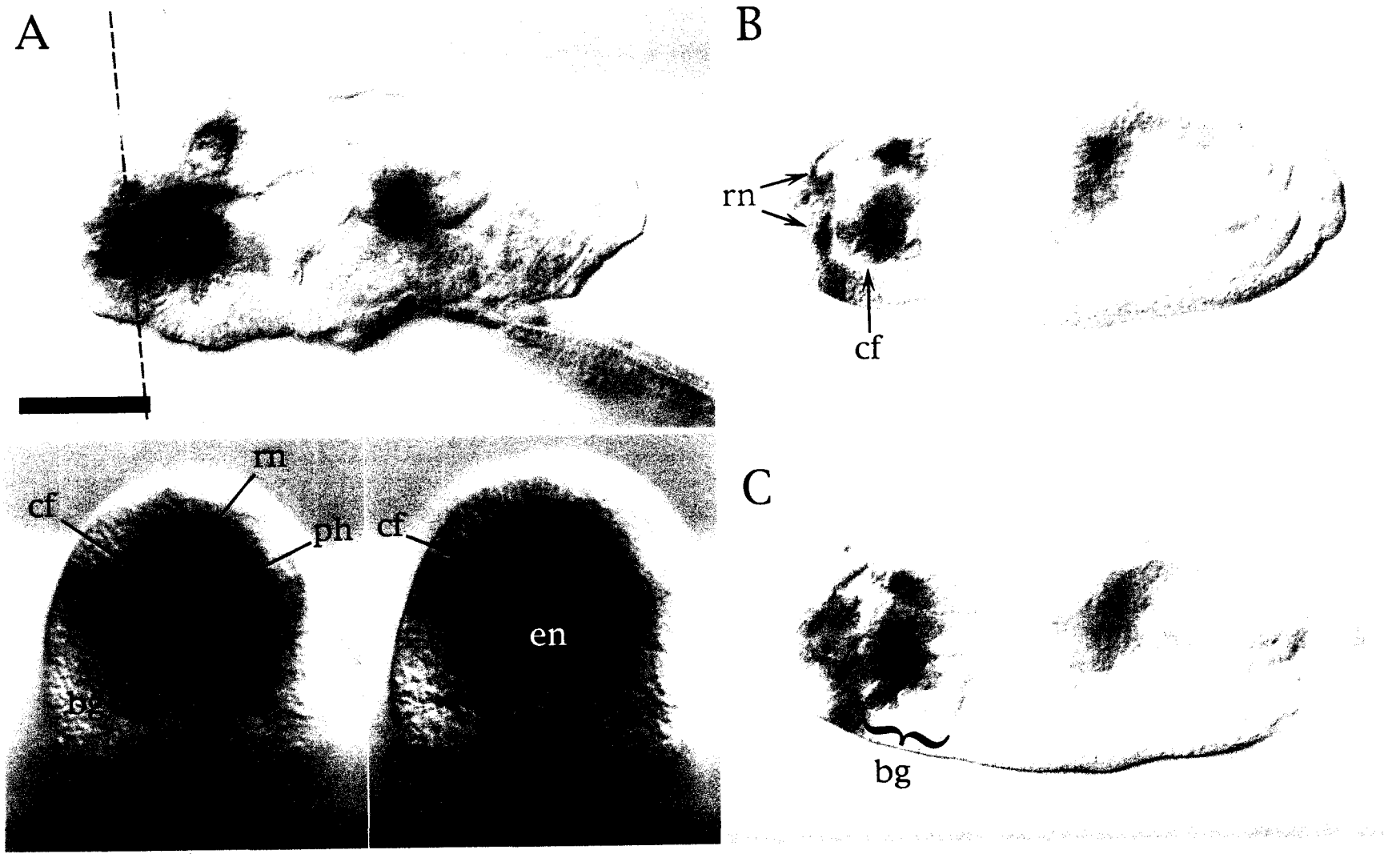


FIG. 19. Other *Pitx* expression domains. (A) Lateral, medial view of a 12 hour old hatchling showing that expression of *Pitx* has spread into the rostral pharynx. The migrating buccal gland tissue is seen in cross-section (arrowhead) at the rear of the endostyle. (B) A more lateral view of a 12 hour old hatchling, showing *Pitx* expression in four epidermal cells (fc) on the fin of the tail, in the stomach connection canal and left stomach lobe (st), and in two cells (*) in the midline of the dorsal epidermis, visible because the trunk is turned slightly toward us. (C) At tailshift stage, *Pitx* is still expressed in the stomach, the brain, and the pharynx. Scale bar, 20μm.

FIG. 20. *Pitx* is expressed in the rostral nerves and the ventral segment of the ciliary funnel in 13 hour old hatchlings. (A) Lateral view showing expression in the rostral nerves, in the pharynx, and in the stomach. Insets show rostral nerve trunk expression in optical cross section. Inset 1 is a section approximately at the level of the dashed line in panel A, and Inset 2 is at a level a little more posterior. *Pitx* is expressed in the anterior end of the buccal gland (bg). (B) Ventral view showing expression in the rostral nerve trunks (rn) and the base of the ciliary funnel (cf). (C) A more ventral view than in panel B showing expression in only the most anterior segment of the buccal gland. (en, endostyle; ph, pharynx).


A
B
rn
cf
rn
cf
ph
cf
en
C
bg

tissue, which forms a horseshoe shape around the posterior margin of the endostyle (Fig. 21, B), expresses *Pitx* in 12 hour old hatchlings (Fig. 21, A, C). Buccal gland expression continues as the tissue medially bifurcates and migrates to bilateral positions at the anterior margin of the endostyle, but expression becomes narrowed to an anterior segment of the presumptive buccal gland (Fig. 20, B-D). Additional expression domains include two cells in the midline, dorsal epidermis, gut cells of the left stomach lobe and the connection between left and right lobes, and in four pairs of epidermal cells along the tailfin edges (Fig. 19, A-C).

Eya is also expressed in domains outside of the rostral epidermis. Two additional regions express *Eya* in young hatchlings: the pharynx at the dorsal ends of the developing gill tubes (spiracles), and the ventral epidermis approximately where the spiracles will break through. This expression complements the pharyngeal domains of *Pax2/5/8* and *Pax1/9*, which are also expressed in the developing spiracles (discussed below).

Six3/6a and Six3/6b Expression in the Ciliary Funnel

Duplicate larvacean homologues of *Six3/6* are expressed in the rostral pharynx. The expression domain of *Six3/6b* is nested within the much broader *Six3/6a* expression. In young hatchlings, expression of these two genes appears comparable (not shown), but in later stages *Six3/6a* expression rapidly expands. While *Six3/6a* is expressed in the entire roof of the anterior pharynx, in the brain, in the esophagus, and in posterior epidermal domains, *Six3/6b* is seen only in a trio of cells on the right side that probably make up the most ventral part of the ciliary funnel where it perforates into the dextro-dorsal wall of the pharynx (Fig. 22). The ciliary funnel appears to grow from the right side of the brain

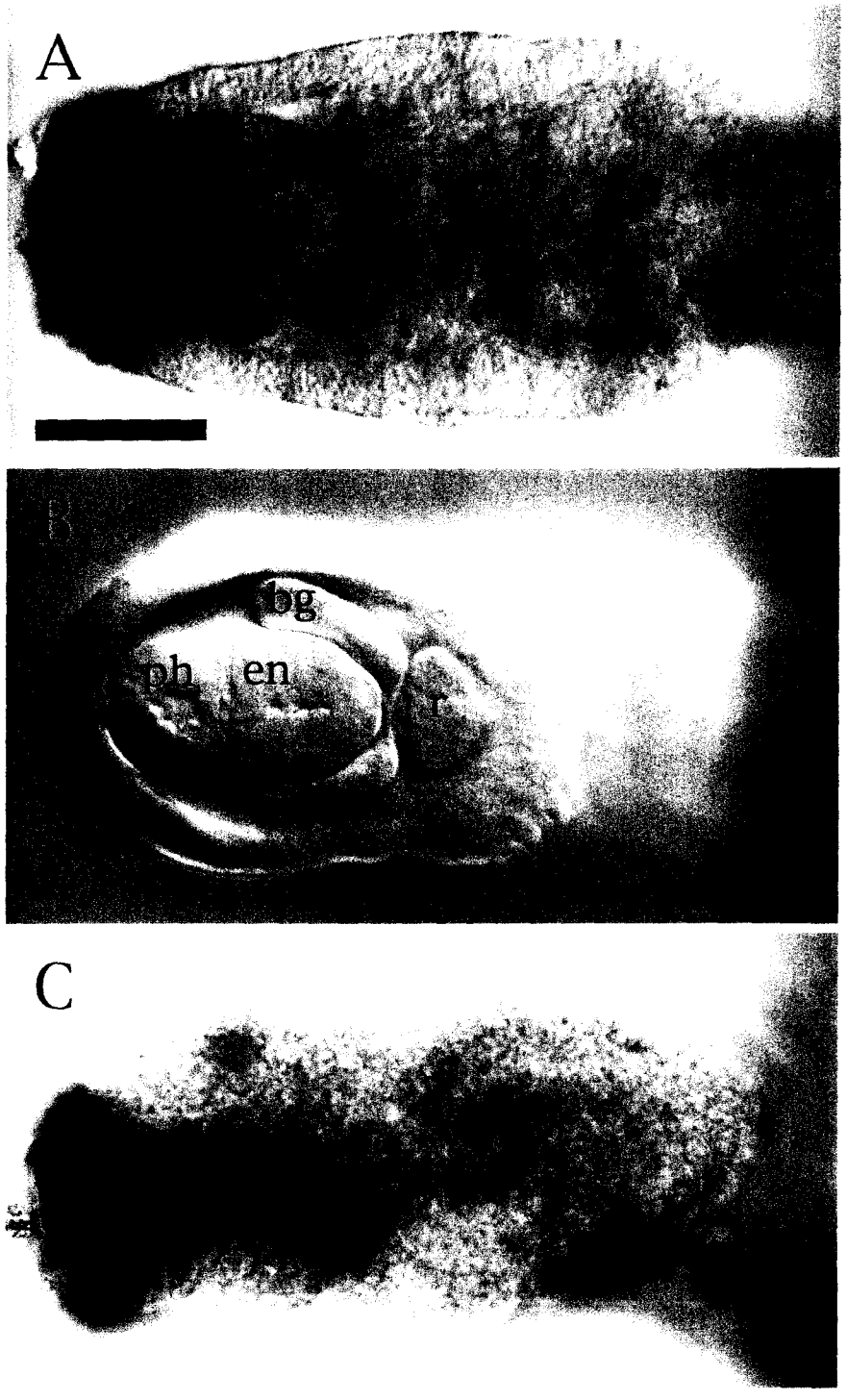


FIG 21. *Pitx* is expressed in the migrating buccal glands. (A) Ventral view of a 12 hour old hatchling, showing expression in a horseshoe of buccal gland tissue stretching around the posterior margin of the endostyle. B) Somewhat oblique ventral view of a live 10.5 hour old hatchling showing the buccal glands in an earlier stage of migration. The middle of the horseshoe becomes attenuated until eventually it splits and the presumptive glands contract forward to their positions on either side of the endostyle. (C) Lateral view of the same 12 hour hatchling. (bg, buccal gland; en, endostyle, ph, pharynx, r, rectum).

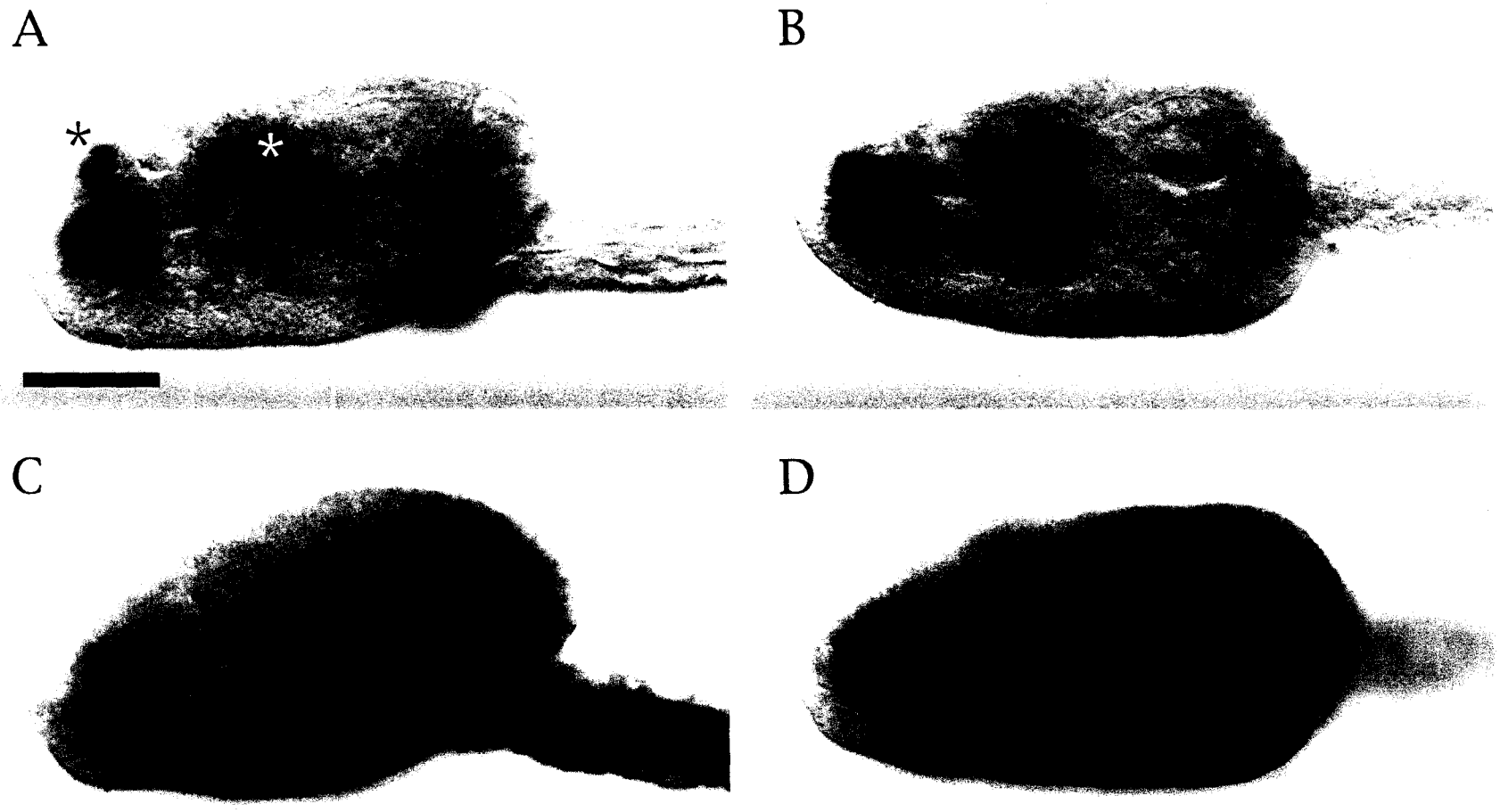


FIG. 22. Larvacean *Six3/6* duplicates are expressed in a nested pattern in 10.5 hour hatchlings. (A, B) Lateral and ventral views showing broad expression of *Six3/6a* in the rostral pharynx. *Six3/6a* is also expressed in the brain (black *), and in the esophagus (white *). (C, D) Lateral and dorsal views showing restriction of *Six3/6b* to just three cells in the roof of the pharynx on the right side. This is the position where the ciliary funnel meets the pharynx.

and forms a perforation into the pharynx (Delsman, 1912). *Six3/6a* and *Six3/6b* are expressed in the developing ciliary funnel. *Six3/6b*, as described above, is expressed only in the ventral part of the funnel that is continuous with the pharynx. In later stages, and finally at tailshift, *Six3/6a* expression marks the ciliary funnel from its pharyngeal base to its apex at the brain (Fig. 23). In addition, *Six3/6a* expression becomes apparent in the rostral nerve (Fig. 23, A-C).

Other Sites of Six3/6a Expression

Six3/6a is expressed in tissues outside the rostral domain. It is expressed in the midgut: medially in the ascending, rostral esophagus, and laterally in the dorsal, posterior margins of the developing spiracles (Fig. 22, A, B; Fig. 23, B, C). *Six 3/6a* may also have a role in the development of epidermal sensory structures. In young hatchlings, the gene is expressed strongly in two lateral trunk cells at the base of the tail in the region where the Langerhans mechanoreceptors develop (Fig. 24, A). This expression disappears, and at 12 hours two more spots appear on the lateral flanks, but dorsal and anterior to the Langerhans receptors (Fig. 24, B, C). Even at earlier stages, there appear to be sensory hairs forming from epidermal cells (Fig 24, D). In addition to these hairs and previously unreported bristles of the upper lip cells, there are potentially other sensory structures in the larvacean epidermis that have gone unrecognized.

Pax1/9 and Pax2/5/8 Expression in the Developing Gills

In addition to the expression of *Eya*, *Pitx*, *Six3/6a* and *Six3/6b*, two genes from the *Pax* gene family are expressed in the larvacean pharynx. *Pax2/5/8* is expressed

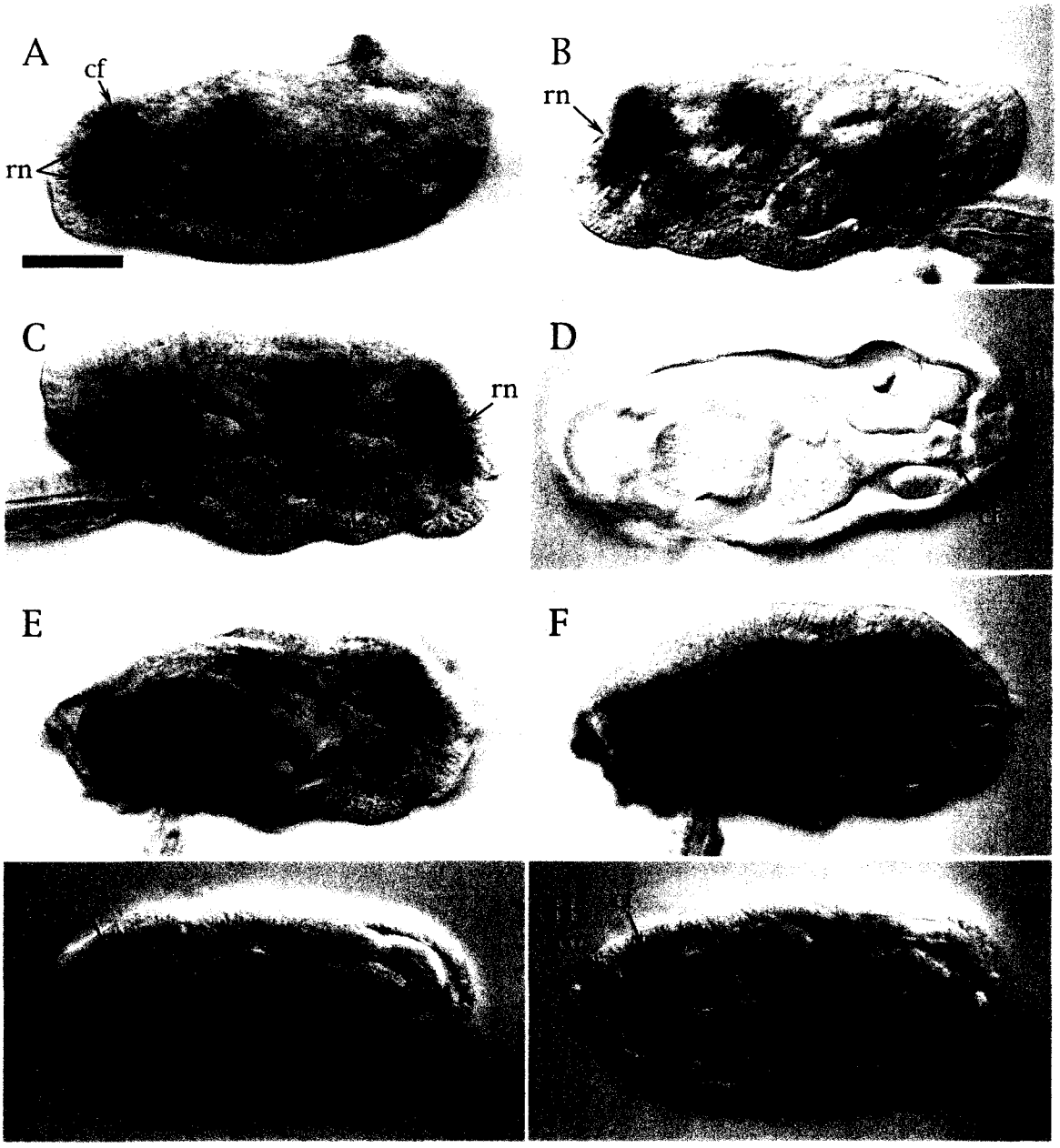


FIG. 23. *Six3/6a* is expressed in the ciliary funnel and the rostral nerve trunks. (A) A 12 hour old hatchling, in dorsal view, showing expression in the rostral nerves, the ciliary funnel. (B) Another 12 hour hatchling, in lateral view, showing the nerve trunk, pharyngeal roof, brain, and esophagus expression. (C) A 13 hour old hatchling, right side view. (D) An oblique dorsal view of a live tailshift stage (15 hour) animal to show the relative positions of the ciliary funnel (seen here in cross-section as it enters the pharynx from the dextro-dorsal side) and the rostral nerves, now detached from the epidermis. (E-H) Views of the same stage as in D. (E) *Six3/6a* expression has narrowed to the funnel by this stage, seen in a superficial, longitudinal view. (F) The only other expression domain besides the funnel is in the rostral brain. (G) A dorsal view. *Six3/6a* has broadened across the rostral brain, as compared with earlier stages. (H) A deeper dorsal view showing expression in the ciliary funnel. (cf, ciliary funnel; rn, rostral nerves) Scale bar, 20μm.

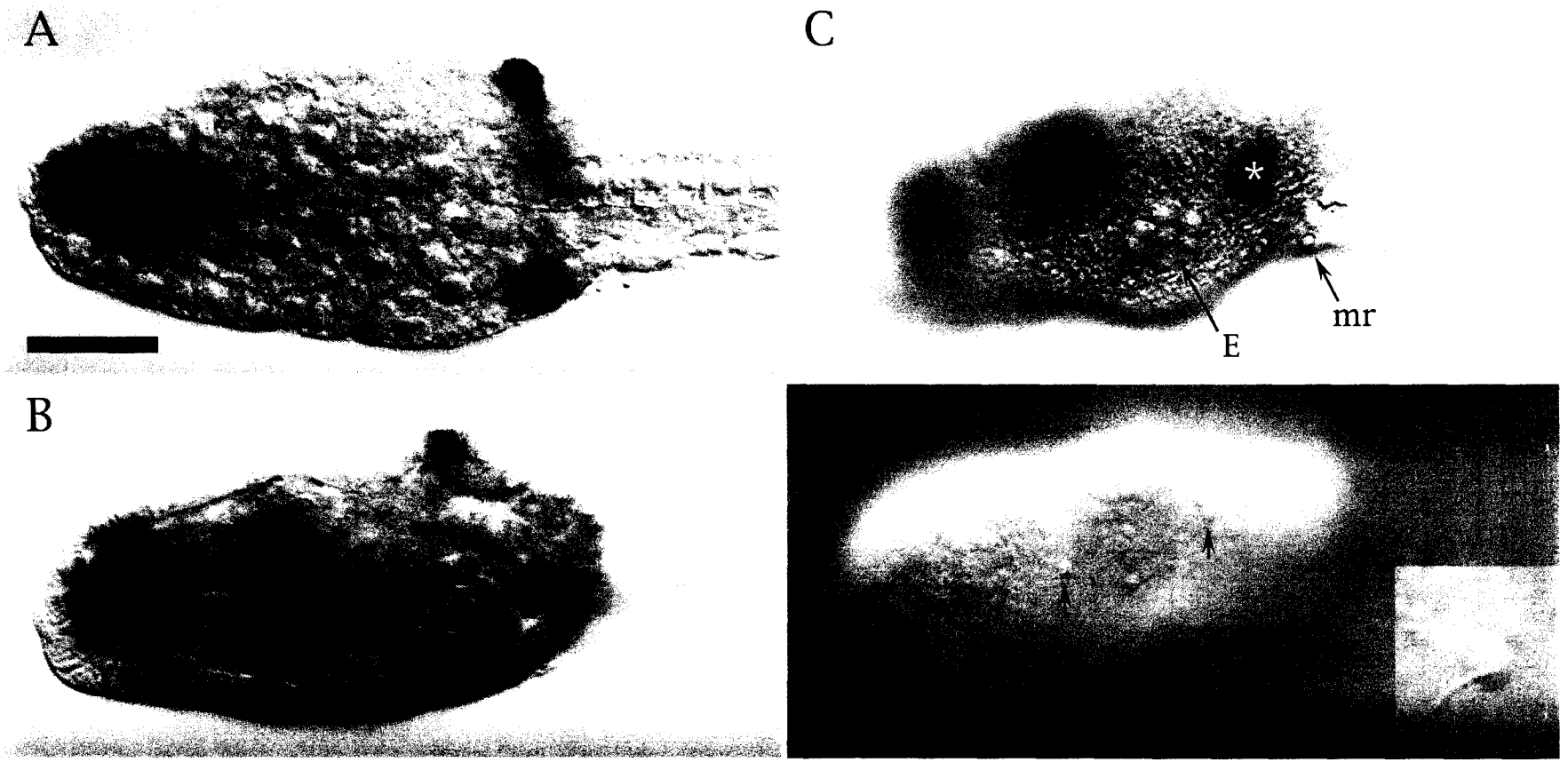


FIG. 24. *Six3/6a* may be expressed in epidermal sensory cells. (A) A young hatchling (6 hours old), ventral view. Two epidermal cells at the junction of the trunk and tail express *Six3/6a*. This expression may mark the same two cells that make the Langerhans mechanoreceptors, long ciliary hairs that protrude into the secreted house. (B) A 13 hour old hatchling, showing different patches of expression in the trunk epidermis. (C) A superficial lateral view of the same 13 hour old hatchling showing that this new expression (*) is not in the Langerhans cell (mr) or in another epidermal landmark of the lateral trunk, Eisen's oikoplast (labeled E). (D) The flank of a live 12 h old hatchling showing what appear to be sensory hairs (arrows). Inset shows and enlargement of one of these. There may be other, undescribed epidermal sensory structures, some of which could specializations of these early life history stages.

circumferentially in the rostral pharynx (Fig. 25, C, D; Fig. 26). There its expression overlaps that of *Pitx* and *Six3/6a*. In the developing spiracle, *Pax2/5/8* is expressed initially in the external part of the spiracle endoderm, while *Pax1/9* is expressed in the saddle between the spiracles in the floor of the pharynx and its expression extends somewhat into the dorsal part of the spiracle walls (Fig. 25). *Pax2/5/8* expression eventually encompasses the length of the spiracle or endodermal pouch during expansion of the gill lumen (Fig. 26 A, C).

Summary of Expression Patterns

Figure 27 diagrammatically summarizes the descriptions of the above expression patterns for *Eya*, *Pitx*, and the *Six* and *Pax* genes.

DISCUSSION

Sequence and Structure of Cloned Genes

Many of the *O. dioica* genes reported in this study diverge significantly from homologous sequences in other organisms. In particular, the two *Six3/6* genes as well as *Pax1/9* and *Pax2/5/8* are more divergent from vertebrate genes in the same subfamilies than are the *Drosophila* orthologues (Fig. 2, Fig. 6). The degree of divergence in the larvacean genes, even within domains conserved among non-chordates, suggests these larvacean genes are evolving very rapidly relative even to the genes of their ascidian relatives.

Drosophila and vertebrate *Six* genes fall into three subfamilies (Seo *et al.*, 1999). In a neighbor-joining gene phylogeny with *Six* family genes from other organisms, the

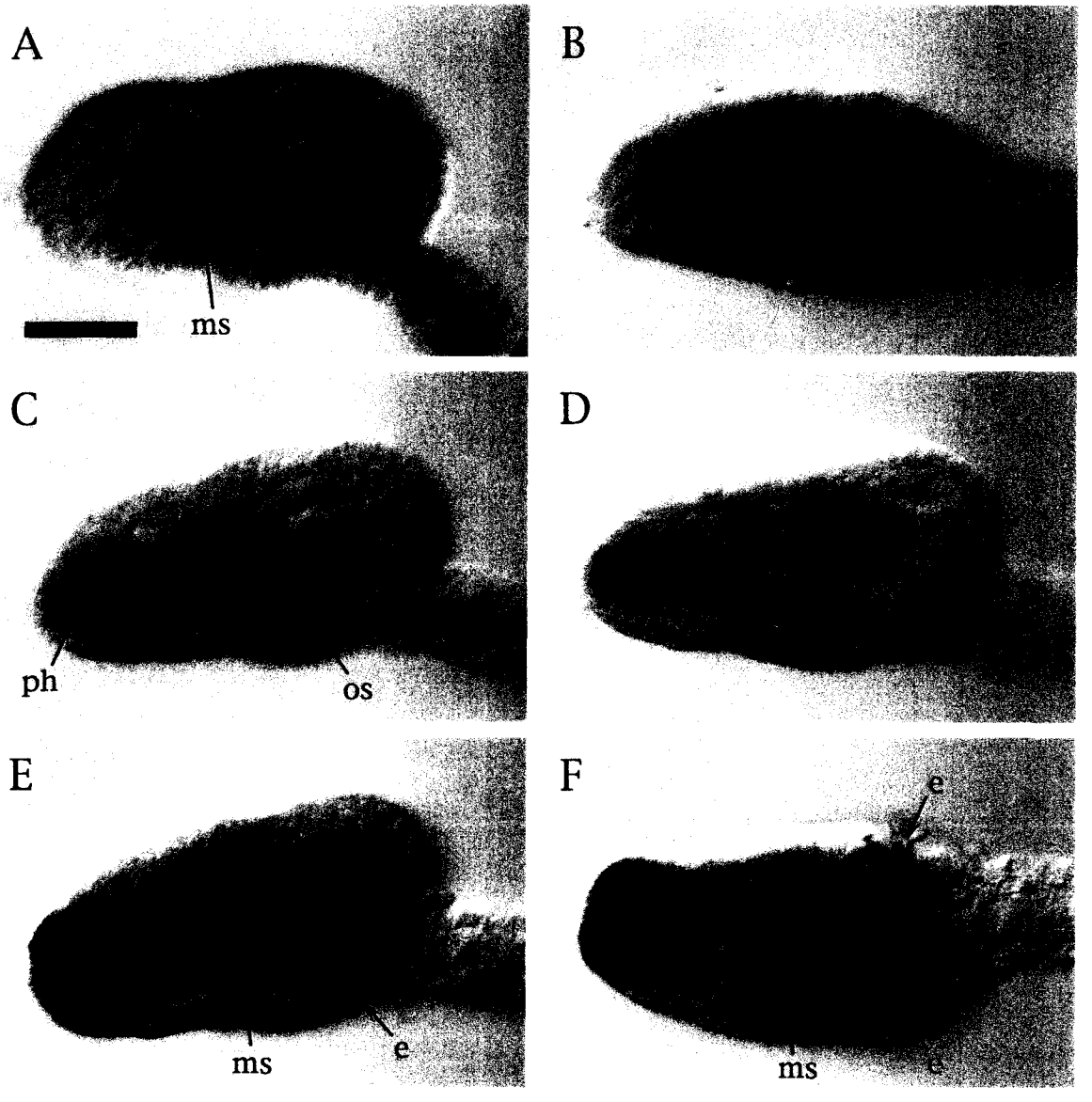


FIG. 25. Comparison of expression *Pax1/9*, *Pax2/5/8*, and *Eya* in the developing spiracles of 8 hour old hatchlings. (A, B) *Pax1/9*, lateral and ventral views showing endodermal expression in the medial part of the spiracles (ms). (C, D) *Pax2/5/8*, lateral and dorsal views showing endodermal expression in the outer part of the spiracles (os) and the rostral pharynx (ph). (E, F) *Eya*, lateral and ventral views showing expression in the medial spiracle endoderm (ms) as well in two ventral patches of epidermis (e) overlying the distal spiracle endoderm. Scale bar, 20μm.

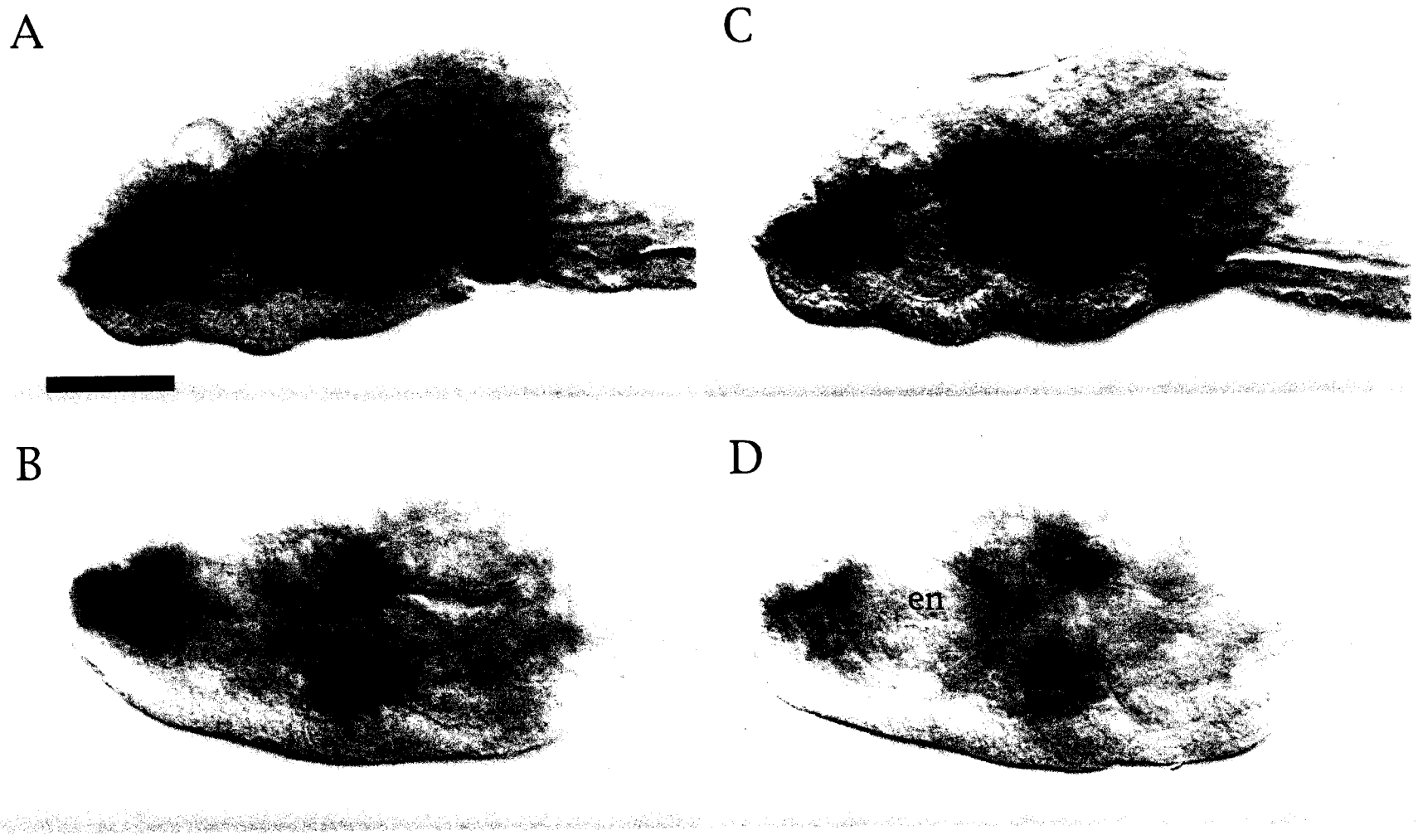


FIG. 26. Expression of *Pax2/5/8* 10.5 hour and 12 hour old hatchlings. Expression becomes stronger in the dorsal/medial parts of spiracle pouches, and remains in both the dorsal and ventral foregut. It is not detectable in the endostyle (en). (A, B) 10.5 hour, lateral and ventral views. (C, D) 12 hour, lateral and ventral views. Scale bar, 20μm.

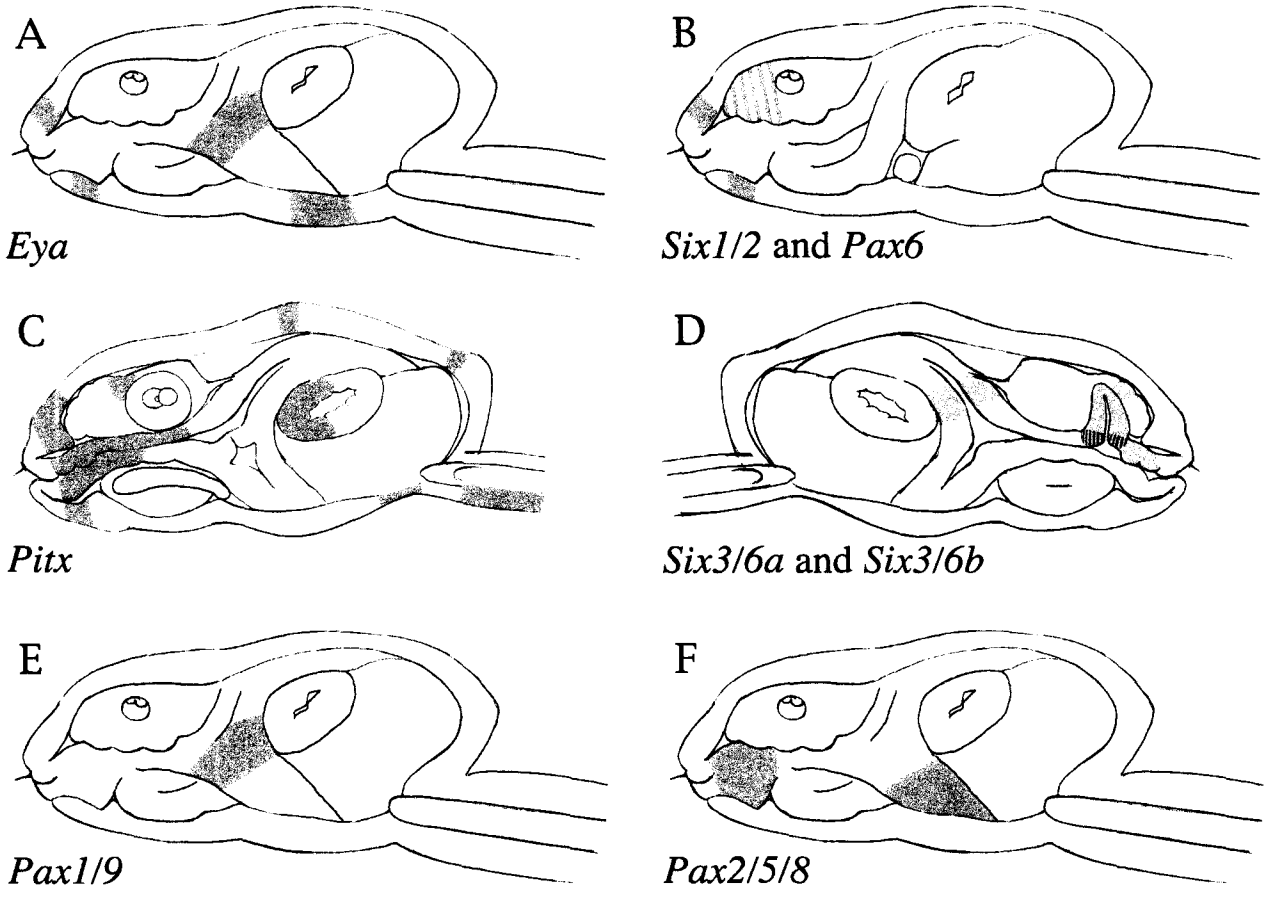


FIG. 27. Summary of expression patterns. (A) *Eya*: rostral epidermis, gill endoderm, ventral epidermis. (B) *Six1/2* (shaded): rostral epidermis. *Pax6* (striped shading): rostral brain. (C) *Pitx*: rostral epidermis, rostral pharynx, brain, migrating buccal glands, stomach and connection canal between stomach lobes, several epidermal domains of the trunk and tail. (D) *Six3/6a* (shaded): ciliary funnel, roof of the rostral pharynx, brain, esophagus, lateral trunk epidermal domains. *Six3/6b* (striped shading): ventral ciliary funnel. (E) *Pax1/9*: medial endoderm of the gill pouches. *Pax2/5/8*: distal endoderm of the gill pouches, rostral pharynx. A, B, E, F represent 8.5 hour hatchlings; C, D represent 12 hour hatchlings.

three larvacean *Six* genes cloned in this study group with and are likely orthologous to genes of only two of the three subfamilies (Fig. 2). One larvacean gene groups with *Six1* and *Six2* genes (Fig. 5), and the other two both group with the *Six3* and *Six6* genes. The two mammalian paralogues *Six3* and *Six6* probably result from one of the genome-wide duplications thought to have taken place in the vertebrate lineage (Kawakami *et al.*, 2000). It is unlikely that each larvacean *Six3/6* gene is singly orthologous to either of the vertebrate paralogues in the *Six3/Six6* clade. The two larvacean *Six* genes, therefore named *Six3/6a* and *Six3/6b*, probably represent an independent duplication in the larvacean lineage or in the urochordate lineage. At least one more *Six* gene, from the *Six4/5* clade, is likely to exist in *O. dioica*. It will be important to examine the expression of this gene, as its vertebrate orthologues are expressed in domains very similar to those of the *Six1/2* genes, including the olfactory and adenohypophyseal placodes.

The larvacean *Pitx* sequences cloned in this study were highly variable. Not only were unprocessed introns present in some clones, but also potential splice variants and polymorphisms affecting protein length and sequence. Two different 5' ends that were non-overlapping by more than 260 base pairs could have been transcribed by different promoters, as occurs in the transcription of mammalian *Pitx1* and *Pitx2* genes (Schweickert *et al.*, 2000; Tremblay *et al.*, 2000). Additionally, at one splice site is a coding sequence length difference not easily explained by known patterns of alternative splicing. The two possible splice acceptor sites are only 7 nucleotides apart in the aligned sequences; this proximity could cause mis-splicing, and it is unlikely that they occur in the same allele. The shift creates a stop codon 2 amino acids from the splice junction,

which would truncate the predicted protein by 86 amino acids. It is necessary to clone the coding region of *Pitx* from genomic DNA of several unrelated animals to determine if intron sequence polymorphisms can explain the shift of the splice acceptor site (Fig. 4).

The *Pitx* genes belong to the *Aristaless*-related subfamily within the *Paired*-class homeobox gene superfamily. Aristaless-related proteins share a characteristic sequence near the carboxyterminus, the Aristaless domain (reviewed in Meijlink, 1999). This highly conserved domain may act as an intrinsic inhibitor of DNA binding activity, with modulation by protein-protein interactions (Amendt *et al.*, 1998). There is no recognizable Aristaless domain in any *O. dioica* sequence cloned in this study. There are examples from other members of the Aristaless-related family that express alternate splice forms lacking the Aristaless-domain (Rao *et al.*, 1997), however, and this possibility cannot be ruled out for the larvacean gene. Cloning the *Pitx* genomic sequence will resolve this question.

Expression of Eya, Six1/2 and Pitx in the Rostral Epidermis

Six and *eyes absent* (*eya*) genes were first recognized in *Drosophila* for their importance in eye development. In vertebrates, *Six* and *Eya* paralogues participate both in eye development and in a number of other embryonic processes. The combined expression domains of the four *Eya* genes identified in mammals include the lens placode, otic vesicles, olfactory and other neurogenic placodes, Rathke's pouch, kidney, somites and hypaxial muscle precursors. The six mammalian *Six* paralogues are expressed in sum in the olfactory placodes, Rathke's pouch, the ventral forebrain, otic vesicles, branchial arches, fore- and midgut, somites, nephrotomes, limb mesenchyme

and other domains (reviewed in Kawakami *et al.*, 2000). The expression patterns of *Eya1*, *Six1* and *Six4* correspond with one another to a large degree, both in mammals and in amphibians. Eya proteins are thought to interact synergistically with other proteins including some Six family proteins, both in vertebrates and in flies, to regulate transcription of unknown downstream genes (Ohto *et al.*, 1999). Six proteins all bear a homeodomain and a conserved Six domain. Eya proteins have transactivation abilities and a conserved "Eya domain" presumed to be a protein interaction domain, but Eya has no recognized specific DNA-binding ability. Both the Six domain of So (fly Six1/2) and the Eya domain are necessary for interaction of these proteins (Pignoni *et al.*, 1997), and coexpression in a transfected cell line of vertebrate *Six2*, *Six4*, or *Six5* with *Eya* genes causes the translocation of Eya1, Eya2 or Eya3 proteins into the nucleus (Ohto *et al.*, 1999). Six3, however, did not effect translocation of any Eya proteins, and the authors suggest that Six3 may interact with other, unknown co-molecules. Similarly, in *Drosophila*, the *Six3* homologue *optix* does not synergize with *Eya* when ectopic eyes are induced (Seimiya and Gehring, 2000).

Pax6 forms part of the *Eya/Six* pathway in flies and probably in vertebrates. In *Drosophila*, eyeless (fly Pax6) directly binds an eye-specific enhancer in an intron of *So* (Chow *et al.*, 1999). So and Eya act together to feed back positively on the transcription of *eyeless* (Pignoni *et al.*, 1997), and murine *Eya1* and *Eya2* require the function of *Pax6* for their expression in the olfactory placode (Xu *et al*, 1997). Vertebrate *Pax6* is expressed in several of the tissues that express *Eya* and *Six* paralogues, including the

olfactory, adenohypophyseal, and lens placodes. *Pax6* is also important in the developing forebrain, but it is not required for expression of *Six3* (Oliver *et al.*, 1995b).

O. dioica Six1/2 and *Eya* expression domains appear to overlap in the rostral epidermis. Both genes are expressed in a ring of epidermal cells surrounding the presumptive mouth (Fig. 15, A, B; Fig 16). The overlap of *Six1/2* and *Eya* suggests that their protein products could interact in the same way their homologues do in vertebrates and flies. The conservation of this pathway, however, does not appear to extend to *O. dioica Pax6*. As in ascidians (Glardon *et al.*, 1997), *Pax6* is not expressed in the rostral epidermis in larvacean embryos. Its expression is limited to the brain rostral to the sensory vesicle and a small domain in the posterior, right side of the brain (Fig. 17). In contrast, amphioxus *Pax6* is expressed in the rostral epidermis, and Glardon *et al.* (1998) argue that sensory receptors that will form in the rostrum could be olfactory.

A third larvacean gene, *Pitx*, is expressed in an overlapping pattern with *Eya* and *Six1/2* in a ring of epidermal cells around mouth (Fig. 15; Fig. 18, B, C). Vertebrate *Pitx1* (pituitary homeobox 1) was originally identified as a transcriptional regulator of pro-opiomelanocortin in the anterior pituitary (Lamonerie *et al.*, 1996). *Pitx1* and *Pitx2* genes are expressed early in the stomodeum and their expression persists in the olfactory and adenohypophyseal placodes as they invaginate. The location and timing of larvacean *Six1/2*, *Eya* and *Pitx* expression in the rostral epidermis are appropriate to mark two cell groups that form by placode-like characteristics: an epidermal thickening that gives rise to neuronal cells which may delaminate. These are the ciliary cells of the ventral organ and the "rostral brain bulb" cells. Martini (1909) first recognized the nerve tracks that

connect the ventral organ to the brain via the ganglion-like swellings (bulbs), and referred to the organ as both an olfactory organ ("Geruchsorgan") and a taste bud ("Geschmaksknospe"). More recently, Bollner *et al.* (1986) analyzed the ultrastructure of the organ and the brain bulbs. They concluded that the somewhat dilated cilia and nervous connections of the receptor cells are consistent with them being primary, olfactory neurons. The authors' interpretation of the brain bulbs as homologues of the vertebrate olfactory bulbs revolved on the assumption that the larvacean bulbs grew out from the brain. Delsman (1912), however, witnessed the development of the bulb cells, which come to lie somewhat linearly along the bundled axons from the sensory and brain cells (Fig. 11, A); he described a row of epidermal cells on each side of the midline that sink in, separate from the epithelium, and ultimately make contact with the brain by a medial stalk (Fig. 11, A; Fig. 14, D). Consistent with Delsman's observations, expression of larvacean *Pitx* is seen first in the epidermal epithelium, and subsequently in the rostral bulb cells (Fig. 18; Fig. 20, A, B). It remains possible, however, that the rostral brain bulbs do not arise by delamination from the *Pitx* domain of the epidermis, but independently express *Pitx*. The best test of whether the *Pitx*-expressing cells of the epidermis give rise to the *Pitx*-expressing cells of the rostral brain bulbs would be to observe the delamination of these cells in a living animal as it expresses a reporter gene driven by the *Pitx* promoter.

I propose, based on the expression of *Eya*, *Six1/2* and *Pitx*, that the brain bulbs represent part of the ventral organ placode. The vertebrate olfactory placode gives rise not only to ciliary sensory cells, but to several other cell types, some of which are

ultimately located at a distance from the placode; these other cell types include supporting cells, two kinds of basal cells, ensheathing cells, and migratory cells, including those that make the terminal nerve. The complete list of placodally-derived cells includes at least three types of neuronal cells and at least seven types of non-neuronal cells (Farbman, 1992). Bollner *et al.* (1986) sorted the four cells of the larvacean rostral bulbs into 3 classes. One of these had no processes itself, but was postsynaptic to fibers from the brain. The other three gave rise to axons extending into the brain, and all were pre- and post-synaptic with fibers from both the brain and the ciliary sensory cells. One cell type ensheathed the bundle axons of both ciliary receptor and brain cells, similar to the way in which the Schwann (ensheathing) cells from the vertebrate olfactory placode bundle the unmyelinated axons of the receptor cells. The olfactory nerve is unique in two ways in vertebrates: the olfactory nerve Schwann cells originate in the placode itself rather than from neural crest, and, unlike in non-olfactory sensory nerves, Schwann cells ensheath axons in bundles, not singly (Farbman, 2000). Migratory neuronal cell types of the terminal nerve arise from the medial part of the vertebrate placode. The cell bodies of a GnRH-positive class of these migratory cells are embedded along the olfactory nerve in most vertebrates, and it has been argued that these act as neuromodulators of signals from the olfactory sensory cells (reviewed in Wirsig-Wiechmann, 2001). Some of the larvacean bulb cells may be inhibitory, as GABA-like immunoreactivity was found in a subset of them (Bollner, *et al.*, 1991). The axons of the larvacean ciliary cells do not appear to reach into the brain, but they make connections

with the brain processes with which they are bundled by the bulb cells (T. Bollner, personal communication).

There is no published information on *Eya*, *Six1/2* or *Pitx* expression from other urochordates, but the expression pattern of *Pitx* has been examined in the cephalochordate *Branchiostoma belcheri*. *BbPtx* expression makes a symmetrical crescent estimated to mark the anterior margin of the neural plate in late gastrulae, but its ectodermal expression after this stage quickly becomes restricted to the left side encompassing the region where the mouth, Hatschek's pit, and the duct of the enigmatic club-shaped gland will form in this asymmetric animal (Yasui *et al.*, 2000). Interestingly, although the rostral epidermis suggested to be olfactory is marked by *AmphiPax-6* in post-gastrula *Branchiostoma floridae* (Glardon *et al.*, 1998), the corresponding tissue does not appear to express this *Pitx* homologue at comparable stages in the congener *B. belcheri*. While *Eya* gene function appears to be conserved in fly and vertebrate eye development, fly *Pitx* (*D-Ptx1*) expression seems to share no developmental similarities with vertebrate gene homologues. Lack- and gain-of-function mutants did not give a developmental phenotype detected by Vorbrüggen *et al.* (1997), and the authors conclude that, unlike the vertebrate *Pitx* genes, the fly orthologue is not likely to be involved in embryonic patterning.

Other Sites of Eya and Pitx Expression

Larvacean *Eya* is expressed early in two patches of ventral epidermis where the gill openings will form and in the gill endoderm (Fig. 25, E, F; Fig. 16, A-C). In mouse embryos, *Eya1* is expressed in the pharyngeal pouch endoderm, and *Eya2* is expressed in

the branchial arch ectoderm (Xu *et al.*, 1997). Because larvacean *Eya* is orthologous to both mouse genes, it is possible that the larvacean gene retains an ancestral pattern amounting to the combined branchial arch expression of murine *Eya1* and *Eya2*.

Several different putative isoforms of the larvacean *Pitx* gene were cloned in this study. The variety of transcripts is mirrored by the complexity and dynamism of *Pitx* expression in larvacean development. In this *in situ* analysis the *Pitx* riboprobe may detect only one of the potential isoforms reported here or it may detect them all. It will be possible to make probes against the variable 5' regions to test whether these isoforms are differentially expressed. One isoform of vertebrate *Pitx2* is part of the cascade that sets up the left-right body axis, affecting such processes as gut and heat looping (Schweickert *et al.*, 2000). Amphioxus *Pitx* becomes restricted during neurula stage to the left side of the embryo, in ectoderm, endoderm and the developing somitocoels (Yasui *et al.*, 2000), and the authors argue for a primitive role for *Pitx* in establishing asymmetry. Larvacean anatomy displays striking asymmetries, in gut looping, the position of the heart, the rotation of the tail, and the brain. No early asymmetry was detected in the expression of the larvacean *Pitx* gene, however. *Pitx* is expressed in muscle cells on both sides of the tail in tailbud stage larvacean embryos (Fig. 18, A). Expression of the larvacean gene is more like that of vertebrate *Pitx1*, which is expressed symmetrically in posterior lateral plate mesoderm (Lanctôt *et al.*, 1997). A careful analysis of the expression of different larvacean isoforms may resolve this conflict.

Only two kinds of tissue have been described as migratory in larvacean development: the subchordal cells of the tail and the buccal gland cells of the anterior

trunk. Delsman (1912), from his observation of live embryos, concluded that both were made up of endodermal strand cells that retract into the trunk in early stages of the extension of the tail. Hirano and Nishida (2000) regard the ascidian caudal endoderm, or endodermal strand, as homologous to the vertebrate intestine. The buccal gland cells migrate forward, and fork around the right and left sides of the endostyle (Fig. 21, B). Ultimately, they split into separate right and left halves that make up the buccal glands, cells that secrete bioluminescent material into the house in which the larvacean lives and feeds. This migratory tissue expresses *Pitx* once it is in a horseshoe shape around the endostyle (Fig. 21, A, C). Probably this represents a co-option of *Pitx* expression to this highly specialized larvacean tissue, although vertebrate *Pitx2* is expressed in the midgut intestine, recalling Delsman's endodermal strand observation.

Expression of larvacean *Pitx* in the left developing stomach may represent a conserved pattern with that of vertebrate *Pitx1* expression in the stomach and other gut compartments (Lanctôt *et al.*, 1997). *Pitx2* is expressed in regions of the vertebrate forebrain (Muccielli *et al.*, 1996), and larvacean *Pitx* is expressed in a pair of dorsal cells to the right of the sensory vesicle. A number of other *Pitx* expression domains are mysteries. Cells in the proximal tail fin edges (Fig. 19, B) and at the tip of the tail (Fig. 18, C) are both locations described by Lohmann (1933) as bearing putative sensory structures in various larvacean species, though this idea has not since been confirmed.

Expression of Larvacean Six3/6 Paralogues in the Ciliary Funnel

The larvacean *Six3/6* duplicates exhibit a nested pattern of expression in the stages examined; *Six3/6a* is more broadly and diversely expressed than is *Six3/6b*. Both

genes are expressed in the primordium of the ciliary funnel, but *Six3/6a* is expressed along the dorso-ventral axis of the funnel while *Six3/6b* is expressed only at the level of its ventral opening into the pharynx (Fig. 22, C, D). *Six3*/6a is expressed in additional domains: in the roof of the rostral pharynx (Fig. 22, A, B), in the developing rostral nerves (Fig. 23, A-C), in the brain (Fig. 22, A; Fig. 23, B, C, F, G), and in pairs of cells in the trunk epidermis (Fig. 24, A-C). Nested expression is also seen among *Six3* family paralogues in vertebrates, in which *Six3* genes are more broadly expressed than *Six6* (Jean *et al.*,1999; Ghanbari *et al.*, 2001). Vertebrate *Six3* genes are expressed in both the adenohypophyseal primordium and in the ventral forebrain, including the hypothalamus (Oliver *et al.*, 1995b; Kobayashi *et al.*, 1998; Ghanbari *et al.*, 2001).

Like early developmental stages of the vertebrate adenohypophysis, the larvacean ciliary funnel forms a pouch that connects the pharyngeal roof with the brain. The larvacean funnel, however, is an outgrowth of the brain (Delsman, 1912). The neural origin of the ciliary funnel and the apparent secretion of a granular substance from the funnel tip into the haemocoel, prompted Holmberg (1982) to propose that this larvacean organ represents a homologue of both the adenohypophysis and neurohypophysis. It was long thought that the glandular cells of the vertebrate pituitary were products of the stomodeal ectoderm. More recently, fate-mapping analyses in chick (Takor and Pearse, 1975) and amphibians (Kawamura and Kikuyama, 1992) have shown that the anterior pituitary placode is a derivative of the apical neural ridge that becomes separated from the neural plate before neural tube closure. These studies found that immediately adjacent neural plate regions give rise to the hypothalamus and the adenohypophysis, and the

authors argue that the pituitary should be regarded as a single developmental entity rather than as a composite organ. Expression of vertebrate *Six3* in both the hypothalamus and Rathke's pouch is a continuation from the gene's earlier expression in the anterior neural plate (e.g., Oliver *et al.*, 1995b). Expression of *Six3/6* homologues in the larvacean ciliary funnel supports this organ's homology with the vertebrate adeno- and possibly also neurohypophysis.

In addition to the *Six*es, vertebrate *Pitx* genes are also expressed during pituitary development. Vertebrate *Pitx1* is an early marker of the stomodeum and the foregut endoderm, and its expression persists in the olfactory and adenohypophyseal epithelia into late development (Lanctôt *et al.*, 1997). Larvacean *Pitx* is also expressed in the rostral foregut (Fig. 19, A; Fig. 20, A), where it overlaps with the expression of *Six3/6a*, *Six3/6b* and also *Pax2/5/8* (discussed below). However, *Pitx* does not appear to be expressed in the funnel between the pharyngeal wall and the brain. Amphioxus *Pitx* is expressed in the epithelium of the developing Hatschek's pit, the cephalochordate adenohypophyseal candidate, and in the ectoderm that includes this structure and the mouth (Yasui *et al.*, 2000). Vertebrate *Pitx1* is expressed in the invaginating stomodeum, marking what Lanctôt *et al.* (1997) term the most anterior domain or "ectomere" of the embryo; it is possible that the absence of larvacean *Pitx* in the dorsal ciliary funnel stems from the funnel outgrowing directly from the brain rather than inpocketing from within the stomodeal ectomere. In contrast, the larvacean ventral organ and bulbs are separate from the neural tube, and their cellular behavior and molecular characteristics may be

more similar to ANR tissue that invaginates in stomodeum formation. A larvacean fate map analysis of the anterior neural plate will be necessary to resolve this question.

Larvacean *Six3/6a* is expressed in domains outside the ciliary funnel. The rostral nerves, a combination of axons from the brain and the bulb cells delaminating from the epidermis, express *Six3/6a* at the stage when the separation of the bulb cells from the rostral epidermis is visible (Fig. 23, A-C). This expression is consistent with the expression of vertebrate *Six3* class genes in the olfactory bulbs and in the olfactory placode (e.g., Ghanbari *et al.*, 2001). Vertebrate *Six1/2* and *Six3/6* class genes overlap in the olfactory placode, the olfactory bulbs, and Rathke's pouch. Larvacean *Six1/2* and *Six3/6a*, however, are expressed in complementary rather than overlapping domains in the olfactory and adenohypophyseal candidates. Is there an evolutionary scenario that could result in the observed difference between vertebrate and larvacean expression patterns? Following the gene duplication events that created the three *Six* paralogues from an ancestral gene, and following the subsequent divergence of the larvacean and vertebrate lineages, some overlapping gene functions could have become sorted into separate larvacean *Six* paralogues, while these ancestral functions continued to overlap among the vertebrate paralogues.

Other Sites of Six3/6a Expression

Six3/6a expression in cells at the rostral boundary of the brain is consistent with expression of its homologues in the vertebrate forebrain, but the identity of these cells is otherwise unknown (Fig. 22, A; Fig. 23, B,C,F,G). On the other hand, expression of *Six3/6a* in the esophageal/pharyngeal endoderm or in isolated epidermal cells of the trunk

does not directly parallel the expression of any known vertebrate *Six* gene (Fig. 22, A; Fig. 23, B, C). Vertebrate *Six3*-related genes are restricted to domains in the head (reviewed in Kawakami *et al.*, 2000). Murine *Six2* is expressed at the esophageal/pharyngeal boundary, but only in the mesodermal, not the endodermal component (Oliver *et al.*, 1995a). Expression of a larvacean *Six* gene in a developing secondary sensory receptor of the trunk (the Langerhans receptor) and possibly other epidermal receptor types could represent a novel deployment of this homeobox gene in the urochordate lineage. In the ascidian fate map, the pharynx and anterior neural tube share a close cell lineage; at the "tissue-restricted stage" (a 110 cell embryo), one cell pair will still give rise to part of both the pharynx and anterior neural tube (Nishida, 1987). *Six3/6a* expression in the dorsal pharynx could reflect this kind of close association of the pharyngeal cell lineage with the brain in the larvacean embryo.

Expression of Pax2/5/8 and Pax1/9 in the Pharynx

Pitx2 marks the developing otic vesicles in mouse (Nornes *et al.*, 1990). The discovery that an ascidian *Pax2/5/8* homologue is expressed in the ectodermal thickenings of the atrial siphon primordia led to the proposal that these are homologous to the vertebrate otic placodes (Wada *et al.*, 1998). This was the first report of molecular evidence for a urochordate placode homologue. The atrial primordia are initially paired thickenings that then fuse in the dorsal midline ultimately to make a single siphon. After metamorphosis, water flows through the ectodermally lined atrium, or exhalent siphon, after it has passed through the pharyngeal gill basket. The expression of *Pax2/5/8* in the pharyngeal endoderm and overlying ectoderm of amphioxus led Kozmik *et al.* (1999) to

propose an alternative hypothesis, that *Pax2/5/8* was ancestrally important for the formation of gill slits, perhaps by causing changes in cell adhesive properties to promote the fusion of the endodermal and ectodermal layers.

Larvacean *Pax2/5/8* is expressed in a pattern resembling the amphioxus *Pax2/5/8* expression in developing gills. Initially, there are two spots of expression in the ventral ectoderm, perhaps marking where this ectoderm overlies presumptive pharyngeal endoderm (not shown). Later, although the ectodermal expression fades, the endodermal pharyngeal pouches express *Pax2/5/8*, initially most strongly in the distal segment (Fig. 25, C, D). This endodermal domain is complementary to the expression of larvacean *Pax1/9*, which marks the gill pouches more medially as well as the pharyngeal saddle between the two pouches (Fig. 25, A, B). In vertebrates, *Pax1* and *Pax9* are important for the development of the pharyngeal endoderm and its derivatives (Peters *et al.*, 1995; Peters *et al.*, 1998). In lamprey, *Pax9* is expressed in pharyngeal pouch endoderm. Similar to the larvacean pattern, the lamprey gene becomes medially restricted and is excluded from the lateral walls, which the authors suggest are the perforating part of the endoderm (Ogasawara *et al.*, 2000). This conserved gill expression of *Pax1/9* genes may be more ancient than the chordates; Ogasawara *et al.* (1999) showed that a hemichordate *Pax1/9* homologue is expressed in the endodermal part of the adult gill. They also report the expression of *Pax1/9* homologues in the gill of two ascidian species. Just as in amphioxus, larvacean *Pax1/9* appears to be expressed only in pharyngeal endoderm (Holland and Holland, 1995); this is unlike expression of the vertebrate homologues of these genes, which are expressed, also in segmental mesoderm and the dorsal neural tube.

Amphioxus *Pax2/5/8* may be involved in the development of another endodermal structure, the endostyle. Kozmik *et al.* (1999) relate the expression of this gene in parts of the developing endostyle to thyroid expression of mammalian *Pax8* (Plachov *et al.*, 1990) or amphibian *Pax2* (Heller and Brändli, 1999). Endostyles occur in both cephalochordates and urochordates and are thought to be homologous to the vertebrate thyroid. The larvacean endostyle, however, does not express *Pax2/5/8* at any stages examined (Fig. 25, C; Fig. 26, A-D).

Conclusions

The "new head" hypothesis for the origin of vertebrates (Gans and Northcutt, 1983) is currently under reevaluation in light of information from urochordate and cephalochordate molecular embryology. While it is still widely held that neural crest and placodes were evolutionary innovations unique to the vertebrate lineage, molecular tools make it possible to test this hypothesis, as well as older hypotheses about the commonality of morphological characters among the chordates. This study of the larvacean *Oikopleura dioica* has focused on candidate homologues of the vertebrate olfactory and adenohypophyseal placodes, and on several genes that are important for the development of these vertebrate tissues.

Primary sensory receptors, which have their cell bodies in the periphery and send axons into the central nervous system, are common in invertebrates, but unusual for vertebrates. The primary sensory cells of the olfactory epithelium are an exception, and it is because of this that they are sometimes regarded as the 'primitive' part of the vertebrate sensory system. The larvacean ventral organ system develops like the

vertebrate olfactory placodes, from paired epidermal thickenings that give rise to peripheral primary receptors and several other cell types. I propose that these larvacean placodes include not only the ventral organ but also *Six1/2*, *Eya* and *Pitx* positive cells that will make the so-called bulb cells. These bulbs, according to Delsman's (1912) and my own observations, develop by ingression from a thickening in the epidermis rather than by outgrowth from the brain. Larvacean *Six1/2* and *Eya* appear to be co-expressed in the developing ventral organ system; therefore the protein products of these genes may interact in the same way their fly and vertebrate orthologues do, as components of a putative transcription factor complex. The *Pax6* homologue I studied cannot act as a transcriptional regulator of *Six1/2*, however, as the two genes are expressed in different tissues. This means that part of a network likely conserved between flies and vertebrates has diverged in *O. dioica*. A duplicate *Pax6* gene that regulates *Six1/2* could exist in the larvacean genome. The presence of *Six3/6* duplicates suggests that, despite the extremely reduced genome size of approximately 72 megabases (Seo *et al.*, 2001), some gene families may be amplified independently in larvaceans.

Gorbman (1995) theorized that in the ancestor of the vertebrates, the olfactory organ and pituitary were components of a single organ that would mediate and synchronize sexual maturation and behavior through detection of environmental chemical cues. He proposed that amphioxus could represent such an evolutionary state, citing that Hatschek's pit is situated to sample the surrounding seawater directly. While larvacean anatomy does not appear consistent with a single, combined olfactory and adenohypophyseal organ, it does conform to the correlate of Gorbman's hypothesis: the

olfactory organ and adenohypophysis share an origin more ancient than the vertebrates. The ciliary funnel is an outgrowth of the larvacean brain that forms a tubular connection with the roof of the pharynx. Although the neural origin of both the larvacean funnel and its presumed ascidian homologue, the neurohypophyseal duct, has seemed inconsistent with their possible homology to the adenohypophysis in vertebrates or to Hatschek's pit in amphioxus, this conflict may dissolve in light of new ideas about the vertebrate fate map. Kawamura and Kikuyama (1992) showed that the adenohypophyseal primordium is not part of the stomodeal ectoderm. The amphibian equivalent of Rathke's pouch starts in the ANR, loses contact with both the neural plate and the ectoderm, and segregates rostrocaudally to form cells of the preoptic hypothalamus and also of the adenohypophysis. The adenohypophyseal primordium integrates into the pharyngeal roof and reestablishes contact with the hypothalamus to form the pituitary. This pattern of morphogenesis could explain Gorbman's observation that the hagfish adenohypophysis is recognizable at a stage before the formation of the stomodeum (Gorbman, 1983). I propose that the larvacean ciliary funnel could represent either a derived or a primitive state in which homologues of the adenohypophyseal and hypothalamic primordia never separate during neural tube closure and remain connected through subsequent development of the funnel. Cell marking in larvacean embryos will be required to test the prediction that the ciliary funnel and ventral organ primordia arise as neighboring domains in the anterior neural plate.

Cupular organs in the atrial siphon wall support of the homology of the ascidian atrial primordia with vertebrate otic placodes; these ciliary receptors resemble the hair

cells of the ear and lateral line (Bone and Ryan, 1978). Given the expression of *Pax2/5/8* in the developing gill of larvaceans and cephalochordates, however, it seems that any evolutionary scenario that includes the homology of ascidian atrial primordia with vertebrate otic placodes must contend with a correlated homology of otic placodes and gill openings. A reasonable alternative hypothesis is that the vertebrate otic placodes independently co-opted the function of the *Pax2* gene.

Caveats

Urochordates and vertebrates have been separated for over 530 million years, and their body plans show a high degree of divergence. Making homology comparisons, therefore, especially on the basis of the expression of a small number of genes, requires caution. When the larvacean ventral organ and ciliary funnel were proposed as homologues of the olfactory epithelium and adenohypophyis, nothing was known of their molecular characteristics, and the argument was based entirely on their topographical relationship to the brain and pharynx, and on the fine structure and arrangement of their cells. All of these data, taken together, suggest the homology of these organs to specific vertebrate counterparts. However, alternative interpretations are possible. In *Drosophila*, the *Six* genes are proposed generally to be regulators of head development; *So* is important for development of the optic system, while *D-Six3* and *D-Six4* are expressed in the developing brain, stomatogastric nervous system, and several sensory organs (Seo *et al.*, 1999). Larvacean expression of *Six* homologues, therefore, could represent the functions of *Six* genes in head sensory structures conserved from the common ancestor of arthropods and chordates. Alternatively, larvacean *Six* expression might merely be the

product of cooption of *Six* genes into novel larvacean structures. A third possibility is that the larvacean ventral organ and/or the ciliary funnel are placodes in the vertebrate sense, but they represent independent placode lineages from those that arose in the vertebrates. In other words, the larvacean organs could have descended from a placode precursor in the common ancestry of the vertebrates and urochordates, but do not directly correspond to any one modern vertebrate placode. Evidence supporting or against a direct correspondence of larvacean and vertebrate organs will depend in part on determining the fate map positions of the larvacean organs and their physiological functions; approaches for the latter might include the cloning and expression of olfactory receptor genes and immunohistochemical analysis with antibodies to adeno- and neurohypophyseal hormones.

Finally, there is the possibility that the larvacean ventral organ and ciliary funnel are homologous to vertebrate olfactory and pituitary, in that they derive from the same embryonic structures in the last common ancestor, but that their functions may have diverged significantly from their predecessor organs in the common ancestor. Larvaceans have a very rapid life cycle, a small repertoire of recorded behaviors, and a relatively low complexity of tissue types commensurate with extremely small body size. The vertebrate olfactory placode gives rise to at least ten neuronal and non-neuronal cell types, and the adenohypophysis generates five classes of neuroendocrine cells that affect functions as diverse as germ cell maturation, reproductive behavior, and metabolism in the liver. It is unreasonable to expect to find all the components of the vertebrate olfactory and pituitary system in miniature form in *O. dioica*. The synthesis of data from molecular embryology

and other comparative approaches, however, has allowed us to detect several levels of apparent homologies between the organs. Further work analyzing additional genes, tracing cell lineages, and investigating organ development and function may tell us with certainty whether the vertebrates can continue to claim placodes as their own.

CHAPTER IV

CONCLUSION

This dissertation is built upon the excellent morphological studies of classical biologists such as E. Martini, H. C. Delsman and H. Lohmann, and on the more recent efforts of anatomists and physiologists such as R. Olsson, T. Bollner, K. Holmberg, and C. Galt. The research of these authors describes many aspects of larvacean biology. Missing from this body of work, however, was the application of molecular techniques in larvacean embryos. The now commonplace tools of gene cloning and *in situ* hybridization in embryos can test the predictions of long-standing hypotheses on evolution and comparative anatomy of the chordates. This dissertation, which focused on the embryology and evolution of chordate innovations, has contributed the first molecular embryology study on a member of the urochordate class Larvacea, and provides a seed for future investigations into the origin of vertebrate placodes.

The larvacean lineage, though it may have branched early within the urochordate subphylum relative to sister classes Ascidiacea and Thaliacea, clearly does not represent an ancestral condition either for the urochordates or for the chordates as a whole. Like any modern taxon, larvaceans display a number of interesting synapomorphies, including novel tissues like the oikoplastic epithelium, which secretes the house. Other synapomorphies include topographical rearrangements like the 90° twisting of the tail relative to the trunk, and the apparent bilateral asymmetry of the brain. Larvaceans may be remarkably derived at the molecular level; the sequence divergence of *O. dioica* genes cloned in this study relative to homologous genes in other organisms implies that some

proteins in this larvacean are evolving rapidly. Despite changes in protein sequence, the expression patterns of many of these genes appear conserved relative to expression of orthologues in other chordates. It is crucial to our understanding of homology relationships to learn whether homologous genes that mark tissues in two organisms are performing the same biological function in those tissues. Now that the expression patterns of several important larvacean gene groups have been studied, it is necessary to test the function of these genes. This question can be approached in several ways, including: 1) determining if larvacean proteins can rescue wild-type phenotypes in vertebrates mutant for the orthologous gene, which result would suggest conservation of the structural properties of the protein, 2) blocking the native function of a larvacean gene in larvacean embryos to test phenotypic predictions based on our knowledge of the function of the gene's product in other organisms, 3) determining whether the larvacean gene is part of a conserved pathway by introducing a reporter construct driven by the larvacean gene's promoter into a vertebrate embryo and observing whether expression pattern is conserved.

In addition to experimental investigations such as those listed above, there are other questions about basic larvacean biology that can be addressed with simple observation. There is no fate map for larvacean development. Delsman (1910) based his assumptions about cell lineage in *O. dioica* on comparisons with the stereotypical patterns of cleavage and blastomere fate in ascidians described by Conklin (1905). Although the expression of the *O. dioica T gene* did appear to confirm one detail of Delsman's fate map, there are likely to be many differences between larvacean and ascidian fate maps. Delsman himself noticed that larvacean embryos begin to gastrulate a full cleavage earlier than do ascidian embryos, with the result that landmark stages of development, such as neural plate stage, occur in a larvacean embryo with many fewer

cells than exist in the corresponding stages of ascidian embryos. Given the small number of cells in the larvacean embryo of any stage, the optical clarity of the embryos, and their extremely rapid development, it will be possible to mark individual cells and not only discover what adult tissues those cells become, but also observe the morphogenetic movements of the marked cells during development by means of time-lapse microscopy. A larvacean fate map will answer questions raised in this dissertation about the developmental origin of the ventral organ system and the ciliary funnel, presumptive larvacean homologues of vertebrate placodes.

The work in this dissertation has shown that *O. dioica* is a tractable subject for the study of morphogenesis and gene expression in larvacean embryos. It also suggests that there is significant correspondence between larvaceans and vertebrates with respect to gene expression, including the expression of several genes essential to the development of supposed vertebrate innovations, the placodes.

APPENDIX

A

Culturing Notes

Abundance of *O. dioica* in sampled near-shore plankton varied greatly, not only between seasons, but also from year to year. In general Spring and late Fall are optimal times for collecting on the central Oregon coast, when the concentration of chain-forming diatoms is lower. A low abundance of spiculated diatoms that foul larvacean houses facilitates using a phytoplankton net for the gentler collection of larvaceans still in their houses. Plankton collections were carried out either from the dock (Charleston, Oregon, or Bamfield, B.C.) at high tide, or from a boat up to 5 miles offshore. Whole plankton was collected by slow, vertical towing using a 35-micron nitex plankton net with a 40 cm opening and a threaded glass cod end. One-liter plankton samples were sometimes diluted immediately into gallon jars so that larvaceans could deploy new houses with less interaction with other organisms and particles. Carboys of dipped seawater were collected at the same time from surface water. Predators such as comb jellies and jellyfish were removed from cultures, and all seawater and larvaceans were stored in gallon glass jars in a 10.5°C or 12°C incubator with constant broad-spectrum illumination. Flakes of 1-hexadecanol (Aldrich) were added to culture jars to create a non-toxic, hydrophobic barrier at the surface; this prevents animals from becoming trapped in the surface tension.

Larvaceans were cultured in natural seawater and their diet was supplemented with cultured microalgae and with cryopreserved Nanochloropsis sp., *Isochrysis* sp. *(T-iso)*, and *Tetraselmis chui (*strain *PLY429)* (glycerol cryopastes from Brineshrimp Direct).

Large animals with swelling, but ambiguously sexed gonads were placed in pairs in small finger bowls and allowed to mature at 10.5 or 12°C. There may be a chemical or pheromonal interaction between maturing animals that results in the ripening of one male and one female when an indeterminately sexed pair is cultured in a small volume of water. Formal experiments are needed to test the hypothesis of a pheromonal interaction. It may be significant that *O. dioica* is the only dioecious larvacean described. All congeners of *O. dioica* are thought to be hermaphroditic. This suggests the possibility that environmental, rather than genetic factors determine sex in *O. dioica*, and that an individual is competent to develop either gonad type.

Sperm was collected by rupturing ripe male gonads in as little water as possible. Sperm could then be stored at 8 to 9°C for up to 24 hours. The average number of spermatozoa per male, calculated from hemacytometer counts from over 30 ripe males, was approximately 2.25×10^6. Animals ripen at a broad range of sizes, however, which greatly influences the number of gametes produced. Females were allowed to spawn naturally; in still water, the shed eggs will sink directly to the bottom of the dish and accumulate in a closely packed monolayer. Eggs were transferred to small tissue culture Petri dishes, where they were fertilized with a dilute sperm suspension and then rinsed after 10 minutes with several water changes. Hatchlings will stick to polystyrene dishes unless the plastic is coated for tissue culture. Swimming hatchlings easily become

trapped in the surface tension; to avoid this, a few small flakes of 1-hexadecanol can be added to the dish.

APPENDIX

B

Hox1 Genomic Organization

When biologists discovered that *Hox* cluster genes are expressed in a broadly conserved pattern between flies and vertebrates, they proposed that the common ancestor of all bilaterians probably employed *Hox* genes in establishing regional identity along the anterior to posterior axis. Additionally, hypotheses have been proposed to explain the maintenance of *Hox* genes in clusters, and to explain the enigma of colinearity, the mirroring of the arrangement of *Hox* genes along the chromosome with the order in which they are expressed chronologically and along the body axis during embryonic development. In the vertebrate *Hox* clusters, for which a great deal of physical linkage and sequence data are known, no non-hox genes have been observed to intervene between the clustered *Hox*es. It is thought that intercalation of other genes would disrupt cis-regulatory mechanisms of within the cluster. There are variations in cluster organization, however. The *Hox* genes of *Drosophila melanogaster* are broken into two separate clusters on different chromosomes. There are also significant differences among taxa in how *Hox*es are regulated. Serial cis-regulation appears to be important in vertebrate clusters, but in flies, the *Hox*es are separated by buffering sequences that create barriers between regulatory elements. In addition, *Hox* expression in several taxa is not

uniformly regulated by position along the body axis, as in vertebrates and flies; in nematodes and leeches, expression of specific *Hox*es may be linked to cell lineage rather than position along the axis (reviewed in Gellon and McGinnis, 1998).

The single amphioxus cluster may have undergone tandem expansion of the posterior *Hox* genes and these posterior genes do not pair as obvious orthologues with the vertebrate posterior paralogy groups (Ferrier *et al.*, 2000). The number of *Hox* genes in the common ancestor of the chordates is still an open question. In an effort to determine the *Hox* cluster organization in *O. dioica*, I made a cosmid library of genomic DNA collected from ripe males. A screen of this library with degenerate primers designed to target anterior *Hox* genes and with *Hox*-specific primers yielded three overlapping *Hox*-containing cosmids. The gene specific primers were based on sequence for *O. dioica Hox1* obtained in a PCR survey that previously yielded 3 different anterior *Hox* fragments. While all three cosmids contain the *Hox1* paralogue, however, no other *Hox*es were found. After extensive subcloning of the cosmids and shotgun sequencing of overlapping fragments, several contigs coalesced that contain a small number of genes. A 12 kb and a 4.6 kb contig are probably linked as they share some overlapping clones. If this linkage within the cosmid insert sequences accurately represents the genomic sequence in *O. dioica*, it means that the *O. dioica Hox1* gene falls between a *Histone 4* gene and a *cellulase* gene. In human, the H4 histone family members reside on chromosome 6p, unlinked to any of the *HOX* clusters, and humans lack a cellulase gene. This unexpected finding may have confirmation in the work of Seo *et al.* (2002) who have identified several *Hox*-containing cosmids in *O. dioica* and found that the *Hox* genes

appear to be scattered in the genome; no cosmid appeared to contain more than one *Hox* gene, implying the *Hox*es are at least tens of kilobases apart. The breaking of the larvacean cluster could indicate a divergent mechanism for *Hox* regulation. Larvaceans are thought to be eutelic, having stereotyped cell-lineages with fixed cell numbers; perhaps expression of *Hox*es is tied to segregation of cell lineages rather than to the sequential and nested activation of *Hox* genes as the axis extends. The most anterior vertebrate *Hox* genes exhibit anterior borders of expression in the rhombomeres of the developing hindbrain. Early expression of *O. diocia Hox1* may exhibit a homologous role in patterning the central nervous system (Fig. B.1).

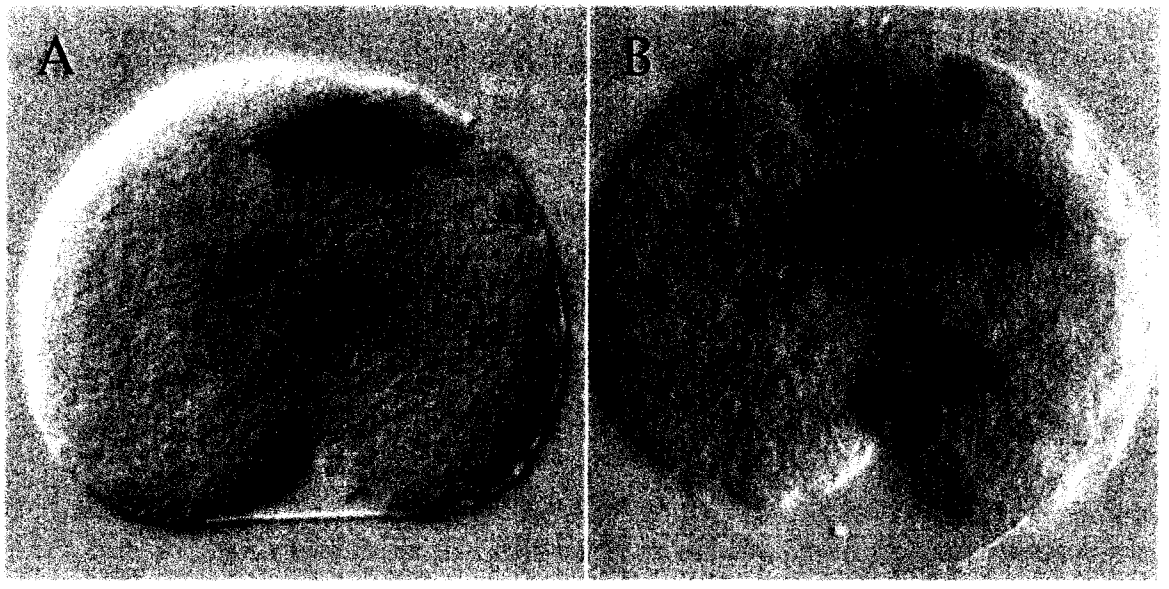


FIG. B.1. Early expression of *Hox1*. (A) A mid-tailbud stage embryo with the trunk oriented to the left and tail to the right. *Hox1* is expressed dorsally, perhaps in the nervous system, at the junction of the trunk and tail. (B) A slightly later tailbud stage embryo, oriented in an oblique dorsal view. Expression is seen in the mediodorsal domain seen in A, but also in two lateral domains.

APPENDIX

C

Pax3/7

In Chapter III of this dissertation, I reported the sequences and expression for larvacean representatives of three of the four Pax gene family subgroups. A larvacean *Pax3/7* was also cloned in this study, and here I report an aspect of the expression (Fig. C.1) and the sequence (Fig. C.2) and of this gene. In hatchling stages, *O. dioica Pax3/7* is expressed in epidermal domains of the dorsal trunk, perhaps marking regions of the developing oikoplastic epithelium. These dorsal domains may correspond to parts of a speciallized array of cells called Fol's oikoplast which secrete the feeding filters of the house. *Pax3/7* could play a role in patterning the complexly arrayed epithelium or may act upstream of various, unrelated "oikosin" genes identified by Thompson *et al.* (2001) to be regionally expressed in oikoplast domains in adults larvaceans.

Larvacean Pax3/7 lacks a recognizable octapeptide sequence that occurs between the paired domain and the homeodomain in the orthologous *Pax*es of other organisms. This is like the ascidian Pax3/7 protein, however, which also lacks the octapeptide (Wada *et al.*, 1997).

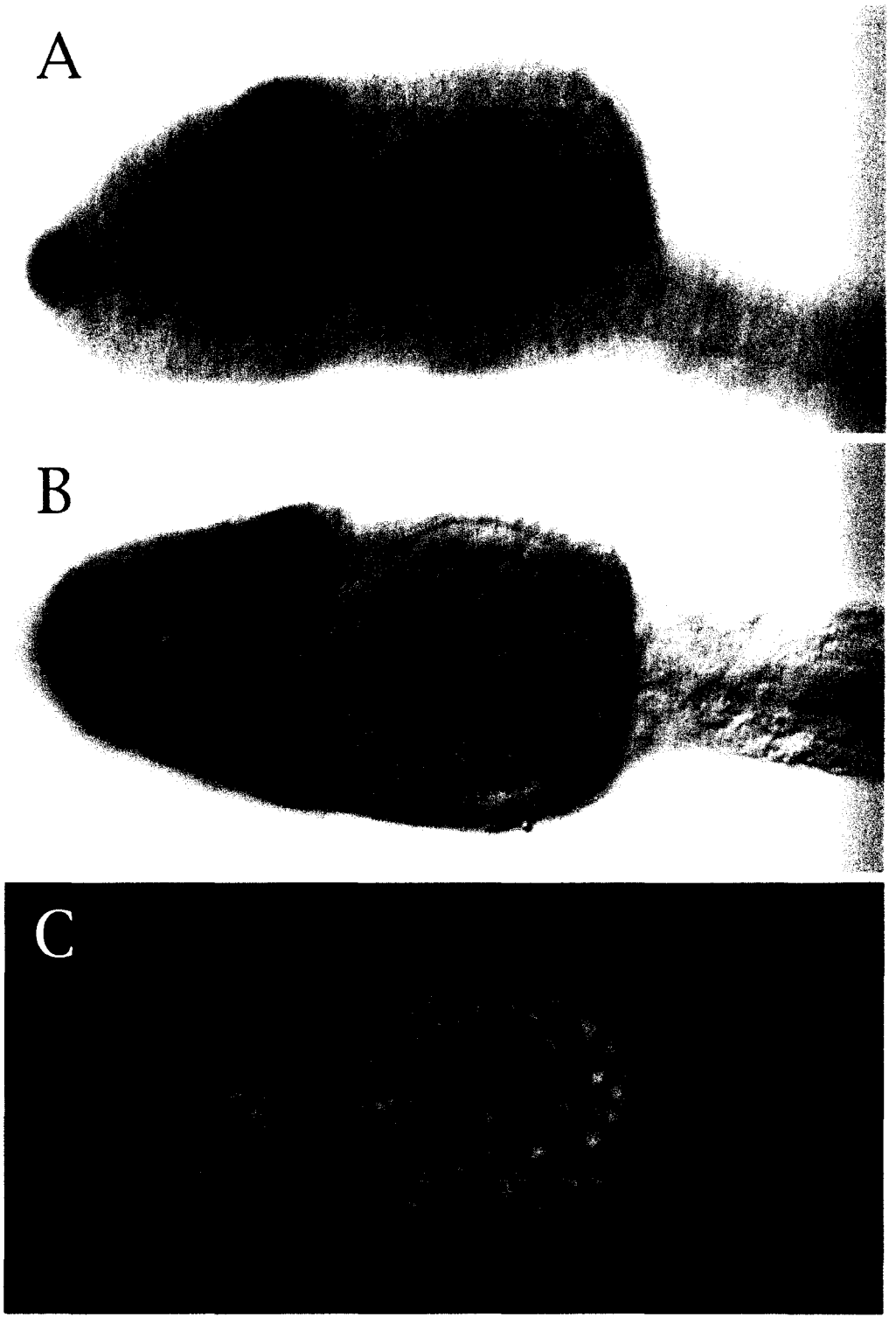


FIG. C.1. Expression of Pax3/7. (A, B) Lateral and dorsal views of 10.5 hour hatchlings. Transcripts are detected in domains in the dorsal epidermis that may be associated with Fol's oikoplast and also in a band dorsal to the mouth. (C) A dorsal view of the same hatchling treated with the nuclear stain DAPI. Cells of the oikoplastic epithelium will become stereotypically arranged to create a prepattern for the architecture of the secreted house. Already at this early stage, some of this orderly arrangement is apparent.

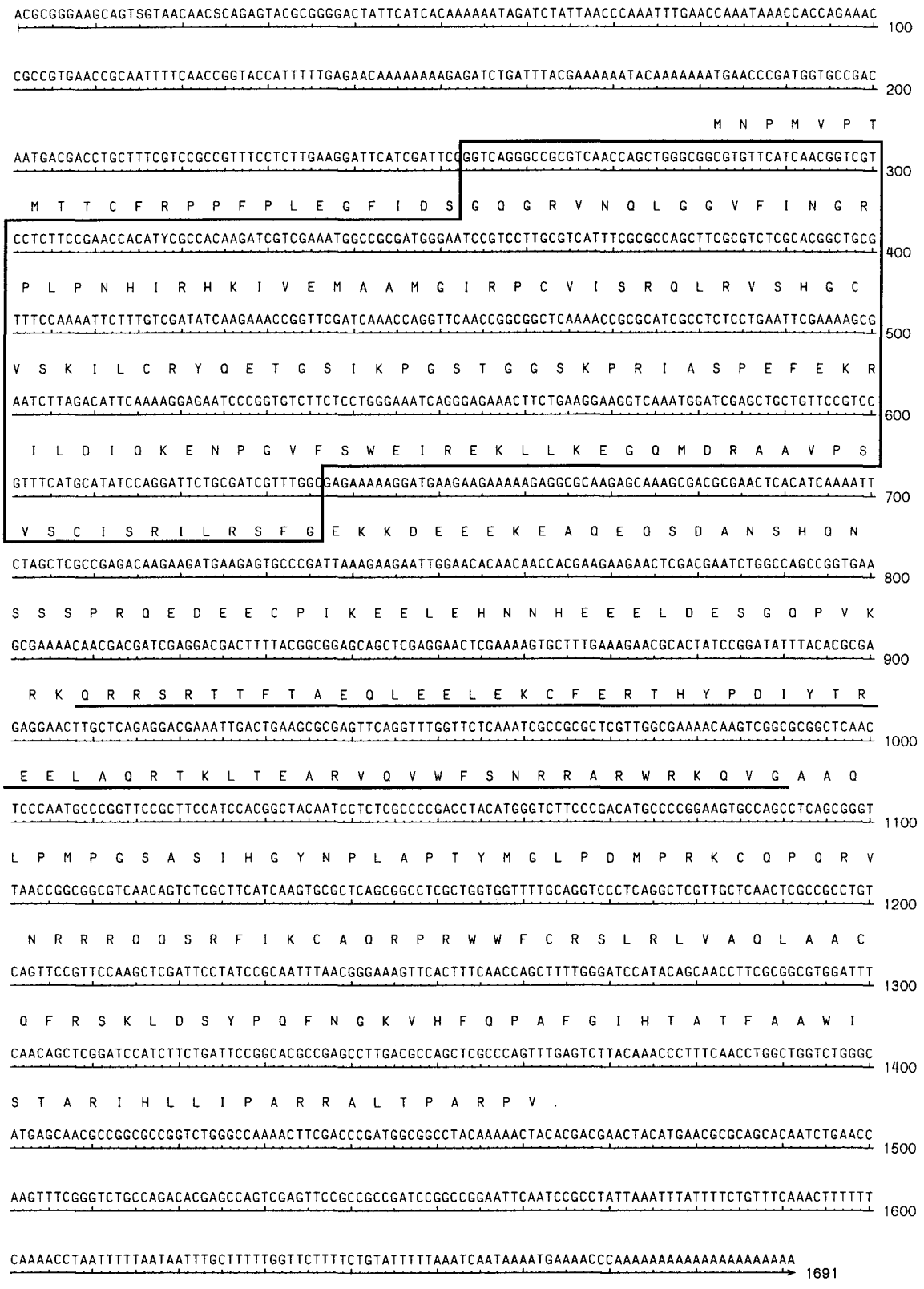


FIG. C.2. Sequence of *Pax3/7*. Paired domain is boxed, and the homeodomain is underlined.