EVOLUTION AND EXPRESSION OF THE ACTIN GENE FAMILY OF THE SEA URCHIN *LYTECHINUS PICTUS*

by

Hung Fang

B.Sc., Peking University, 1983

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the Department

of

Biological Sciences

SIMON FRASER UNIVERSITY

March 1994

APPROVAL

NAME: HUNG FANG

DEGREE: DOCTOR OF PHILOSOPHY

TITLE OF THESIS:

EVOLUTION AND EXPRESSION OF THE ACTIN GENE FAMILY OF THE SEA URCHIN LYTECHINUS PICTUS

Examining Committee:

Chair: Dr. A. T. Beckenbach, Associate Professor

Dr. B. P. Brandhorst, Professor, Senior Supervisor,
Department of Biological Sciences,
Simon Fraser University

Dr. M. J. Smith, Professor,
Department of Biological Sciences,
Simon Fraser University

Dr. D. L. Baillie, Professor
Department of Biological Sciences
Simon Fraser University

Dr. B. Honda, Associate Professor
Department of Biological Sciences
Simon Fraser University
Public Examiner

Dr. W. R. Crain, Professor,
McLaughlin Research Institute, University of Montana
External Examiner

Date Approved March 14, 1994

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

EVOLUTION AND EXPRESSION OF THE ACTIN GENE FAMILY OF THE SEA URCHIN LYTECHINUS PICTUS

Author: _____
(signature)

_____
(name)

March 15, 1994
(date)

Abstract

Several genomic DNA and cDNA clones corresponding to actin genes of the sea urchin *Lytechinus pictus* were recovered by screening genomic and complementary DNA libraries with actin coding sequence hybridization probes. Five different actin genes were identified on the basis of differences in 3' noncoding and amino acid coding sequences. DNA sequence and DNA gel blot hybridization analyses indicate that the actin gene family of *L. pictus* consists of five genes: a single muscle actin gene designated as LpM, and four cytoskeletal actin genes designated as LpC1, LpC2, LpC3, and LpC4. The origin and relationship of these actin genes to members of the actin gene family of the sea urchin *Strongylocentrotus purpuratus* were considered by molecular phylogenetic analysis. Comparison of deduced amino acid sequences indicated a close relationship between the LpC1 and the CyI-CyII subfamily of *S. purpuratus* actin genes, and between the LpC2 and the CyIII subfamily of *S. purpuratus* actin genes; the muscle actin genes are orthologous. The temporal and spatial patterns of *L. pictus* actin gene expression were determined by RNA gel blot hybridization and whole mount *in situ* hybridization with gene-specific probes. All the *L. pictus* actin genes were expressed during embryogenesis, each with a distinct pattern of expression. The LpC1 and LpC2 messenger RNAs were most abundant; LpC1 mRNA accumulated in several spatial territories, while LpC2 mRNA accumulated predominantly in the aboral ectoderm. None of the *L. pictus* actin genes has a pattern of expression identical to any *S. purpuratus* actin gene.

To My Parents

Acknowledgments

I wish to thank my thesis supervisor, Dr. Bruce Brandhorst for providing me the opportunity of doing my thesis project in his laboratory, which means a lot to me. I have greatly benefited from his expert guidance, knowledgeable suggestions; I enjoyed the stimulating discussions, and the autonomy he gave me throughout the work. The most important things he has taught me were critical thinking and a serious attitude to scientific work.

I thank members of my supervisory committee, Drs. Michael Smith and David Baillie, for their interest, appropriate advice and guidance of the thesis. Dr. Smith has given many useful suggestions on the molecular phylogenetic analyses involved in the thesis; Dr. Baillie has patiently advised me on the computer problems I encountered.

I am indebted to all the members of our laboratory over the past five years. Dr. Zhiyuan Gong initiated this work; he screened a cDNA library of *L. pictus* and partially sequenced several actin cDNA clones. I thank Dr. Zhiyuan Gong, Dr. Arianna Lee, Francis Ouellette, Dr. Peter Cserjesi, Isabelle Montpetite and Becky Dolan for their enthusiastic assistance and friendship which made my starting time in the lab easy and enjoyable. I thank Peter Fairly, Diana Richer, Ruth Davenport, John Boom, Minghua Fu, Graham Corley-Smith, Diane Driver, Bryan Crawford, and David Lum for their unselfish help and interesting discussions.

I also gratefully acknowledge many faculty members and graduate students from the Biological Sciences Department and IMBB. In particular, I want to thank Dr. Gabe Kalmar, Dr. Andy Beckenbach, Karen Beckenbach, Marco Marra, Steven Jones, Alan Arndt, Sarah Baldry and Dr. Vic Bourne.

I wish to thank my whole family, my father, mother, husband, brother and all relatives, for their love, encouragement, support, and for sharing all my frustration and joy.

Table of Contents

List of Tables

List of Figures

CHAPTER I GENERAL INTRODUCTION

1.1 ACTIN GENE FAMILIES

A multigene family is a set of related genes descended by duplication and divergence from a common ancestral gene. These gene families are useful models for studying the evolution of eukaryotic genes in relation to their function and regulation. The actin gene family is particularly well characterized.

1.1.1 Structure and function of actin polypeptides

Actin is ubiquitous in eukaryotes, and is the most abundant of all the cytoskeletal proteins in the cytoplasm of cells. Thin filaments had been seen in cells by electron microscopy before they were identified as actin microfilaments by Ishikawa and co-workers (1969). In animal muscles, actin is primarily involved in contraction of differentiated muscle cells. In nonmuscle animal cells, actin is involved in a variety of processes including cytoskeletal structure, cellular motility, cell-surface mobility, intracellular transport, cytoplasmic streaming, cytokinesis, endocytosis, exocytosis, clot retraction, microvillar movement and, possibly, chromosomal condensation and mitosis (Schliwa, 1981; Lloyd, 1983; Ponte *et al.*, 1983; Stossel, 1984).

An actin polypeptide consists of 375 or 376 amino acids, depending on the isoform of the polypeptide. The region located between codons 10 and 230 is more conserved through evolution than adjacent regions. The functional domains and various binding sites, such as those for ATP, calcium ions, actin polymerization and various actin-associated proteins (e.g, myosin, tropomyosin, troponin, actinin, cofilin), have been elucidated (Hambly *et al.*, 1986). Actin-binding proteins probably impose constraints on the actin primary sequences. As a consequence, protein domains corresponding to each

binding site are highly conserved in actins of both animals and plants (Tellam *et al.*, 1989; McElroy *et al.*, 1990). Actins can be categorized into two major types based on diagnostic amino acid residues: cytoskeletal and muscle (Vandekerckhove and Weber, 1984). The biological significance of amino acid residues which distinguish cytoskeletal and muscle isoforms remains unknown. However, high resolution models of the tridimensional structure of actin polypeptides, such as the one proposed for vertebrate skeletal muscle actin (Kabsch *et al.*, 1990), should improve our knowledge of the *in vivo* functioning of actins.

1.1.2 Divergent actin gene families

The primary structure of actin genes and proteins has been examined across broad phylogenetic distances. Actins are encoded by a multigene family in all animals, all plants, and most protozoa so far examined. The genomes of *Tetrahymena* (Cupples and Pearlman, 1986) and *Saccharomyces cerevisiae* (Gallwitz and Seidel, 1980) contain only a single gene, although an actin-like gene has been found in yeast recently (Schwob and Martin, 1992). The organization of the actin gene family has been studied in *Dictyostelium discoideum* (McKeown and Firtel, 1981), *Caenorhabditis elegans* (Files *et al.*, 1983), *Drosophila melanogaster* (Fyrberg *et al.*, 1981), *Strongylocentrotus purpuratus* (Crain *et al.*, 1981; Davidson *et al.*, 1982), chicken (Fornwald *et al.*, 1982), rat (Zakut *et al.*, 1982) and human (Engel *et al.*, 1981).

Acrasiomycota: *Dictyostelium discoideum*

Dictyostelium has 17 actin genes (Kindle and Firtel , 1978; McKeown *et al.*, 1978). None of the actin genes sequenced by Firtel *et al.* (1979) appears to have introns within their protein-coding regions. At least one actin gene, actin 2-sub 2, is a pseudogene. The actin genes can be categorized into subfamilies by their 3' noncoding regions (McKeown and Firtel, 1981).

Nematode: *Caenorhabditis elegans*

The nematode *Caenorhabditis elegans* has four actin genes, three clustered and one unlinked (Files *et al.*, 1983). The cluster of three actin genes has been mapped to linkage group V (Files *et al.*, 1983), while the fourth unlinked actin gene has been mapped to the X chromosome (Albertson, 1985). The four *C. elegans* actin genes are highly conserved at the nucleotide level, the lowest similarity seen in pairwise comparisons being 92%. Two of the genes are identical within the coding region. Among the four actins, only three conservative amino acid replacements are found (Krause *et al.*, 1989). The amino acid sequences of the four actins are typical of those for invertebrates: they resemble more closely vertebrate cytoskeletal actins rather than vertebrate muscle actins.

Arthropod: *Drosophila melanogaster*

Six actin genes have been found in the genome of *Drosophila melanogaster*, encoding at least five different proteins (Fyrberg *et al.*, 1981). The deduced amino acid sequences of the six actin genes resemble vertebrate cytoskeletal actins. Sequences of Drosophila actins characterized to date are extremely conserved; most of the amino acid replacements are conservative. Most striking is the variation in the positions of introns. For instance, one intron within codon 13 of the Drosophila actin gene DmA4 is not present in any other Drosophila actin genes; one within codon 307 of the DmA6 gene is not present in at least two other Drosophila actin genes; at least two genes, DmA3 and DmA6, are not interrupted by an intron between codons 121 and 122 where many actin genes, including most sea urchin actin genes, have an intron.

Echinoderm: *Strongylocentrotus purpuratus*

The actin gene family of the sea urchin *Strongylocentrotus purpuratus* is the most completely characterized among echinoderms (Cooper and Crain, 1982, Davidson *et al.*, 1982). The genome of *S. purpuratus* consists of at least eight nonallelic actin genes, which have been assigned to four subfamilies (Lee *et al.*, 1984). There are seven

cytoskeletal actin genes, including a single CyI actin gene, three CyII genes (CyIIa, CyIIb and CyIIc), three CyIII genes (CyIIIa, CyIIIb and CyIIIc), and a single muscle actin gene M. Two actin genes, CyIIc and CyIIIc, have been identified as pseudogenes. The other five cytoskeletal actin genes are closely linked into two clusters on chromosomes: cluster CyI-CyIIa-CyIIb and cluster CyIIIa-CyIIIb (Scheller *et al.*, 1981; Lee *et al.*, 1984; Akhurst *et al.*, 1987; Flytzanis *et al.*, 1989). Intron locations are conserved among the actin genes of *S. purpuratus*: all the actin genes possess two introns in the protein-coding region between the codon positions 121 and 122, and within the codon 204, while the muscle actin gene M possesses two extra introns at the codon positions 41/42 and 267/268.

Mammals: human and mouse

Mammalian genomes, such as mouse and human, have been reported to include 20-30 actin genes (Humphries, *et al.*, 1981; Engel *et al.* 1981). However, the number is likely to be overestimated because of the inaccuracy of the methods involved in those early studies. Slight sequence differences located primarily in the amino terminal region of mammalian actins result in six distinct actin isoforms (Vandekerckhove and Weber 1978), which fall into three major categories: cytoskeletal types (β and γ in both brain and thymus), smooth muscle types (vascular and nonvascular) and striated muscle types (α skeletal and α cardiac). The primary sequences of smooth-muscle and striated-muscle actins are more closely related to each other than they are to cytoskeletal actins. Muscle and cytoskeletal actins differ in fewer than 7% of their amino acids. The positions of introns in human actin genes characterized is conserved but the sizes of introns and untranslated regions are different (Miwa *et al.*, 1991).

In summary, protein coding regions of actin genes are highly conserved among eukaryotes. The number of actin genes varies widely among diverse species, while the

number and location of introns within actin genes are also highly variable (Shah *et al.*, 1983; Hightower and Meagher, 1983); invertebrate actin genes resemble more closely vertebrate cytoskeletal actin genes.

1.2 DIFFERENTIAL REGULATION OF THE ACTIN GENE EXPRESSION

In all organisms studied to date, members of actin multigene families are differentially expressed either temporally, spatially or both. The differential expression of members of a conserved gene family makes actin genes particularly attractive for studies of developmental gene regulation, since they can be readily cloned and distinguished.

1.2.1 Differentially expressed actin genes in different organisms

Dictyostelium

At least six of 17 actin genes of *Dictyostelium discoideum* are expressed. Four (actin genes 5, 6, 8 and M6) show different patterns of expression during development, although three (actin genes 5, 6, and 8) of these appear to encode the same protein (McKeown and Firtel, 1981). The transcripts of actin gene 8 are present at a high level in vegetative cells. Actin gene 6 is expressed at a high level during vegetative growth, but the level of the mRNA drops dramatically during development. Actin gene 5 is represented by 5 to 10% of actin RNA at all times examined (McKeown and Firtel, 1981).

Drosophila

Although the six actin genes of *D. melanogaster* encode similar polypeptides (Fyrberg *et al.*, 1981; Sanchez *et al.*, 1983), they are differentially expressed during development (Fyrberg *et al.*, 1983; Sanchez *et al.*, 1983; Tobin *et al.*, 1990). The 5C and 42A actin genes encode cytoskeletal actins; transcripts from both genes are evenly distributed in preblastoderm embryos, becoming localized to the periphery of the

blastoderm. Later in development, the two cytoskeletal actin genes are differentially expressed in specific tissues. Transcripts of actin 42A are enriched in the midgut, brain, nerve cord, and gonad. Transcripts of actin gene 5C show more complicated expression patterns, and separate into two tissue-specific subsets. RNAs produced by utilization of the 5C distal promoter are enriched in the cells of proventriculus and midgut, while those transcribed from the 5C proximal promoter accumulate in all tissues at low levels (Burn *et al.*, 1989). The 87E and the 57B transcripts accumulate in the developing larval body wall musculature, but at differing levels and in differing patterns (Tobin *et al.*, 1990). Transcripts of the 79B and the 88F actin genes are undetectable in embryos; both genes are expressed in the thorax of the fly at about the same time during development (Fyrberg *et al.*, 1983), but in two nonoverlapping sets of muscle tissues. The 88F actin gene is active in the indirect flight muscle of the fly, while the 79B actin gene is strongly expressed in leg muscles, direct flight muscles, and muscles which support the abdomen and the head (Mahaffey *et al.*, 1985; Hiromi and Hotta, 1985). The 79B gene is also expressed in the scutellar pulsatile organ and in abdominal muscles which are present only in males (Courchesne-Smith and Tobin, 1989).

Amphibians

In the *Xenopus laevis* embryo, cytoskeletal and muscle-specific actin genes show different patterns of expression. Transcription of the α-actin genes is restricted exclusively to those regions of the mesoderm that are committed to form embryonic muscle. The α-cardiac actin transcripts and their α-skeletal counterparts are absent in the early stages of development, being first detected in neurula embryos. By contrast, a low, constant level of cytoskeletal actin mRNAs was detected in all stages before gastrulation. In later stages, the levels of both cytoskeletal and sarcomeric actin mRNAs increase in a similarly dramatic manner. One cytoskeletal actin gene, the type-5 actin gene, is expressed in nonmuscle cells of all regions of the developing embryo. However, the other

cytoskeletal actin (type-8), is correlated with the muscle-specific skeletal actin gene and is restrictively activated in embryonic muscle tissue, in contrast to other vertebrate cytoskeletal actin genes which are exclusively expressed in nonmuscle cells (Mohun and Garrett, 1987).

Mammals

In mammals, individual actin isoforms show tissue and developmental stage specificity. Pairs of actin isoforms are coexpressed under some circumstances (Gunning *et al.*, 1983; Vandekerckhove and Weber, 1981). For instance, the cytoskeletal β- and γ-actin genes are coexpressed in mouse tissues and organs, but their mRNA levels are subject to differential regulation among different cell types (Erba *et al.*, 1988).

1.2.2 Actin gene expression during embryogenesis of sea urchins

Sea urchin embryogenesis

Embryogenesis of indirect developing sea urchins is very similar among species. Most experiments involved in this thesis research were performed with the embryos of the sea urchin *Lytechinus pictus*. Therefore, the embryogenesis of this species is described. Embryo culturing is carried out in artificial sea water at 18°C. Eggs of sea urchins are meiotically mature when shed (Giudice, 1973, 1986). Once fertilized, the sea urchin egg initiates a period of rapid cleavage. The first cleavage occurs about 90 minutes after insemination. The first two cleavages divide the embryo equally along the animal-vegetal axis and the third cleavage plane is equatorial, dividing the embryo into animal and vegetal halves. The fourth cleavage is equal in the animal half but unequal in the vegetal half, producing eight mesomeres of equal size in the animal half but four large macromeres and four small micromeres in the vegetal half. The cells of the animal half of the embryos give rise to much of the ectoderm of the larva, the macromeres give rise to ectoderm, endoderm, and mesoderm, and the micromeres give rise to mesoderm, some of their

progeny being responsible for skeletogenesis. The rate of cell division slows at about 10 hours postfertilization at blastula stage, with approximately 256 cells. The ciliated epithelial cells of the blastula are arranged in a single layer around the blastocoel. Hatching usually begins around 12 hours PF, brought about by secretion of a hatching enzyme, producing a free swimming hatched blastula. The vegetal pole flattens at late blastula stage to form the vegetal plate. Primary mesenchyme cells, derived from the 4 large micromeres formed at 32 cell stage, ingress into the blastocoel. At about 24 hours PF, the vegetal plate invaginates to form the archenteron. At late gastrula stage, the tip of the archenteron contacts the ectoderm near the animal pole, in part the result of contractile activity of the secondary mesenchyme cells which originate at the tip of the archenteron (Gustafson and Kennander, 1956; Hardin and McClay, 1990; Crawford and Burke, 1994). Primary mesenchyme cells fuse with each other and form a ring on the inner layer of the ectoderm around the archenteron, and elaborated the triradiated spicules of the skeleton, composed of $CaCO_3$, $MgCO_3$, and several spicule matrix glycoproteins (Wilt and Benson, 1988). The archenteron bends toward the ventral side and opens into a stomadaeum (mouth), while the blastopore becomes the anus, as in all deuterostomes. The gut then differentiates into esophagus, stomach, and intestine. Some of the secondary mesenchyme cells, including some derived from the four small micromeres formed at the 5 th cleavage, form a coelomic sac on both sides of the esophagus. Meanwhile, the skeletal spicules have grown and increased in number to form two oral arms, two anal arms, two transverse rods and two body rods of the pluteus larva. A ciliary band is formed at the boundary of the oral and aboral ectoderm. Until this stage of development, the embryo has maintained the same mass. The pluteus larva begins to feed and grows considerably during larval development. At metamorphosis the specifically larval structures initially formed during embryogenesis are destroyed or discarded, including the larval skeleton, the anal, oral and esophageal structures, the ciliary band, and most of the collapsed larval ectoderm (Czihak, 1975; Hörstadius, 1973).

In the undisturbed embryo there is a strict relationship between cell lineage and cell fate (Cameron and Davidson, 1987; Cameron and Davidson, 1991; Davidson, 1989). In *S. purpuratus*, five spatially restricted territories are established during the first six cell divisions: the small micromeres, the skeletogenic primary mesenchyme cells derived from the large micromeres, the vegetal plate, the oral ectoderm, and the aboral ectoderm. Each territory is derived from several founder cells, and show territory-specific gene expression and regulation (for review: Davidson, 1989; Brandhorst and Klein, 1992). Several *L. pictus* genes show patterns of expression and territorial restrictions similar to their homologues in *S. purpuratus* embryos, indicating that the *L. pictus* embryo may establish territories in a manner similar to the *S. purpuratus* embryo.

The expression of the actin gene family of sea urchin *S. purpuratus*

Six of the eight known *S. purpuratus* actin genes (CyI, CyIIa, CyIIb, CyIIIa, CyIIIb and M) are functional and are distinguishable by sequences present in the 3' nontranslated region of each mRNA (Lee *et al.*, 1984). The temporal and spatial domains of expression of the actin genes are differentially regulated during the *S. purpuratus* life cycle (Crain *et al.*, 1981; Garcia *et al.*, 1984; Shott *et al.*, 1984; Cox *et al.*, 1986; Lee *et al.*, 1986). CyI and CyIIb messages initially appear in all cells, but then become restricted to the lineages destined to form oral ectoderm and gut at the mesenchyme blastula stage. The major difference in the expression of the two genes is that the CyIIb transcripts are present at a much lower absolute concentration than are those of the CyI gene (Shott *et al.*, 1984). The CyIIa transcripts are observed in the vegetal plate before ingression of primary mesenchyme cells. However, by late gastrula stage, expression of the CyIIa gene is limited to secondary mesenchyme cells; in the pluteus the CyIIa gene is expressed in the stomach and intestine as well. The CyIIIa and CyIIIb genes are expressed exclusively in the aboral ectoderm, and their expression is confined to the embryo. Gene CyIIIa

provides the major embryonic actin mRNA, contributing approximately 65% of actin poly(A)$^+$ RNA at gastrula stage. Another 25% of actin transcripts correspond to the CyI gene (Shott et al., 1984). M actin transcripts appear only in the muscle cells surrounding the esophagus beginning at early pluteus stage (Cox *et al.*, 1986).

Gene expression in sea urchin hybrid embryos

Hybrid embryos constructed from different sea urchin species have been used in variety of classical and modern studies (Morgan, 1927; Chen and Baltzen, 1979, for review). *S. purpuratus* and *L. pictus* can be reciprocally crossed, forming healthy plutei having a predominantly maternal morphology (Brandhorst and Klein, 1992). Some mRNAs accumulate to reduced levels in the hybrid embryos (Conlon *et al.*, 1987), while others are present at roughly normal levels (Crain and Bushman, 1983). The *S. purpuratus* actin genes (CyI, CyIIIa and M) show correct temporal and spatial pattern of expression in reciprocal hybrid embryos, indicating that the *trans* regulatory factors necessary to regulate the temporal and spatial expression of these genes are present in the hybrid embryos (Bullock *et al.*, 1988; Nisson *et al.*, 1989). For the gene family consisting of Spec1 and 2 genes of *S. purpuratus* and the two LpS1 genes of *L. pictus* (which encode aboral ectoderm specific calcium binding proteins), the Spec1 gene is expressed at nearly normal levels in hybrid plutei, while mRNA of the homologous LpS1 genes is at only 2-5% of its normal level (Brandhorst *et al.*, 1991). The spatial patterns of gene expression of the Spec1 and LpS1 genes show that one or the other is exclusively expressed in patches of aboral ectoderm cells of hybrid plutei; most plutei exclusively express Spec1 mRNA. The expression in hybrid embryos of other tissue-specific *L. pictus* genes, for instance the homologues of the *S. purpuratus* actin genes, remains to be investigated.

1.3 REGULATION SYSTEMS OF ACTIN GENE FAMILIES

5' flanking sequences as transcriptional regulators

Regulation of the actin multigene family involves the recognition of *cis*-acting regulatory sequences that specify the tissue type and developmental program of expression for each actin isoform. Studies carried out on a variety organisms have revealed that, in general, regulatory elements necessary for proper expression of actin genes are located in the 5' flanking region.

The 5' flanking sequences of human skeletal α-actin gene are sufficient to direct tissue-specific and differentiation-regulated expression when transfected into rat L8 myogenic cells, indicating a regulatory system conserved between humans and rats. Most notable among the conserved sequences are the CC(A/T rich)6GG (CArG box) motifs which have demonstrable interactions with *trans*-acting transcriptional factors. This same motif has been identified in several other genes and in some also serves as binding sites for transcription factors (Taylor *et al.*, 1988).

Xenopus laevis cytoskeletal actin gene promoters contain a 20-bp sequence homologous to the serum response element (SRE) required for transient human c-*fos* gene transcription in response to serum factors. Sequences containing the SRE homology are essential for constitutive activity of the actin promoter in both *Xenopus* and mouse cells, and a synthetic SRE can function as a promoter element. In mouse cells, transcription of both transfected *Xenopus* actin and actin/c-*fos* fusion genes is activated following serum stimulation. These data suggest that the SRE and its cognate protein form part of a regulatory pathway that has been highly conserved during vertebrate evolution (Mohun *et al.*, 1987). Study of the expression of a transgenic *Xenopus* cytoskeletal actin gene indicated that as little as 485 nucleotides of upstream sequence is sufficient for proper temporal control of activation of the transgenic gene (Brennan, 1990).

The major cytoskeletal actin gene of *Drosophila melanogaster*, actin 5C, has two promoters which are differentially controlled and possess distinct sets of regulatory elements (Chung and Keller, 1991). The proximal promoter controls constitutive synthesis of the 5C actin mRNA in all growing tissues (1990a), while the distal promoter controls synthesis of the RNA in a tissue- and developmental stage-specific manner (Chung and Keller, 1990b). Interestingly, the distal promoter has a TATA motif and depends on it for its correct function, while the proximal one does not (Chung and Keller, 1990a). Several positive and negative regulatory elements have been identified as required for the specific activities of each promoter (Chung and Keller, 1990a, b).

Recent studies of actin gene expression in sea urchin embryos have been focused on the mechanisms regulating differential gene expression. Deletion constructs of the 5' flanking sequence of an investigated actin gene fused to a reporter gene, such as the CAT (bacterial chloramphenicol acetyl transferase) or β-gal (β-galactosidase) genes, are expressed in developing embryos following microinjection into eggs. These studies revealed segments of the 5' flanking regions which are necessary for embryonic expression of the gene. Among the best characterized sea urchin *cis*-acting systems is that controlling expression of the CyIIIa actin gene. The CyIIIa regulatory domain includes approximately 20 sites of DNA-protein interaction, serviced by about ten different factors. Some of these sites are known to negatively control spatial expression, while others positively regulate temporal activation and the level of CyIIIa gene expression (Davidson 1989, for review). Transgenic analysis of the CyIIIb actin gene reveals that at least five DNA:protein interaction sites in the 5' flanking sequence are necessary for quantitative expression; one of these elements (E1) is involved in the temporal regulation of the CyIIIb gene (Niemeyer and Flytzanis, 1993). A great deal of progress has been made in the characterization of the *trans*-acting regulatory proteins that interact with *cis*-acting DNA sequence elements of echinoid cytoskeletal actin genes (Calzone *et al.*, 1988; Theze *et al.*,

1990; Thiebaud *et al.*, 1990; Franks *et al.*, 1990; Hough-Evans *et al.*, 1990). Sea urchin transcription factors, interacting with the *cis*-acting elements of actin genes, have been identified by gel shift assays and footprinting, cloned with the aid of affinity chromatography or nucleotide sequence analysis, and classified with respect to their target sites and the genes they regulate. The CyIIIa-CAT fusion gene is expressed appropriately in the *S. purpuratus* embryos, but is expressed ectopically in the embryos of a distantly related sea urchin species, *Lytechinus variegatus* (Franks *et al.*, 1988). By contrast, the CyI actin fusion genes are expressed properly in the embryos of *S. purpuratus* (Katula *et al.*, 1987) as well as *L. pictus* (Collura and Katula, 1992). These results indicate that the regulatory systems of some actin genes have diverged.

3' noncoding region:

Conservation of 3' noncoding sequences has been observed in several kinds of genes encoding structural proteins such as actin, α-tubulin and type I collagen; they are all members of multigene families. Conservation of 3' noncoding sequences has been reported between skeletal-actin genes (Ordahl and Cooper, 1983; Kedes *et al.*, 1984; Gunning *et al.* 1984); among the human, rat, mouse, and chicken β-actin and /or γ-actin 3' noncoding sequences (Ponte *et al.*, 1983, 1984); between the *X. borealis* type 1 actin gene and a human γ-actin gene (Cross *et al.*, 1988); and between *S. purpuratus* cytoskeletal CyIIa and CyIIb genes as well as CyIIIa and CyIIIb genes. The conservation of the 3' noncoding sequence of the *S. purpuratus* CyI actin gene extends to sea urchin species which have been separated more than 40 million years ago, while sequences homologous to the 3' noncoding sequence of the CyII and CyIII genes have been detected in other species in the family of Strongylocentrotinae (Lee *et al.*, 1984). The 3' noncoding regions of the human skeletal and cardiac actin genes have not evolved as complete units but rather as discrete subdomains. The 3' halves of the 3' noncoding regions of both genes show greater cross-species conservation than does the 5' halves of the 3' noncoding

regions (Gunning *et al.*, 1984). In contrast with this, the sequence similarity of sea urchin actin genes CyIIa and CyIIb is restricted to a limited region that does not extend further than ~70 nt beyond the stop codon (Durica *et al.*, 1988).

Little is known about the function of the conserved 3' noncoding sequences of actin genes. It has been suggested that this region is involved in translational control, or may be important in mRNA stability (Yaffe *et al.*, 1985; Cross *et al.*, 1988). The rearranged 3' noncoding region of *C. elegans* actin gene 1 implicates the involvement of the sequence in the steady-state regulation of the corresponding transcript (Krause *et al.*, 1989). There are other examples of 3' noncoding regulatory regions in *C. elegans* mRNAs. For instance, Pulak and Anderson (1989) have identified a mutation in the 3' noncoding region that influences the level of expression of the myosin gene. On the other hand, no detectable similarity has been found in the 3' noncoding regions of actin genes of insects *Bombyx* and *Drosophila* (Mounier *et al.*, 1992). The functional role of the 3' noncoding region of actin genes in vertebrates and echinoderms suggested by their conservation might be absent in insect actin genes.

1.4 EVOLUTION OF ACTIN GENE FAMILIES

The different genes of a multigene family likely arise by duplication and subsequent divergence from an ancestral gene. Several selective factors may have favored the evolution of some genes into multigene families. For instance, multiple copies of the gene may be needed to fulfill a requirement for a large amount of gene product (Finnegan *et al.*, 1977). In other cases, individual gene family members may have diverged, producing variant gene products (isoforms) which have specialized functions. A third possibility is that regulation of expression is facilitated by the diversity of promoters conferring

differential regulation. All three possibilities may account for the evolution of actin gene families. Actin is a major component of all eukaryotic cells. Different isoforms such as muscle actin isoforms and cytoskeletal actin isoforms have appeared during evolution, and may have acquired distinct functional properties. During the evolution of the actin gene family, different family members have acquired different transcriptional regulatory domains and have become subject to the differential regulation of gene expression during growth and development. It has been suggested that the pathways of gene evolution can be reconstructed by comparison of the amino acid and nucleotide sequences of the coding region, the 5' and 3' noncoding regions, and exon-intron arrangement in related genes (Efstradiatis *et al.*, 1980).

Comparison of the amino acid sequences of actins from different animal phyla shows three major subfamilies: the cytoskeletal actins, the muscle actins of vertebrates and echinoderms, and the muscle actins of insects. These three subfamilies appear to represent three evolutionary branches of actin gene families (Mounier *et al.*, 1992). It has been proposed that insect muscle actin genes emerged from an ancestral cytoplasmic actin gene within the arthropod phylum, whereas muscle actin genes of vertebrates and echinoderms appear to share a common evolutionary origin from an ancestral cytoskeletal actin gene (Mounier *et al.*, 1992). From the molecular structures of the six characterized human actin isoform genes, Miwa *et al.* (1991) proposed an evolutionary pathway for the human actin gene family: extant cytoskeletal actin genes have directly evolved from the presumed ancestral cytoskeletal actin gene through loss of introns. A prototype of muscle actin genes had been created by duplication of an ancestral cytoskeletal actin gene followed by replacements of several amino acids. Subsequently striated muscle actin and smooth muscle actin genes evolved from this muscle prototype by loss and acquisition of introns at different sites. Multigene families sometimes evolve together in a concerted manner via unequal crossing over and gene conversion (Li and Graur, 1991). Crain *et al.* (1987) have

provided evidence for a gene conversion event between a muscle and a cytoskeletal actin gene in the sea urchin *S. purpuratus*.

1.5 SPECIFIC AIMS OF THE THESIS RESEARCH

Comparative studies have proven to be a strong and useful research tool and have significantly enriched our knowledge of gene regulation and gene evolution. We are interested in the extent to which mechanisms regulating the spatial specificity of gene expression are conserved among related genes, in particular, actin genes. The sea urchin embryo is an ideal organism for the purpose of our study since large amount of sea urchin eggs and synchronously developing embryos are easily obtainable. The development is rapid and extensively described both on the morphological and the molecular level (Davidson, 1989). Phylogenetically, sea urchins belong to phylum Echinodermata, subphylum Echinozoa, and class Echinoidea. Like chordates, echinoderms are deuterostomes. The existence of a true coelom suggests that echinoderms and chordates have a common ancestor. Therefore, investigation of echinoderm genes and their regulation may shed light on the origins of genes and gene regulation in chordates, including vertebrates. The specific objectives of this thesis research were: (1) Isolation of cDNA clones and genomic DNA clones encoding *L. pictus* actins; (2) Characterization of the sequence of the isolated cDNA/genomic DNA clones; (3) Investigation of the evolutionary history of sea urchin actin gene families by comparison with actin genes of other echinoids; (4) Determination of the temporal and spatial patterns of actin gene expression during embryogenesis of *L. pictus*. Of particular interest was the identification of *L. pictus* orthologues of *S. purpuratus* actin genes. Such genes would enable the comparative investigations of the structure and function of actin genes, and might also shed light on the peculiar regulation of gene expression in hybrid embryos.

CHAPTER II THE ACTIN GENE FAMILY OF THE SEA URCHIN *L. PICTUS*: GENE NUMBER AND GENE SEQUENCES

2.1 INTRODUCTION

Actin genes have been cloned and characterized in a wide variety of phylogenetically diverse species (Chapter I, 1.1). In most species actin isoforms are encoded by a multigene family, where gene number and genomic organization vary among species. Individual members of actin gene families are frequently differentially expressed during development or are tissue-specific (Chapter I, 1.2). The actin multigene family is thus a useful system for study of gene regulation and evolution.

The actin gene family of the sea urchin *Strongylocentrotus purpuratus* is the most completely characterized among echinoderms (Cooper and Crain, 1982, Davidson *et al.*, 1982). It consists of at least eight nonallelic actin genes (Lee *et al.*, 1984). A single functional muscle (M) actin gene has been identified, as well as five functional cytoskeletal actin genes, designated as CyI, CyIIa, CyIIb, CyIIIa and CyIIIb (Durica *et al.*, 1980; Scheller *et al.*, 1981; Lee *et al.*, 1984; Akhurst *et al.*, 1987; Flytzanis *et al.*, 1989). Two actin genes, Sf15a and Sf15b, which are similar to the *S. purpuratus* actin genes CyI and CyII, have been cloned and sequenced in the sea urchin *Strongylocentrotus franciscanus* (Foran *et al.*, 1985). A gene apparently homologous to CyI, TgCyI, has been cloned and characterized in the sea urchin *Tripneustes gratilla* (Wang *et al.*, 1994). The sea urchins (class Echinoidea) *S. purpuratus* and *S. franciscanus* belong to the subfamily Strongylocentrodinae, while *T. gratilla* and *L. pictus* belong to the subfamily Toxopneustinae (Smith, 1988). The organization and expression of actin gene families in sea stars has also been investigated (Kovesdi and Smith, 1985; Kowbel and Smith, 1989).

The primary structures of actin genes and proteins are highly conserved. The region located between codons 10 and 230 are more conserved through evolution than adjacent regions. The C-terminal coding region is relatively less constrained. Several nonconservative amino acid replacements, clustered in the latter 40% of the primary sequences of *S. purpuratus* CyIIIa and CyIIIb actins , were identified as gene specific replacements (Durica *et al.*, 1988). Phylogenetic conservation of the 3' noncoding sequences has been reported among actin genes encoding same actin isoforms. For instance, regions of high similarity have been observed between the 3' noncoding sequences of the human and *X. laevis* cardiac muscle actin genes (Gunning *et al.*, 1984), and among the human, rat, and chicken β cytoskeletal actin genes (Kost *et al.*, 1983; Nudel *et al.*, 1983; Ponte *et al.*, 1984). Sequence features of the 3' noncoding termini have been used to classify *S. purpuratus* actin genes into subfamilies (Lee *et al.*, 1984).

We are interested in the regulation and evolution of the actin gene families in sea urchins. In contrast to the knowledge regarding to the organization, expression and regulation of the actin gene family of the sea urchin *S. purpuratus*, little was known about the actin gene family of *L. pictus*. In this chapter, I present a characterization of the structure and organization of the actin gene family of *L. pictus*.

2.2 MATERIALS AND METHODS

2.2.1 Preparation of sea urchin genomic DNA

L. pictus sea urchins were obtained from Marinus, Long Beach, CA. Genomic DNA was isolated from sperm of individual sea urchins according to the procedure of Blin and Stafford (1976) with modifications. Briefly, 0.1 ml of concentrated sperm was suspended in 12 ml of TNS solution (TNS: 400 mM NaCl; 100 mM Tris, pH 8.0; 2% SDS). The sample was digested with Proteinase K (Sigma) at concentration of 25 μg/ml

at 42°C overnight on a rocker, and then was digested for an additional 5 hrs at same temperature after adding 150 µl of 10 mg/ml Proteinase K. After the digestion, DNA was gently extracted three times with phenol/chloroform, and once with isoamyl/chloroform. Extractions were carried out in 50 ml Corning polypropylene tubes. During each extraction, the organic phase was drained from a hole punched at the bottom of the tube with a 18G1$^1/_2$ needle to prevent damage to DNA molecules during transfer of solution. DNA was precipitated by addition of two volumes of 100% ethanol and spooled out with a flamed glass pipette. DNA precipitant was resuspended in 10 ml TEN buffer (50 mM Tris, pH 7.5; 0.1 M NaCl; and 5 mM EDTA) at room temperature overnight. The resuspended DNA sample was extracted with phenol/chloroform and isoamyl/chloroform and precipitated as described above. DNA precipitant was again resuspended in 10 ml of TEN buffer, quantified by spectophotometry, aliquoted and stored at 4°C.

2.2.2 Isolation of actin cDNA clones

Two libraries of *L. pictus* gastrula stage cDNA (Xiang *et al.*, 1988; P. Cserjesi, PhD Thesis, McGill University, 1991) were screened with ^{32}P-dCTP-labeled inserts prepared from the *S. purpuratus* actin clone pSpG17 (Durica *et al.*, 1980; Cooper and Crain, 1982). The inserts from selected phage containing actin cDNA were subcloned into the plasmid vector Bluescript KS(+) (Stratagene) at the EcoR I site.

2.2.3 Isolation of genomic actin gene clones

A *L. pictus* genomic library was constructed essentially as described in PROTOCOLS AND APPLICATIONS GUIDE, Promega (second edition, 1991). *L. pictus* sperm DNA from a single male was partially digested with restriction enzyme Sau3A. The DNA was ligated into phage vector Gem12 (Promega). DNA fragments

ranging in size from 9-23 Kb can be cloned in this vector; no size fractionation is required in this cloning system. The recombinants were packaged with GIGAPACK II extracts (Stratagene), and titered with *E. coli* strain KW251 (Promega). The unamplified library was composed of 4×10^6 recombinant phages.

Since the genome size of sea urchins is about 5×10^8 to achieve a 99% probability of having a given DNA sequence represented in a library of 16 Kb fragments, the necessary number of recombinants need to screened is about 10^5 (Maniatis *et al.*, 1982). Approximately 10^5 phages were screened with ^{32}P-labeled pLpA7D3 inserts. The insert corresponds to 670 bp of the 5' coding sequence of the cDNA clone pLpA7 (Figure 2.1, d). Hybridizations were carried out in 6 × SSPE, 5 × Denhardt's solution, and 0.3% SDS at 60°C overnight. The filters were washed with 1 × SSC, 1% SDS at 60°C, and exposed to film. Selected clones were purified by additional two rounds of screening with the same probe.

2.2.4 Oligonucleotide primers used for PCR and sequencing

Oligonucleotides used were synthesized on an ABI 392 DNA/RNA Synthesizer using phosphoramidites and are listed in Table 2.1.

2.2.5 Subcloning

a) Subcloning the third exons

For simplicity, we refer to DNA sequences from codons 204 through 375 as "the third exon". For the cytoskeletal actin genes of *S. purpuratus*, this region corresponds to the third amino acid coding exon. For the known muscle actin genes of sea urchins, this region corresponds to the third and fourth amino acid coding exons.

To clone the third exons of the actin genes which are non-allelic with the two *L. pictus* actin cDNA clones (pLpA7 and pLpA5) initially classified as LpC1 and LpC2, *L. pictus* genomic actin DNA clones were prescreened to identify those corresponding to

Oligonucleotides used

Name	nt	Sequence (5' to 3')	Location	Types
#198	22	CGAGGATAAAAATCAACATCCA	3' non-coding regions of the genes LpC1 and LpC2	USP
#878	23	CCGGTGTATCATAGATTATTCTC	3' non-coding regions of the genes LpC1 and LpC2	DSP
#418	20	CTGAGCGTGAAATCGTTCGT	the beginning of the third exon	USP
#419	21	GAAGCACTTGCGGTGAACAAT	the end of the third exon	DSP
#403	19	ATCAACGGATAAGGGCTC	3' non-coding region of LpM	USP
#404	17	GTTAGACGCTCTCAAG	3' non-coding region of LpM	DSP
#420	21	ACTACTCAACCGTTACTTGTT	3' non-coding region of LpC4	USP
#219	22	ATGGATCGGAGGCTCTATCCTT	close to the end of the third exon	USP
T7	17	AATACGACTCACTATAG	upstream of Bluescript vector	USP
T3	17	ATTAACCCTCACTAAAG	downstream of Bluescript vector	DSP

USP: up stream primer; DSP: down stream primer.

Table 2.1

LpC1 and LpC2. The prescreening was done by PCR: a pair of oligonucleotide primers, Oligo#198 and Oligo#878 (Table 2.1), specifically amplify a fragment from the 3' noncoding regions of the LpC1 and LpC2 actin genes. Three genomic clones, λLpA4, λ LpB6 and λLpD3, which corresponded to neither LpC1 nor LpC2, were selected for further characterization. The EcoR I fragments which include the *L. pictus* actin DNAs were cloned into the Bluescript KS(+); the plasmid subclones were designated pLpA4, pLpB6 and pLpD3.

DNA sequences of the third exon were obtained by PCR amplifications with a pair of oligonucleotide primers (Oligo #418 and #419) which anneal to the 5' and 3' ends of the third exons (Figure 2.1). DNA prepared from the plasmids pLpA4, pLpB6 and pLpD3 were used as templates for PCR reactions. The DNA fragments amplified from pLpD3 and pLpA4 were subcloned into Bluescript KS(+) at the EcoR V site and designated as pLpD3E and pLpA4E (Figure 2.1, j and m).

b) Subcloning the 3' non-coding regions

The strategies for subcloning the 3' noncoding region of clones pLpA5, pLpA7, pLpB6, pLpD3 and pLpA4 are illustrated in Figure 2.1. The insertion sites in the Bluescript KS(+) vector, the orientation of the constructed subclones, and the assigned names for each construct are indicated on the maps.

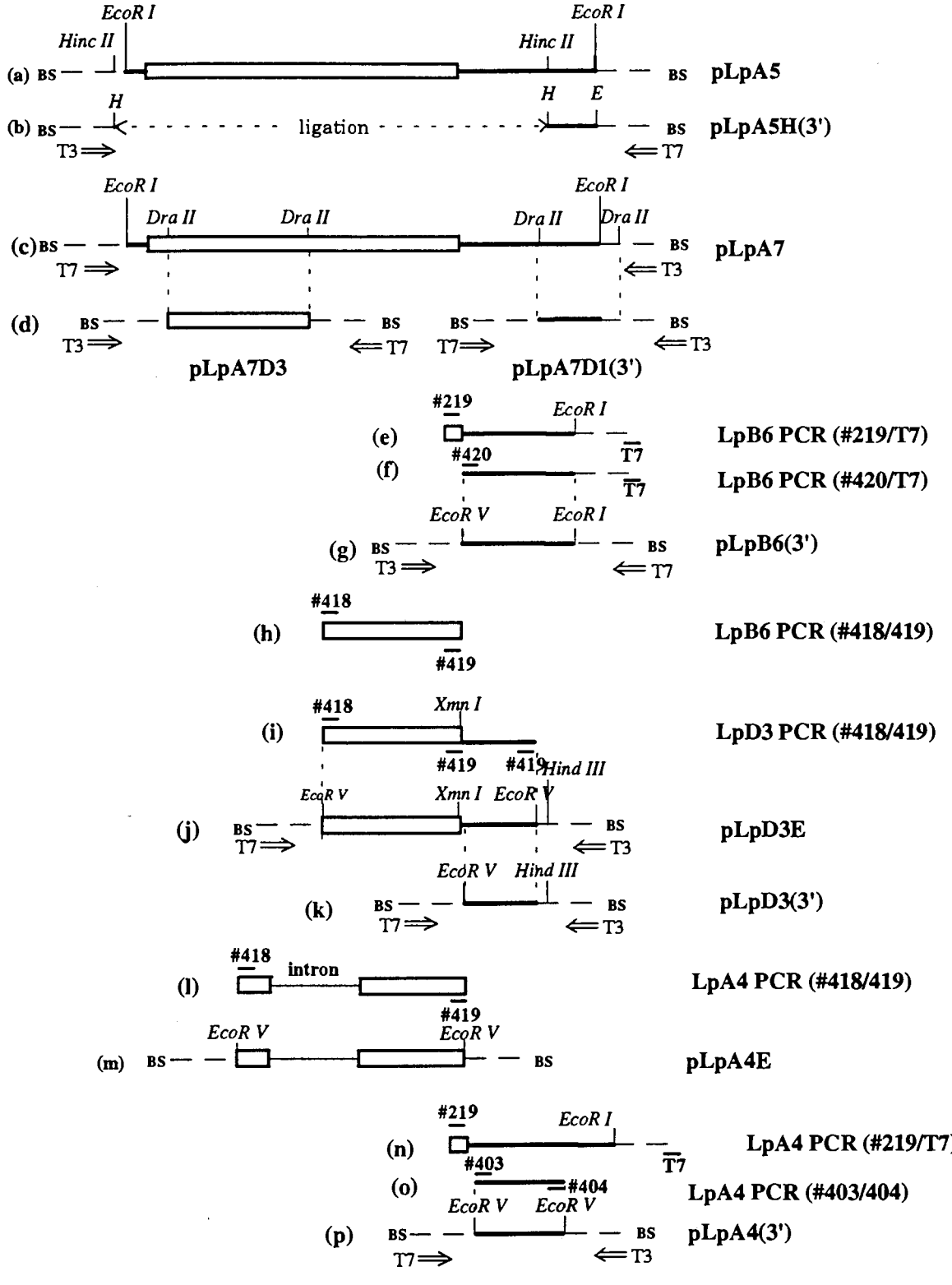
2.2.6 Nucleotide sequencing

Sequencing of plasmids pLpA5 and pLpA7 was done manually by using the dideoxynucleotide chain termination methods with the Sequenase Sequencing Kit (US Biochemicals). Sequencing reactions for plasmids pLpA4E and pLpD3E were performed using the Taq Dye Primer Cycle Sequencing Kit (ABI). Two different clones of pLpA4E and three of pLpD3E were sequenced from both directions, with T3 and T7 primers, to eliminate possible artifacts generated by PCR or sequencing. The LpB6 PCR (418/419)

Fig. 2.1 Construction of subclones of the third exon and 3' noncoding sequence. Strategies for subcloning are schematically shown in diagrams. Diagrams are not drawn to scale. Only the restriction sites involved in subcloning are shown in diagrams. T7 and T3 represent primers of T7 and/or T3 RNA polymerase in Bluescipt (KS+) DNA, and are used to indicate the orientation of each subclone together with arrows which indicate transcription orientations. Numbers (#) placed above or beneath the short solid lines indicate specific oligonucleotide primers used for the PCR reactions (see Table 2.1).

 : a.a coding region; —— : 5' or 3' noncoding region;– : PCR primer;

— — : Bluescript DNA



▭: a.a coding region; ▬: 5' or 3' noncoding region; — -: Bluescript DNA; —: PCR primer.

fragment was sequenced directly using Taq DyeDeoxy™ Terminator Cycle Sequencing Kit (ABI) from both directions. Automated sequencing was performed on the ABI 373 DNA Sequencer (ABI).

2.3 RESULTS

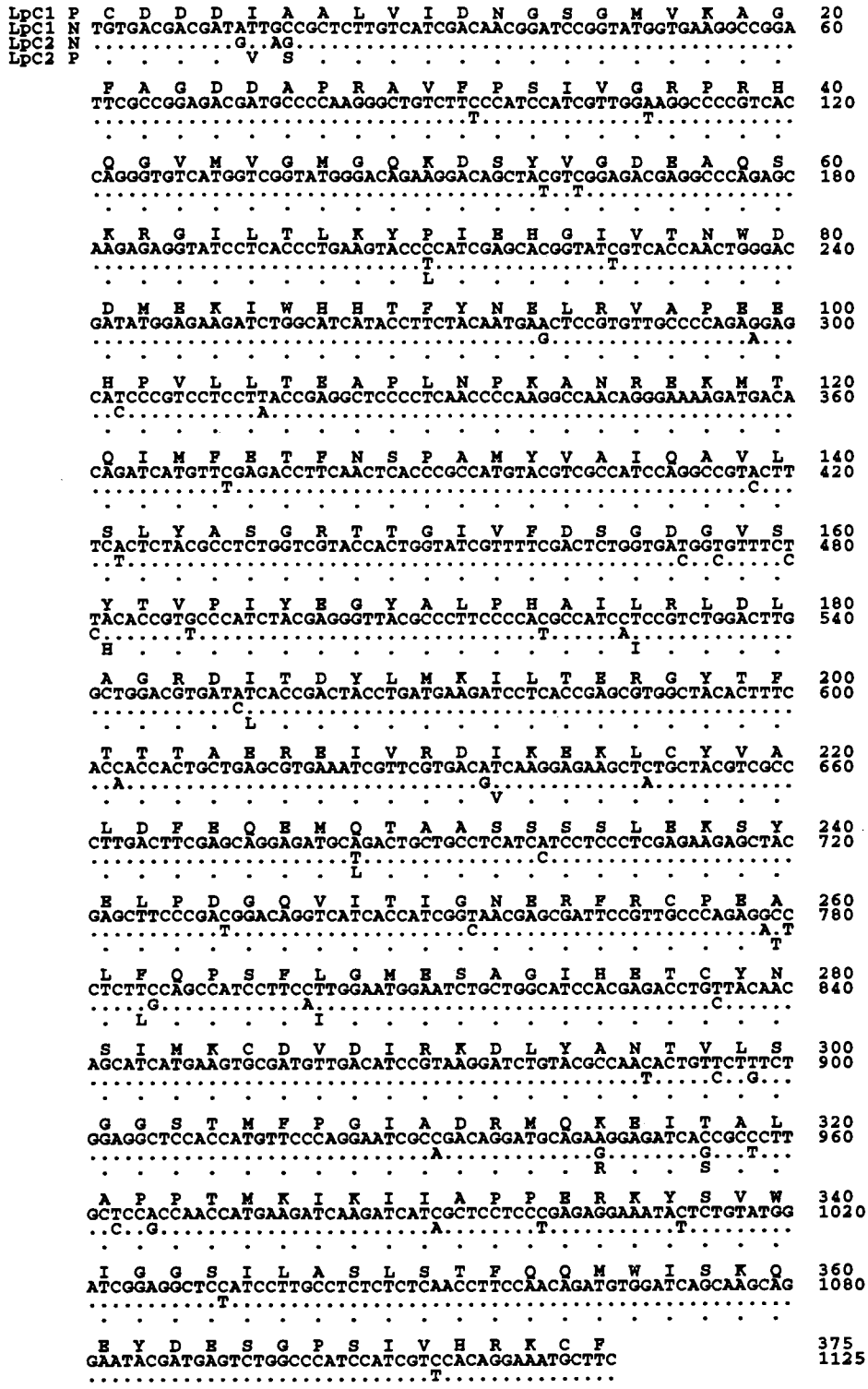
2.3.1 Entire coding sequences of two *L. pictus* actin mRNAs

Actin cDNA clones of *L. pictus* were isolated by screening two gastrula stage cDNA libraries with a *S. purpuratus* actin cDNA clone, pSpG17 (Durica *et al.*, 1980). The inserts of two selected *L. pictus* recombinant plasmids, pLpA7 and pLpA5, were sequenced and correspond to *L. pictus* actin genes referred to henceforth as LpC1 and LpC2, respectively. The nucleotide coding sequences and deduced amino acid sequences of the LpC1 and LpC2 genes are shown in Fig. 2.2. The identification of the coding sequences was based on comparison with the *S. purpuratus* cytoskeletal actin genes which can be aligned with no gaps. The LpC1 and LpC2 actin genes contain an 1125-nt coding region having conventional ATG start and TAA stop codons and encode proteins of 375 amino acids. The nucleotide coding sequences of LpC1 and LpC2 are 95% similar. Comparison of the coding regions of LpC1 and LpC2 to several cytoplasmic and muscle actin genes of sea urchins and a sea star indicate that both LpC1 and LpC2 encode cytoskeletal actins (see Chapter III).

2.3.2 Identification and characterization of the remaining *L. pictus* actin genes

Genomic DNA blot hybridization assays indicated the presence of 4-5 actin genes per *L. pictus* haploid genome (see below). However, the sequence of each of several

Fig. 2.2 Comparison of coding regions of the LpC1 and LpC2 genes. Nucleotide (N) and deduced amino acid (P) sequences of the LpC1 and LpC2 coding region are shown.

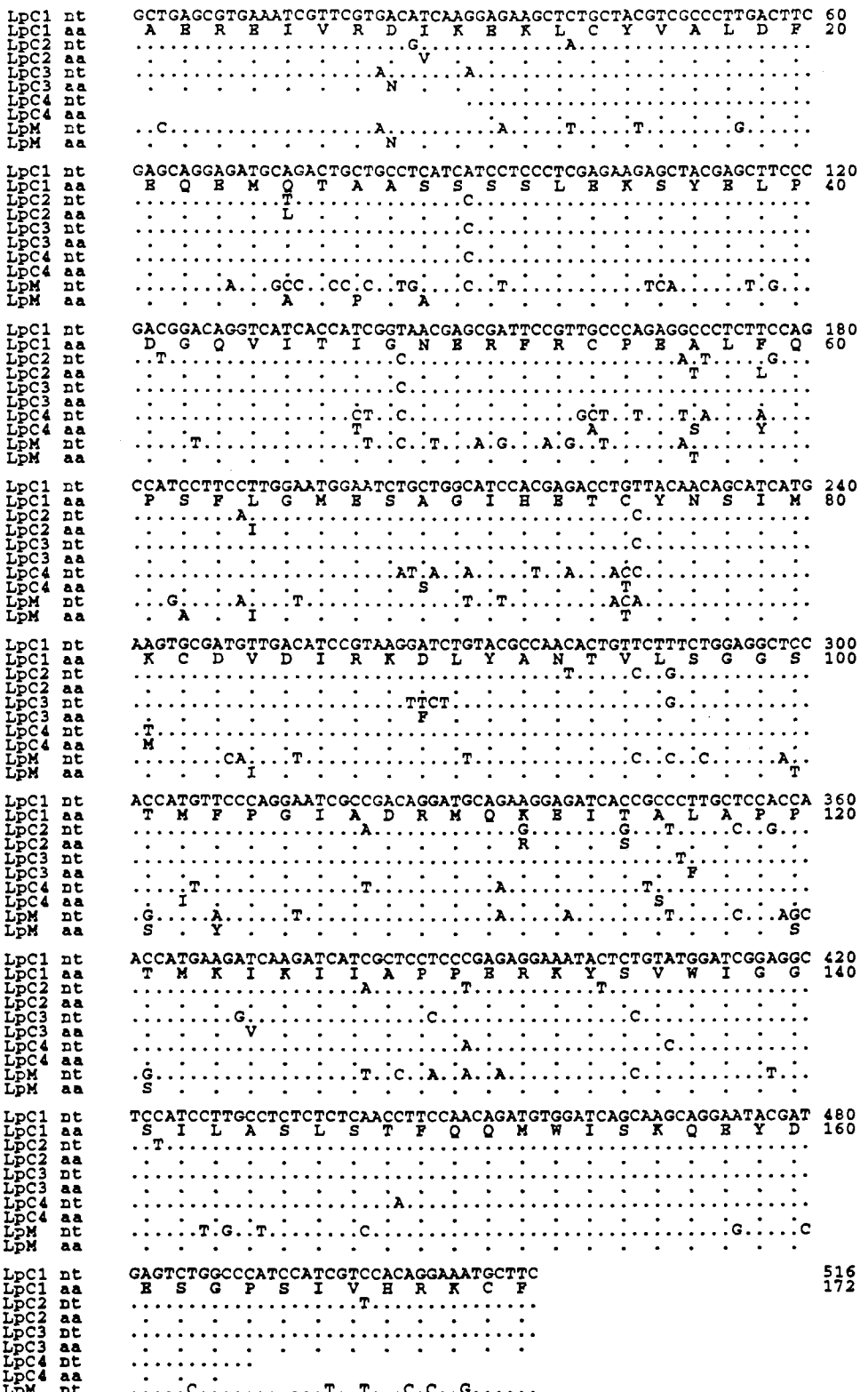

LpC1 P C D D D I A A L V I D N G S G M V K A G 20
LpC1 N TGTGACGACGATATTGCCGCTCTTGTCATCGACAACGGATCCGGTATGGTGAAGGCCGGA 60
LpC2 N ..........G..AG.....
LpC2 P . . . . V S . . . . . .
F A G D D A P R A V F P S I V G R P R H 40
TTCGCCGGAGACGATGCCCCAAGGGCTGTCTTCCCATCCATCGTTGGAAGGCCCCGTCAC 120
.....T.....T.....
. . . . . .
Q G V M V G M G Q K D S Y V G D E A Q S 60
CAGGGTGTCATGGTCGGTATGGGACAGAAGGACAGCTACGTCGGAGACGAGGCCCAGAGC 180
.....T..T.....
. . . . . .
K R G I L T L K Y P I E H G I V T N W D 80
AAGAGAGGTATCCTCACCCTGAAGTACCCCATCGAGCACGGTATCGTCACCAACTGGGAC 240
.....T.....T.....
. . . . . L . . . . . .
D M E K I W H H T F Y N E L R V A P E E 100
GATATGGAGAAGATCTGGCATCATACCTTCTACAATGAACTCCGTGTTGCCCCAGAGGAG 300
.....G.....A...
. . . . . .
H P V L L T E A P L N P K A N R E K M T 120
CATCCCGTCCTCCTTACCGAGGCTCCCCTCAACCCCAAGGCCAACAGGGAAAAGATGACA 360
..C.....A.....
. . . . . .
Q I M F E T F N S P A M Y V A I Q A V L 140
CAGATCATGTTCGAGACCTTCAACTCACCCGCCATGTACGTCGCCATCCAGGCCGTACTT 420
.....T.....C...
. . . . . .
S L Y A S G R T T G I V F D S G D G V S 160
TCACTCTACGCCTCTGGTCGTACCACTGGTATCGTTTTCGACTCTGGTGATGGTGTTTCT 480
..T.....C..C.....C
. . . . . .
Y T V P I Y E G Y A L P H A I L R L D L 180
TACACCGTGCCCATCTACGAGGGTTACGCCCTTCCCCACGCCATCCTCCGTCTGGACTTG 540
C.....T.....T.....A.....
H . . . . . I . . . . .
A G R D I T D Y L M K I L T E R G Y T F 200
GCTGGACGTGATATCACCGACTACCTGATGAAGATCCTCACCGAGCGTGGCTACACTTTC 600
.....C.....
. . . . . L . . . . . .
T T T A E R E I V R D I K E K L C Y V A 220
ACCACCACTGCTGAGCGTGAAATCGTTCGTGACATCAAGGAGAAGCTCTGCTACGTCGCC 660
..A.....G.....A.....
. . . . . V . . . . . .
L D F E Q E M Q T A A S S S S L E K S Y 240
CTTGACTTCGAGCAGGAGATGCAGACTGCTGCCTCATCATCCTCCCTCGAGAAGAGCTAC 720
.....T.....C.....
. . . . . L . . . . . .
E L P D G Q V I T I G N E R F R C P E A 260
GAGCTTCCCGACGGACAGGTCATCACCATCGGTAACGAGCGATTCCGTTGCCCAGAGGCC 780
.....T.....C.....A.T
. . . . . T
L F Q P S F L G M E S A G I H E T C Y N 280
CTCTTCCAGCCATCCTTCCTTGGAATGGAATCTGCTGGCATCCACGAGACCTGTTACAAC 840
.....G.....A.....C.....
. L . . . . I . . . . . .
S I M K C D V D I R K D L Y A N T V L S 300
AGCATCATGAAGTGCGATGTTGACATCCGTAAGGATCTGTACGCCAACACTGTTCTTTCT 900
.....T.....C..G...
. . . . . .
G G S T M F P G I A D R M Q K E I T A L 320
GGAGGCTCCACCATGTTCCCAGGAATCGCCGACAGGATGCAGAAGGAGATCACCGCCCTT 960
.....A.....G.....G...T...
. . . . . R . . S . .
A P P T M K I K I I A P P E R K Y S V W 340
GCTCCACCAACCATGAAGATCAAGATCATCGCTCCTCCCGAGAGGAAATACTCTGTATGG 1020
..C..G.....A.....T.....T.....
. . . . . .
I G G S I L A S L S T F Q Q M W I S K Q 360
ATCGGAGGCTCCATCCTTGCCTCTCTCTCAACCTTCCAACAGATGTGGATCAGCAAGCAG 1080
.....T.....
. . . . . .
E Y D E S G P S I V H R K C F 375
GAATACGATGAGTCTGGCCCATCCATCGTCCACAGGAAATGCTTC 1125
.....T.....
. . . . . .

selected gastrula stage cDNA clones corresponded to either the LpC1 or LpC2 genes. To clone the remaining *L. pictus* actin genes I made and screened a recombinant phage library of genomic DNA by plaque hybridization to LpC1 DNA. The collection of actin recombinant phage was further classified by PCR amplification of portions of the third exon (using primers 418 and 419; see Methods), where most of the amino acid substitutions occur, as well as portions of the 3' noncoding region (primers involved are listed in Table 2.1; also see Fig. 2.1). Sequence analysis of the PCR fragments resulted in the identification of 3 more *L. pictus* actin genes having distinctive features.

The PCR fragment cloned into plasmid pLpB6 included 516 bp corresponding to the third exon of a cytoskeletal actin gene; the gene is referred to as LpC4. The sequence of the pLpD3E indicates that the downstream PCR primer #419, instead of annealing to the 3' end of the third exon, annealed to a 21 bp sequence located 243-263 nt following the stop codon; the resulting plasmid includes 263 bps corresponding to the 3' noncoding region of this gene. The deduced protein has cytoskeletal actin features and the gene is referred to as LpC3. The PCR fragment of the plasmid pLpA4E has a sequence most similar to the third exon of muscle actin genes and includes an intron (Figure 2.1, l); we refer to the corresponding gene as LpM. All other genomic actin DNA clones characterized could be classified as one of these five *L. pictus* actin genes.

Alignment of the nucleotide sequences of the third exons of all five *L. pictus* actin genes, LpC1, LpC2, LpC3, LpC4, and LpM, is presented in Figure 2.3. The amino acid sequences of all five actin genes were deduced from nucleotide sequences based on the universal code. The amino acid sequences are compared and discussed in Chapter III.

Fig. 2.3 Comparison of nucleotide sequences (nt) and deduced amino acid sequences (aa) of the third exon of cytoskeletal and muscle actin genes of the sea urchin *L. pictus*. Coding sequences from nt 610 through nt 1125, corresponding to amino acid codons 204-375, of each actin genes are compared to that of *L. pictus* LpC1 actin gene. Only the differences are shown, with similar nucleotides and amino acids denoted by dots.



LpC1 nt GCTGAGCGTGAAATCGTTCGTGACATCAAGGAGAAGCTCTGCTACGTCGCCCTTGACTTC 60
LpC1 aa A E R E I V R D I K E K L C Y V A L D F 20
LpC2 nt .....G.....A.....
LpC2 aa .....V.....
LpC3 nt .....A.....A.....
LpC3 aa .....N.....
LpC4 nt .....
LpC4 aa .....
LpM nt ..C.....A.....A....T....T.....G.....
LpM aa .....N.....

LpC1 nt GAGCAGGAGATGCAGACTGCTGCCTCATCATCCTCCCTCGAGAAGAGCTACGAGCTTCCC 120
LpC1 aa E Q E M Q T A A S S S S L E K S Y E L P 40
LpC2 nt .....T.....C.....
LpC2 aa .....L.....
LpC3 nt .....C.....
LpC3 aa .....
LpC4 nt .....C.....
LpC4 aa .....
LpM nt .....A...GCC..CC.C..TG...C..T.....TCA.....T.G...
LpM aa .....A.....P..A.....

LpC1 nt GACGGACAGGTCATCACCATCGGTAACGAGCGATTCCGTTGCCCAGAGGCCCTCTTCCAG 180
LpC1 aa D G Q V I T I G N E R F R C P E A L F Q 60
LpC2 nt ..T.....C.....A.T.....G...
LpC2 aa .....T.....L..
LpC3 nt .....C.....
LpC3 aa .....
LpC4 nt .....CT..C.....GCT..T...T.A....A....
LpC4 aa .....T.....A.....S.....Y..
LpM nt .....T.....T..C..T...A.G...A.G..T.....A.....
LpM aa .....T.....

LpC1 nt CCATCCTTCCTTGGAATGGAATCTGCTGGCATCCACGAGACCTGTTACAACAGCATCATG 240
LpC1 aa P S F L G M E S A G I H E T C Y N S I M 80
LpC2 nt .....A.....C.....
LpC2 aa .....I.....
LpC3 nt .....C.....
LpC3 aa .....
LpC4 nt .....AT.A..A.....T..A...ACC.....
LpC4 aa .....S.....T.....
LpM nt ...G.....A...T.....T..T.....ACA.....
LpM aa .A...I.....T.....

LpC1 nt AAGTGCGATGTTGACATCCGTAAGGATCTGTACGCCAACACTGTTCTTTCTGGAGGCTCC 300
LpC1 aa K C D V D I R K D L Y A N T V L S G G S 100
LpC2 nt .....T.....C..G.....
LpC2 aa .....
LpC3 nt .....TTCT.....G.....
LpC3 aa .....F.....
LpC4 nt .T.....
LpC4 aa M.....
LpM nt .....CA...T.....T.....C..C..C.....A..
LpM aa .....I.....T

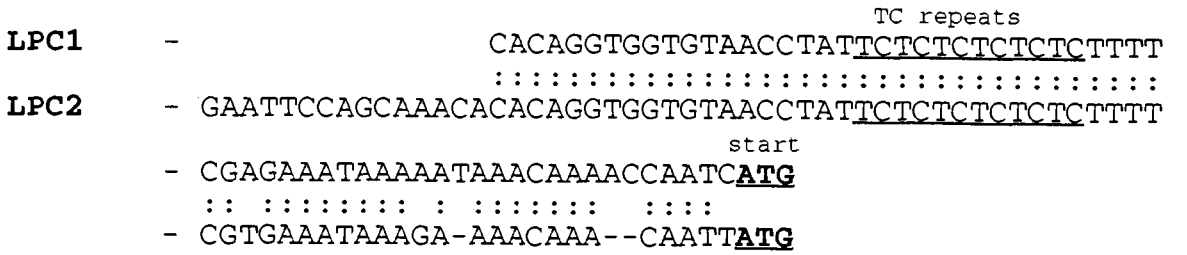
LpC1 nt ACCATGTTCCCAGGAATCGCCGACAGGATGCAGAAGGAGATCACCGCCCTTGCTCCACCA 360
LpC1 aa T M F P G I A D R M Q K E I T A L A P P 120
LpC2 nt .....A.....G.....G...T.....C..G...
LpC2 aa .....R...S.....
LpC3 nt .....T.....
LpC3 aa .....F.....
LpC4 nt ....T.....T.....A.....T.....
LpC4 aa .I.....S.....
LpM nt .G.....A.....T.....A.....A.....T.....C...AGC
LpM aa S.Y.....S

LpC1 nt ACCATGAAGATCAAGATCATCGCTCCTCCCGAGAGGAAATACTCTGTATGGATCGGAGGC 420
LpC1 aa T M K I K I I A P P E R K Y S V W I G G 140
LpC2 nt .....A.....T.....T.....
LpC2 aa .....
LpC3 nt .....G.....C.....C.....
LpC3 aa .....V.....
LpC4 nt .....A.....C.....
LpC4 aa .....
LpM nt .G.....T..C..A..A..A.....C.....T...
LpM aa S.....

LpC1 nt TCCATCCTTGCCTCTCTCTCAACCTTCCAACAGATGTGGATCAGCAAGCAGGAATACGAT 480
LpC1 aa S I L A S L S T F Q Q M W I S K Q E Y D 160
LpC2 nt ..T.....
LpC2 aa .....
LpC3 nt .....
LpC3 aa .....
LpC4 nt .....A.....
LpC4 aa .....
LpM nt .....T.G..T.....C.....G.....C
LpM aa .....

LpC1 nt GAGTCTGGCCCATCCATCGTCCACAGGAAATGCTTC 516
LpC1 aa E S G P S I V H R K C F 172
LpC2 nt .....T.....
LpC2 aa .....
LpC3 nt .....
LpC3 aa .....
LpC4 nt .....
LpC4 aa .....
LpM nt .....C.....T..T...C.C..G.....
LpM aa .....

Fig. 2.4 Alignment of the 5' noncoding nucleotide sequences of LpC1 and LpC2. The initiation codon ATG and the 12 base TC repeats are underlined. Sequences compared have 90.5% identity.



Identity : 57 (90.5%)

2.3.3 The 5' noncoding regions of the actin genes LpC1 and LpC2

A comparison of the 5' noncoding regions of the LpC1 and LpC2 genes is shown in Figure 2.4. The 63 nt of LpC1 and 60 nt of LpC2 immediately 5' to the start codon are 95% similar if two gaps are introduced in the LpC2 sequence. There is sequence identity between the two genes from 27 to 63 nt 5' to the initiation codon. The LpC1 and LpC2 actin genes include 12 bp of TC repeats from 33 to 44 nt 5' to the initiation codon; this element is found in *S. purpuratus* actin genes (Akhurst *et al.* 1987; Flytzanis *et al.* 1989).

2.3.4 The 3' noncoding regions of *L. pictus* actin genes

The 3' noncoding sequences of the genes LpC1 and LpC2 were obtained by sequencing the cDNA clones; the alignment of the two sequences is shown in Fig. 2.5. The two sequences are 89% similar, if 10 gaps are introduced for LpC1 and 11 for LpC2; the alignment gaps range from 1 to 9 bp. These observations indicate that the 3' noncoding sequences of these two genes are highly similar but that many insertion or deletion events have occurred as the genes diverged.

The 3' noncoding regions of LpC3, LpC4 and LpM have been partially sequenced, and include 263 bps, 443 bps and 280 bps, respectively shown in Figure 2.6.

2.3.5 Number of genes in the actin gene family of *L. pictus*

Sperm genomic DNA from three individuals (A, B, C) was digested with five restriction enzymes. The blotted DNA was hybridized with the labeled insert of pLpA7D3, which corresponds to the 5' side of the coding region of the LpC1 actin gene (Fig. 2.1, d), under two criteria of stringency: (1) 0.33 M Na^+ (2 X SSC), 55°C or (2) 0.033 M Na^+ (0.2 X SSC), 65°C. DNA prepared from individuals A and B show virtually

Fig. 2.5 Alignment of the 3' noncoding regions of *L. pictus* LpC1 and LpC2 genes. 681 nt of LpC1 and 672 nt of LpC2 immediately distal to the stop codon are aligned. Dashes represent gaps introduced to optimize the alignment. Alignment was performed by using the NALIN program of the PCGene package.

```
LPC1  - ACAACTCATTTT---GTTTTACTTCTGTTGAGCACAATAACGAACTCCTG -47
LPC2  - ACAACTCGTCTTTTCGTTTTACTTCTGTTGAGCACAATAACG-------- -42

      - GGGTGGGTTAAGTGGGTCACCTTTATCAAATGGGAAAATTCTGGAACAGT -97
      - GGGTGGGTAAAGTGGGTCACCTTTATTAAAATGGAAAATTCTGGAACAGT -92

      - TGGGTTGATCAGGAATTATCTAGCTATCCATCCCCGATAATAGTCGCATC -147
      - TAGGT-GATCAGGAATTATCTAGCTATCGATCTCCGATAATAGTC-CATT -140

      - ATGGCTATTGATTGCTCAAACACTTGACATATTATGGCGATTTCTTTGCC -197
      - CTGGTTATCAATTGCTCAAACACTCGACATA--ATGTCGATTTCTTTGCC -188

      - CATACTGGTGACCACTTTACCA---------ATTTGTTAAC----ATAAA -234
      - CATATTGGTGACCACTTTACCATGTTTAGGAATTTGTTAAAGAAAATAAA -238

      - AACAAATAGGCCCTCAACCTGCCTCAACTTAGGCAAA-GTAAAGAGAGAG -283
      - ATCAAATATGTCCTGAACCTGCCTCAACTTGGGCAAAAGTAAAGAGAGAG -288

      - AAATATTGTGAGAAAAGGTTGACATTATCTTCTGGTTCTGAGGTTGGCAA -333
      - AACTGTTGTGAGAAAAGGTTGACATTATCTTCTGGATCCGAGGTTGATAA -338

      - TGAGGATAAAAA-TCAACATCCAATTGATCCAAACTTATTCTAACTTAAT -382
      - TGAGGGTAAAAAATCAACATCCAA--------AACTTATTCTAACTTAAT -380

      - AAAAAAATAGTTCCCTTTTAGTTCAACAATTTTTTGTTCAACTTAAATTT -432
      - AAAAAA-TAGTTCCCTTT-AGTTCACTAATTT--TGTTCAACTAAAATTT -426

      - CTTTGAGCCAATCCGTGATTCTACGGAATAACATTCTTGCAGA-GCACAC -481
      - CTTTGAGCCAATCCGTGATACAACGGAATAACATTCTTGCAAAAGCACAC -476

      - GATTTTGTATCTTTATAT----TTATAAGTAAATTGCTTTAGAGAAA-CT -526
      - GATTTTGTATCTTTATATATATTTATAAGTAAATTGCTTTAGATAAAACT -526

      - TATGATTGCATTATAATCTTGGAAAGAAATCTTAAAGAGAATAATCTATG -576
      - TATGATTGCATTATAATCTT------AAA----AAAAAGAATAATCTATG -566

      - ATACATAATGATAAAAATGCGCGTTC-AACTTGGATCATGCATATTTTTG -625
      - ATACATAATGATAAAAACGCGCGTTTTAACTTGGATTATGCATATTTTTG -616

      - TTTGATATGTACACTGTATATATTTTGTAACTCTAAAAAAAAAA-AAAAA -674
      - TTTGATATG-ACACTAAATATATTTTGTAACTATAAAAATATACGAAAAA -665

      - AAAGGAA -681
      - AAAAAAA -672

Identity: 597 ( 88.8%)
```

Fig. 2.6 Nucleotide sequences of the 3' noncoding regions of *L. pictus* cytoskeletal actin genes LpC3, LpC4 and LpM. (a) 264 nt of the LpC3 gene immediately distal to the stop codon. A suspected EcoR I or EcoR I* restriction site is underlined (see Results for discussion).(b) 443 nt of the LpC4 gene immediately distal to the stop codon. (c) 316 nt of the LpM gene immediately distal to the stop codon.

(a): the 3' non-coding region of LpC3

```
          10         20         30         40         50
           |          |          |          |          |
  1 ACTTTGAAAAACTCTCCTATAAAAACTGTGCAATGCGATGTTTTTTATTT
 51 TCTCCTCACCAGACATAAGTCGTAGCAAAAATAAAATCATACCCTGAATA
101 AGTGGTAGATCACTTCCTAACTCTGCTTCTCCTAAAATTGGTGGAAATGA
151 CTGTTATACCGTTAAATCAGATATGTTTATCACAACAATATAATAGTACT
201 ATTATATAATACTTGAATGATTACTGGGGAATTGTGTTATAGATTGTTCA
251 CCGCAAGTGCTTC
```

(b): the 3' non-coding region of LpC4

```
          10         20         30         40         50
           |          |          |          |          |
  1 ACTACTCAACCGTTACTTGTTGAACAAACACTCGAACATTACCATTTTCA
 51 GTGTGAGAAATCTTGGGAGATAGAAGAGACTTCCATCTCCATGAGATTCA
101 TTGTCTTGTTTTTGATGGTTCATCTAGTTGTCACAGTAAATTCAATAGAA
151 GGACAACCAAACCGATGCATTCTTAAAAGTTGGATACTGCCAAGATAAAT
201 TGGTTTGTGCCAAGTTAACCAGGAAGATATTCATATAATTNGTTTCCCAT
251 TTTCTTGAAAATTGTGCATTTTCCTTTTTTGGGAAGTTTTGCAACATTAT
301 TCCTACCAAATTTGCAAACATTTTGGCCTAGTATTTATTTTTGTTCTACC
351 TGAATGCCGAAATTCTGCCAAAGGAACCTTGGATTATTTTGGATCACTTT
401 TGAAAGGTNACATTTTCGGCCCCTTGCAATTCCTGGTCCTGGT
```

(c): the 3' non-coding region of LpM

```
          10         20         30         40         50
           |          |          |          |          |
  1 TACATCAACGGATAAGGGCTCACTGGTCTAGGAGGGCTGACATTGGCAGT
 51 TATTCTTTGTGNAANCTGTAGTCGGAGCCCTGTAGGNTCTATATTAATTT
101 TAGAGTATCATTTATAGAATATTGTGACGTCACACTCTTTCTGTCATTCA
151 NCTTCCAAGATNACGCCCACGACATGGGTCGCCCCTCAGGGGTAAGCATT
201 GAGAGCGTCTAACTAAGGTTGATTATGTTGAATGTCCGATTTTCTTACAT
251 GATGGATCCTCTAAATGATGGAAAGTCAACTGGATATAAGATCTAATGGA
301 TTGGTAATTGGAATTC
```

identical patterns of fragments for all of five restriction enzymes used (Fig. 2.7); this is very unusual for sea urchins, which normally show a high level of DNA polymorphism (Britten *et al.*, 1979; Lee *et al.*, 1984; Minor *et al.*, 1987). However, the pattern of restriction fragments shown for individual C is very different from that of A and B (Fig. 2.7).

To assess the number of actin genes in the *L. pictus* genome, genomic DNAs from urchins B and C were digested with EcoR I and blotted. The blots were hybridized with the actin coding sequence (pLpA7D3, Figure 2.1, d), and with five 3' probes selective for the actin genes LpC1, LpC2, LpC3, LpC4 and LpM (Figure 2.1, d, b, k, g and p). Hybridizations were carried out at two criteria of stringency. Figure 2.8 shows that the individual B is homozygous for the EcoR I restriction fragments of the gene LpC2 (panel c, B), but heterozygous for the genes LpC1 (panel b, B), LpC3 (panel d, B), LpC4 (panel e, B) and LpM (panel f, B); individual C is homozygous for the genes LpC2 (panel c, C), LpC3 (panel d, C), LpC4 (panel e, C) and LpM (panel f, C), but heterozygous for the gene LpC1 (panel b, C). Cross-hybridization at low stringency (0.33 M Na^+, 55°C) was observed between the EcoR I fragments of the genes LpC1 and LpC2 (data not shown), not surprising given the similarity of the hybridization probes. When hybridization reactions were carried out at higher stringency (0.033 M Na^+, 65°C), these signals were differentially retained (panel b, c). Hybridization with the 3' probe specific for the gene LpC3 revealed four different EcoR I fragments represented in the two haploid genomes of individual B (Figure 2.8, panel d, B). We surmise that there is an EcoR I site in the 3' non-coding region of the gene LpC3 and this site is included in the probe we used (pLpD3(3')). Two heterozygous alleles should show four EcoR I fragments when hybridized with the pLpD3(3') probe if it includes an EcoR I site. While no EcoR I site was identified from the sequence of the pLpD3(3') hybridization probe, a six bp sequence (GAATTG) having a single mismatch compared to the EcoR I site (GAATTC) was

Fig. 2.7 Genomic DNA gel blot analyses of actin genes. Sperm genomic DNA prepared from three *L. pictus* sea urchins was digested with the restriction enzymes, EcoR I, Hind III, Pst I, Xba I and Xho I, and hybridized the probe pLpA7D3, which corresponds to a coding region of the actin gene LpC1. A, B, and C represent each of three sea urchins. Hybridization conditions are described in Methods and Results.

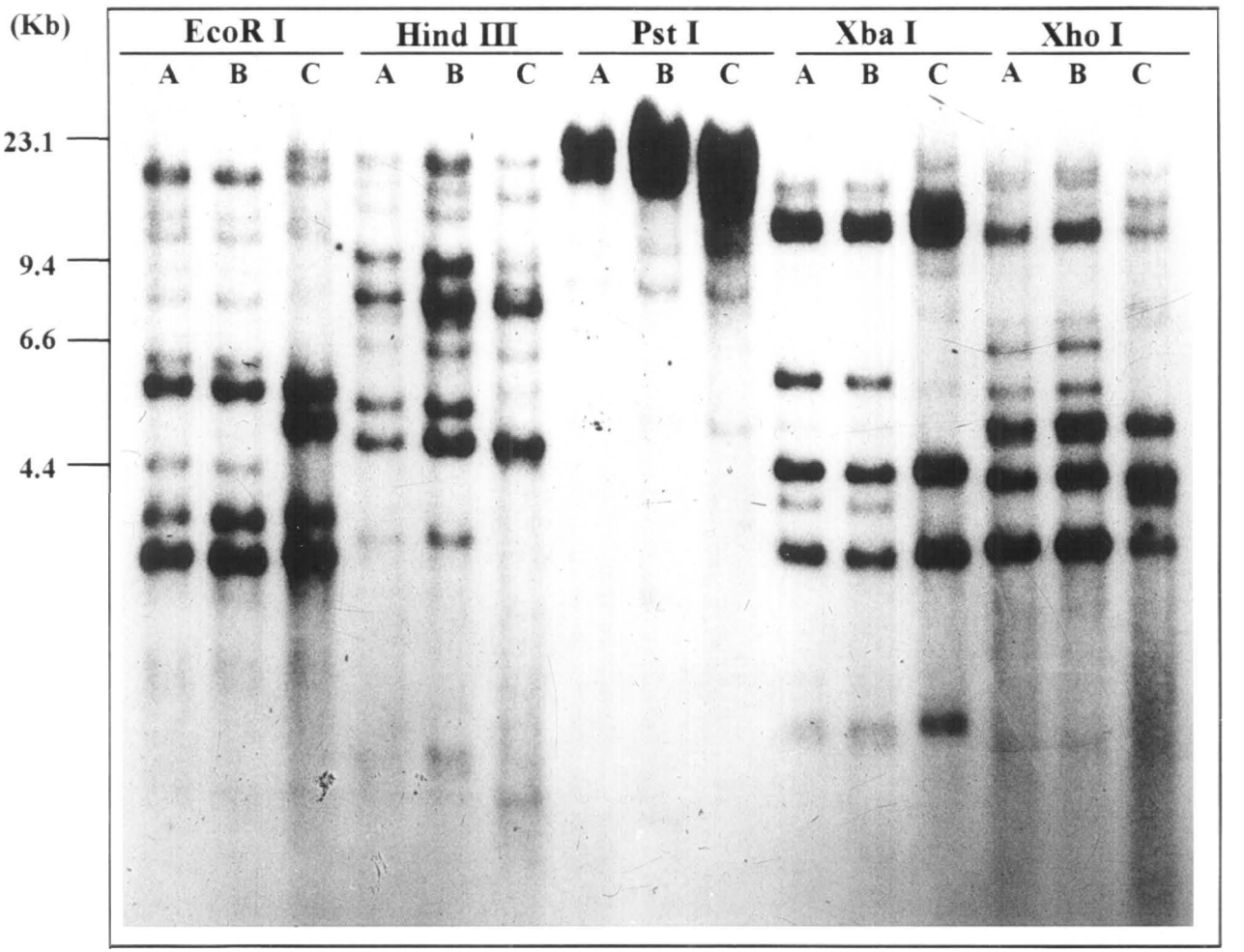

(Kb)
EcoR I
Hind III
Pst I
Xba I
Xho I
A B C A B C A B C A B C A B C
23.1
9.4
6.6
4.4

Fig. 2.8 Analysis of the *L. pictus* actin gene family with gene-specific hybridization probes. Sperm DNA of individuals B and C digested with EcoR I was used. Blots were hybridized with actin gene coding region probe (pLpA7D3) (panel a) or gene specific probes, pLpA7D1(3') for LpC1 gene (panel b), pLpA5H(3') for LpC2 gene (panel c), pLpD3(3') for LpC3 gene (panel d), pLpB6(3') for LpC4 gene (panel e) and pLpA4(3') for the muscle actin gene LpM (panel f). DNA fragments detected by the coding sequence probe (panel a) are labeled for each corresponding actin gene and shown on the left of panel (a). Hybridization was carried out at 65°C, 0.033 M Na^+ (0.2 X SSC).

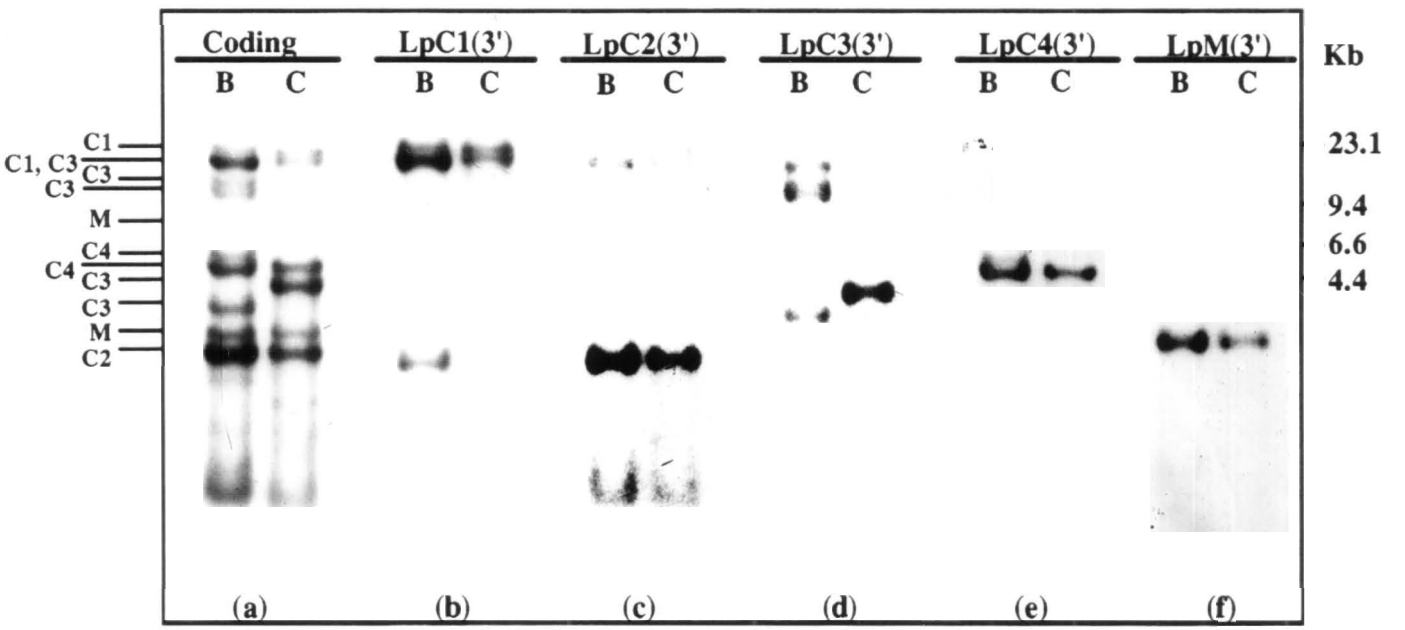
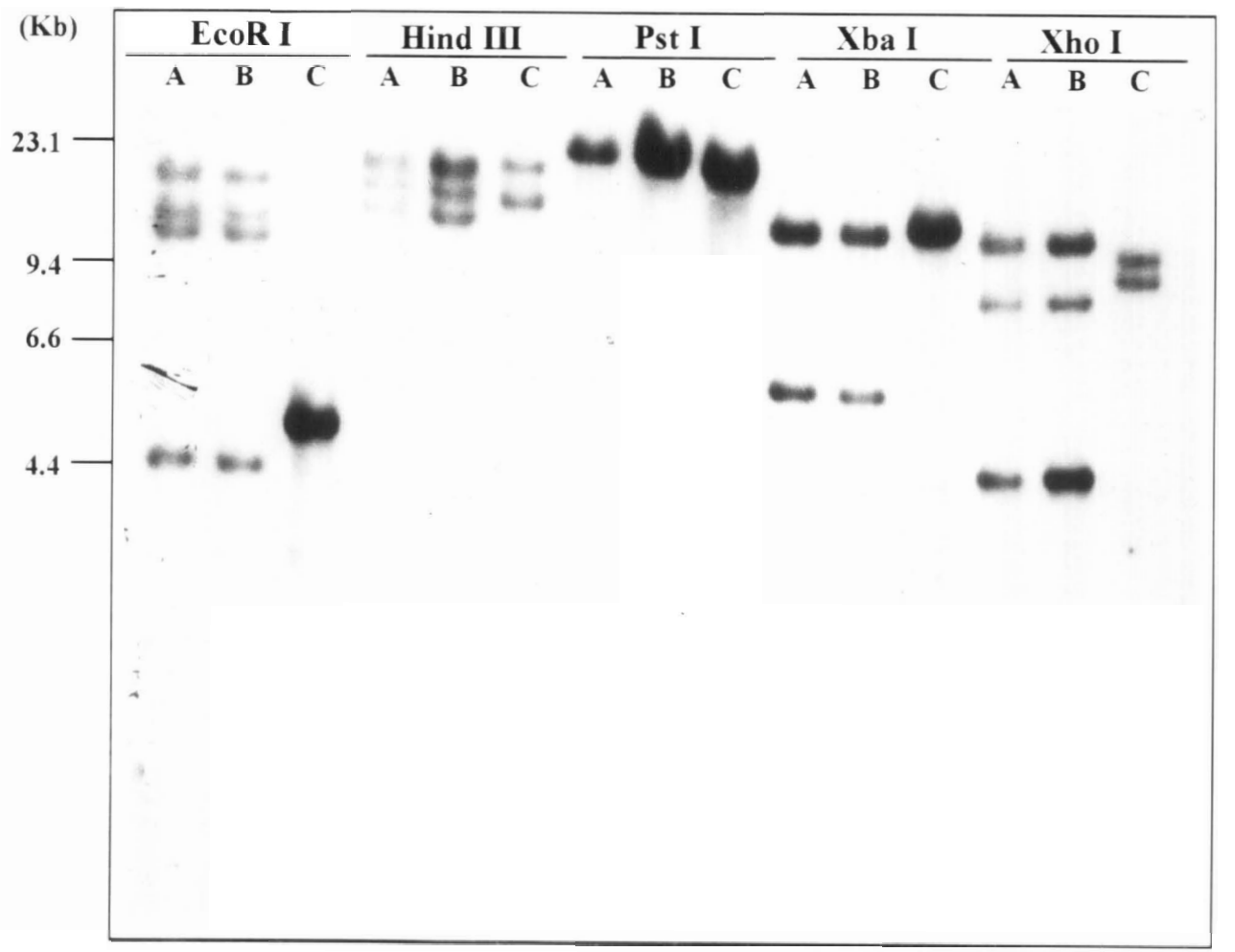

Coding
LpC1(3')
LpC2(3')
LpC3(3')
LpC4(3')
LpM(3')
B C
Kb
23.1
9.4
6.6
4.4
C1
C1, C3
C3
C3
M
C4
C4
C3
C3
M
C2
(a)
(b)
(c)
(d)
(e)
(f)

Fig. 2.9 Genomic DNAs from three sea urchins A, B, and C were digested with EcoRI, Hind III, Pst I, Xba I and XhoI, and hybridized with probe pLpD3(3') specific for the LpC3 gene at 65°C, 0.033 M Na^+ (0.2 X SSC).


(Kb)
EcoR I
Hind III
Pst I
Xba I
Xho I
A B C A B C A B C A B C A B C
23.1
9.4
6.6
4.4

present (Fig. 2.7). We suspect that the sequence mismatch may be a PCR artifact. Alternatively, Eco R1* activity (which recognizes AATT) might account for digestion at this site. A blot of genomic DNA's digested with 5 enzymes was hybridized with the probe pLpD3(3') at a criterion of stringency of 0.033 M Na^+, 65°C. Results shown in Figure 2.9 indicate that there is only one copy of the gene LpC3 per haploid genome. In summary, each haploid *L. pictus* genome contains a single copy of each of the actin genes LpC1, LpC2, LpC3, LpC4 and LpM. No other actin genes were detected by Southern blotting under conditions which allowed considerable cross-hybridization between the quite divergent muscle and cytoskeletal actin coding sequences.

2.4 DISCUSSION

Previous studies have shown that the organization of the actin gene family varies among species in gene number, gene linkage, and intron position (Johnson *et al.*, 1983). The number of actin genes per haploid genome varies widely: 1 for *Saccharomyces cerevisiae* (Nellen *et al.*, 1981), 6 for *Drosophila melanogaster* (Fyrberg *et al.*, 1981), 17 for *Dictyostelium discoideum* (McKeown and Firtel, 1981) and at least 6 for mouse and human genomes (Miwa *et al.*,1991). The number of actin genes also varies among echinoderms. We have identified five non-allelic actin genes in the sea urchin *L. pictus*. There are at least 8 non-allelic actin genes in the genome of the sea urchin *S. purpuratus* (Lee *et al.*, 1984). The number of actin genes is 5 in the sea stars *Pisaster ochraceus* and *P. brevispinus*, and 8 in *D. imbricata* (Kovesdi and Smith, 1985). The number of actin genes can thus vary between related species, indicating that gene duplication and/or deletion events occur on occasion. The cytoskeletal actin genes of *S. purpuratus* are clustered into two linkage groups (Lee *et al.*, 1984; Minor *et al.*, 1987); a similar arrangement appears to maintained in *S. drobachiensis* and *S. franciscanus* (Lee *et al*,

1984). No linkage of actin genes has been detected for *L. pictus* (Johnson *et al*, 1983; our results).

The nucleotide sequences of *L. pictus* actin genes obtained from this study allow us to perform sequence comparisons and gene phylogenetic analyses in order to examine evolutionary relationship among echinoderm actin genes. The 3' noncoding sequences characterized in this study for the LpC3, LpC4 and LpM actin genes, provide useful gene-specific hybridization probes for each of the *L. pictus* actin genes for an investigation of the differential regulation of actin gene expression in *L. pictus*.

CHAPTER III EVOLUTION OF SEA URCHIN ACTIN GENE FAMILIES

3.1 INTRODUCTION

Amino acid replacements accumulate in actin proteins at a very low rate, and there is a high degree of similarity of nucleotide coding sequence among actin genes (Hightower and Meagher, 1986). The conservation of the actin gene family provides an opportunity to investigate its molecular phylogeny. All chordates, echinoderms, and insects examined to date include at least one actin gene exclusively expressed in muscle and multiple cytoskeletal (nonmuscle) actin genes expressed in other types of cells. The vertebrate muscle and cytoskeletal actin genes can be distinguished on the basis of a set of diagnostic amino acid differences (Vandekerckhove and Weber, 1984). The muscle actin genes of chordates and echinoderms share some of the diagnostic amino acids which differentiate vertebrate muscle and cytoskeletal actins and apparently arose in a common deuterostome ancestor, probably from an echinoderm cytoskeletal actin gene (Kovilur *et al.*, 1993). The muscle actin genes of insects arose independently (Mounier *et al.*, 1992).

Differences in 3' noncoding sequences and organization of linkage groups have been used to classify the cytoskeletal actin genes of the sea urchin *S. purpuratus* into the CyI, CyII, and CyIII subfamilies (Lee *et al.*, 1984). The related species *S. drobachiensis* and *S. franciscanus* include CyI, linked CyIIa and CyIIb, and two CyIII actin genes (Lee *et al.*, 1984). This indicates that the actin gene subfamilies of *S. purpuratus* were present before these three species diverged.

We are interested in the extent to which mechanisms regulating the spatial specificity of gene expression are conserved among sea urchin species and among related

genes. In order to interpret the comparative observations on actin gene expression in *L. pictus* and *S. purpuratus* embryos and their reciprocal hybrids (see Chapter I), it is important to know whether, at the time *S. purpuratus* and *L. pictus* shared a common ancestor (30-40 mya; Smith 1988), there were the same subfamilies of actin genes having the same spatially restricted patterns of expression displayed by the contemporary actin genes of *S. purpuratus*. The purpose of the study presented in this chapter was to determine the relationship among sea urchin actin genes. We propose a model to account for the extant actin genes of these sea urchins in which the common ancestor already had two cytoskeletal actin genes which, after divergence of the species, duplicated to give rise to the CyI, CyII and CyIII subfamilies of *S. purpuratus* actin genes.

3.2 MATERIALS AND METHODS

All actin gene sequence data involved in this study, other than those of the *L. pictus* actin genes (LpC1-LpC4, and LpM), were obtained from the Genbank database release 77.0 (Benson *et al.* 1993). The accession numbers are, for sea star *P. ochraceus*, M26500 for PoM and M26501 for PoCy (Kowbel and Smith, 1989); for the sea urchin *S. purpuratus*, V01350 for CyI (Schuler *et al.*, 1983), V01349 for CyIIa (Schuler *et al.*, 1983), M35323 for CyIIb (Durica *et al.*, 1988), M29808 and M30511 for CyIIIa (Akhurst *et al.*, 1987), M35324 for CyIIIb (Durica *et al.*, 1988), X05739-X05744 for SpM (Crain *et al.*, 1987); for the sea urchin *S. franciscanus* X03075 for Sf15A and X03076 for Sf15B (Foran *et al.*, 1985); and L13787 for the ascidian *Styela clava* α-muscle actin gene ScM (Beach and Jeffery 1992).

Sequences were aligned and compared with the aid of the eyeball sequence editor, ESEE (Cabot and Beckenbach, 1989) and the PCGene software package (IntelliGenetics,

ver.6.5). Rates of synonymous and nonsynonymous nucleotide substitutions were computed with the LWL program (Li *et al.*, 1985).

Phylogenetic analysis was done by the neighbor-joining method (Saitou and Nei, 1987), using the NEIGHBOR program of PHYLIP (version 3.5; Felsenstein, 1989). Bootstrap analyses were performed by using the SEQBOOT, DNADIST, PROTDIST, NEIGHBOR, and CONSENSE functions of the PHYLIP package. The maximum parsimony algorithm was also used to generate phylogenetic trees using the DNAPARS and PROTPARS functions of PHYLIP.

3.3 RESULTS

3.3.1 Comparison of the actin entire coding sequences

Table 3.1 shows pairwise comparisons of amino acid and nucleotide sequences of two sea urchin *L. pictus* actin genes LpC1 and LpC2; five *S. purpuratus* actin genes CyI, CyIIb, CyIIIa/b and SpM; two *S. franciscanus* actin genes Sf15A and Sf15B; and two sea star *P. ochraceus* actin genes PoCy and PoM. The data in Table 3.1 confirm that cytoskeletal actin is extremely conserved within sea urchins, with more than 90% similarity at the nucleotide level, and more than 96% at the amino acid level. The nucleotide sequence divergence among cytoskeletal actin genes within a species ranges from 1.62% (CyI-CyIIb) to 7.78% (CyI-CyIIIa, CyIIb-CyIIIa) in *S. purpuratus*, is 4.60% between LpC1 and LpC2 of *L. pictus*, and is 1.71% between Sf15A and Sf15B of *S. franciscanus*. Although the nucleotide sequence divergences indicate that the two *L. pictus* cytoskeletal actin genes LpC1 and LpC2 are more closely related to CyI than to other *S. purpuratus* actin genes, pairwise comparison of amino acid sequences suggests that LpC1 is most similar to CyI (with 5 amino acid replacements), while LpC2 is most

Pairwise comparisons of cytoskeletal and muscle actin genes from the sea urchins *L. pictus*, *S. purpuratus*, *S. franciscanus* and the sea star *P. ochraceus*.

	S. p				S. f		L. p		P. o	muscle	
	CyI	CyIIb	CyIIIa	CyIIIb	Sf15a	Sf15b	LpC1	LpC2	PoCy	SpM	PoM
CyI		1.62/6.5	7.78/35.9	6.86/30.1	3.19/12.5	2.92/13.0	3.66/15.5	6.06/24.2	12.74/66.7	12.90/52.6	13.91/68.1
CyIIb	1		7.78/37.2	7.25/31.4	2.82/12.3	2.54/12.4	3.84/15.7	6.34/24.9	13.13/68.2	12.86/49.8	14.63/71.4
CyIIIa	9	8		3.10/10.7	6.97/31.0	7.37/34.5	8.48/39.1	9.50/44.6	16.31/93.4	15.65/67.7	15.38/76.6
CyIIIb	10	11	7		6.36/25.3	6.84/29.0	7.44/32.2	8.25/37.5	15.66/86.5	15.25/63.9	14.95/70.8
Sf15a	3	2	10	12		1.71/6.8	3.94/15.2	6.44/24.6	13.46/67.9	13.93/55.2	14.28/64.5
Sf15b	1	0	9	11	2		4.41/18.7	7.03/28.9	13.67/70.8	13.72/56.8	14.29/66.2
LpC1	5	6	13	13	7	6		4.60/16.2	12.61/62.9	13.72/53.4	14.08/65.3
LpC2	12	13	12	11	14	13	13		14.88/72.5	15.61/64.1	15.97/76.5
PoCy	7	8	13	13	11	8	10	17		16.39/75.8	11.37/52.7
SpM	26	27	30	29	29	27	31	33	24		12.78/56.7
PoM	14	15	17	17	17	15	19	21	13	16	

Percentage DNA sequence divergence using the Kimura two parameters correction (above) and synonymous site substitution rates (below) are given above the diagonal. Below the diagonal are the number of amino acid replacements. All amino acid sequences were translated from corresponding nucleotide sequences. There are 3 uncertain amino acids in each of the sequences of CyI, CyIIIa and SpM; since these amino acids are very conserved in all known actin amino acid sequences of sea urchins and sea stars, we did not include these sites as replacements.

Table 3.1

Pairwise comparisons of the third exon of actin genes of the sea urchins *L. pictus*, *S. purpuratus*, *S. franciscanus* and the sea star *P. ochraceus*

	S. p				S. f		L. p				P. o	muscle		
	CyI	CyIIb	CyIIIa	CyIIIb	sf15a	sf15b	LpC1	LpC2	LpC3	LpC4	PoCy	SpM	LpM	PoM
CyI		1.67	8.13	7.42	4.92	3.82	4.04	6.71	4.70	7.15	12.30	15.68	16.10	17.11
CyIIb	6.2/0.5		7.90	5.59	4.04	2.96	4.68	7.38	5.13	6.69	12.76	15.89	16.30	18.39
CyIIIa	33.7/2.7	34.3/2.1		5.59	7.19	6.73	8.59	10.97	10.26	10.96	17.75	19.45	20.42	18.27
CyIIIb	29.6/2.4	32.0/2.9	18.1/2.7		7.40	7.16	7.65	9.53	9.30	10.90	17.37	19.64	21.20	18.65
Sf15a	19.0/1.3	17.5/0.8	27.1/2.7	24.1/3.4		2.30	5.13	8.31	6.26	8.30	15.01	18.51	18.11	18.34
Sf15b	17.9/0.5	15.1/0.0	27.0/2.2	25.2/2.9	8.3/0.8		5.35	8.30	6.25	8.07	14.76	17.46	17.09	17.83
LpC1	20.6/0.3	21.1/0.8	39.6/2.4	32.6/2.1	21.1/1.3	24.9/0.8		5.13	2.52	5.78	14.74	19.64	18.13	18.63
LpC2	28.3/1.8	29.3/2.4	52.8/3.0	45.6/1.8	32.7/2.9	34.9/2.4	20.4/1.6		6.03	8.97	12.78	17.21	17.63	16.53
LpC3	19.0/1.3	18.6/1.8	43.1/3.5	36.7/3.2	22.6/2.4	23.8/2.1	8.3/1.0	20.5/2.6		7.14	15.03	18.79	19.47	18.64
LpC4	25.4/3.0	24.6/2.4	47.0/3.6	39.0/4.6	31.4/3.0	32.7/2.4	18.6/2.7	30.1/4.0	21.0/3.8		12.76	18.00	18.42	17.04
PoCy	58.6/3.7	56.9/4.1	104.6/5.9	90.0/5.5	71.7/4.7	73.7/4.1	66.2/3.4	66.4/5.1	56.6/4.4	71.9/5.0		17.97	18.36	14.32
SpM	61.1/7.5	53.6/8.3	73.7/9.7	74.9/9.2	69.2/9.2	68.1/8.3	69.4/8.0	89.1/8.3	63.6/9.2	69.1/9.5	61.1/9.8		8.22	16.21
LpM	72.0/5.9	66.0/6.8	93.9/8.1	105.6/7.6	77.8/7.6	75.0/6.8	81.1/6.5	81.2/6.6	74.6/7.9	86.7/7.8	71.4/8.5	37.5/1.9		16.84
PoM	85.9/6.2	89.0/6.9	83.7/7.3	83.0/7.2	80.8/7.7	82.0/6.9	76.7/6.7	96.8/7.0	68.2/7.7	82.0/8.1	61.1/5.8	69.2/6.4	74.9/6.5	

Percentage DNA sequence divergence using the Kimura two parameters correction are given above the diagonal. Below the diagonal are fractions (in percentage) of synonymous (above) and non-synonymous (below) substitutions.

Table 3.2

similar to CyIIIb (with 11 amino acid replacements).There are 13 amino acid replacements between LpC1 and LpC2 actins. The *S. purpuratus* muscle actin gene SpM and the two sea star actin genes PoCy and PoM are most divergent from sea urchin cytoskeletal actin genes. However, comparison of the two muscle actins to the cytoskeletal actins listed in Table 3.1 indicates that the sea star *P. ochraceus* muscle actin PoM is more similar to cytoskeletal actins than is the *S. purpuratus* muscle actin SpM.

3.3.2 Comparison of the nucleotide sequences of the third exons

Table 3.2 shows comparisons of the sequences of the third exon of all five *L. pictus* actin genes with other echinoderm actin genes; the third exons have the highest frequency of amino acid replacements among the exons. Above the diagonal are corrected pairwise nucleotide sequence divergences. Below the diagonal of Table 3.2 are rates of nucleotide substitutions for synonymous and non-synonymous sites. Pairwise comparisons shown in Table 3.2 indicate that the corrected nucleotide divergences range from 1.67% (CyI-CyIIb) to 10.96% (CyIIIa-LpC4) among cytoskeletal actin genes of sea urchins. The nucleotide divergence between the cytoskeletal actin gene of sea star *P. ochraceus* and the cytoskeletal actin genes of sea urchins range from 12.30% (PoCy-CyI) to 17.75% (PoCy-CyIIIa). Sea stars and sea urchins last shared a common ancestor about 500 mya (Sprinkle 1987, p570). The rates of nucleotide substitution at non-synonymous sites of sea urchin cytoskeletal actin genes are very low, ranging from 0.5% (CyI-CyIIb) to 4.6% (CyIIIb-LpC4).

3.3.3 Comparison of the amino acid replacements deduced from the third exons

The amino acid sequences of all the actin genes used in this study were deduced from the nucleotide sequences based on the universal code. Table 3.3 shows amino acid

replacements in the third exons when actins are compared to the amino acid sequence of the *S. purpuratus* CyI actin. Several informative features were noted: (1) Four cytoskeletal actin genes of *L. pictus* (LpC1-LpC4) and two *S. purpuratus* cytoskeletal actin genes (CyIIIa and CyIIIb) encode serine at codon 265. (2) The three muscle actin genes of the two sea urchins and the sea star encode the same distinctive amino acids at codon 260 (threonine), 267 (isoleucine), 278 (threonine), 287 (isoleucine), 303 (threonine), 305 (serine), 306 (tyrosine). Among these amino acid replacements those corresponding to codons 260 and 267 have been suggested as diagnostic for distinguishing muscle actins from cytoskeletal actins in deuterostomes (Collins and Elzinga, 1975, Kovilur et al., 1993). In addition to these features, four additional distinctive amino acids were shared by muscle actins of the two sea urchins *S. purpuratus* and *L. pictus*: alanine at codons 228 and 232 and serine at codons 323 and 324. (3) Like the muscle actins, CyIIIa, CyIIIb, and LpC2 cytoskeletal actins have two muscle-like "diagnostic" amino acids: threonine at the codon 260 and isoleucine at codon 267. Similar to the sequences of muscle actins, LpC4 and PoCy actins have a threonine at codon 278 and CyIIa has a tyrosine at codon 306. (4) A replacement of cysteine with alanine at codon 257 is found in the three *S. purpuratus* actins CyIIa, CyIIb, CyIIIa, the two *S. franciscanus* actins Sf15a and Sf15b, and the *L. pictus* actin LpC4.

Among the four cytoplasmic actins of *L. pictus*, the third exon of LpC1 actin is the most similar to *S. purpuratus* CyI actin, having only one amino acid difference at codon 265 (serine in place of alanine). All of the *L. pictus* cytoskeletal actins, as well as the *S. purpuratus* actins CyIIIa and CyIIIb, share this serine.

Among 12 amino acid differences between the actins LpC2 and CyI (Table 3.1), eight are found in the third exon. Four conservative substitutions were observed at codons 212 (valine for isoleucine), 262 (leucine for phenylalanine), 267 (isoleucine for

Comparison of Amino Acids of the Third Exon of Actin Genes

Actin Genes	Amino Acid Position Number																			
	211	212	219	222	225	227	228	230	231	232	250	251	257	258	260	262	265	267	272	278
CyI	D	I	V	D	Q	M	Q	A	A	S	I	G	C	P	A	F	A	L	A	C
CyIIa	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.
CyIIb	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.
CyIIIa	.	.	.	.	.	.	.	.	.	.	.	.	A	S	T	.	S	I	.	.
CyIIIb	.	.	.	.	E	.	.	.	.	.	.	.	.	S	T	L	S	I	.	.
Sf15a	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	P	.	.	.
Sf15b	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.
LpC1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.
LpC2	.	V	.	.	.	.	L	.	.	.	.	.	.	.	T	L	S	I	.	.
LpC3	N	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.
LpC4	.	.	.	.	.	.	.	.	.	.	T	.	A	.	S	Y	S	.	S	T
PoCy	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	T
LpM	N	.	.	.	.	.	A	P	.	A	.	.	.	.	T	.	.	I	.	T
SpM	.	.	T	N	R	.	A	.	.	A	.	.	.	.	T	.	.	I	.	T
PoM	.	.	T	.	.	.	.	.	S	.	.	.	.	.	T	.	.	I	.	T

Actin Genes	281	284	287	292	295	297	302	303	304	305	306	315	317	318	319	320	323	324	325	327	367
CyI	S	K	V	D	A	T	G	S	T	M	F	K	I	T	A	L	P	T	M	I	P
CyIIa	.	.	.	.	.	S	.	.	.	.	Y										
CyIIb	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
CyIIIa	R	.	.	.	V	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	.
CyIIIb	.	.	.	.	.	.	A	.	.	.	.	.	.	V	.	.	.	.	.	.	.
Sf15a	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Sf15b	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LpC1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
LpC2	.	.	.	.	.	.	.	.	.	.	.	R	.	S	.	.	.	.	.	.	.
LpC3	.	.	.	F	.	.	.	.	.	.	.	.	.	.	.	F	.	.	.	V	.
LpC4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.
PoCy	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.
LpM	.	.	I	.	.	.	.	T	S	.	Y	.	.	.	.	.	S	S	.	.	.
SpM	.	.	I	.	.	.	.	T	S	.	Y	.	.	.	.	.	S	S	V	.	A
PoM	.	.	I	.	.	.	.	T	S	.	Y	.	.	Q	.	.	.	.	.	.	.

Standard abbreviations of amino acids are used. Dots represent amino acids which are similar to those of CyI actin. Since the sequence of CyIIa is incomplete, amino acids corresponding to codons 307-375 were not compared and are shown as blanks in the table.

Table 3.3

leucine) and 318 (serine for threonine). Four additional replacements involved changes between nonpolar and polar amino acids: leucine for glutamine at codon 228, threonine for alanine at codon 260, serine for alanine at codon 265 and arginine for lysine at codon 315. A striking observation is that four identical amino acid replacements were also found in the sequences of CyIIIa actin at codons 260, 265, 267 and 318, and in CyIIIb actin at codons 260, 262, 265 and 267, making the amino acid composition of the third exon of LpC2 actin most similar to the CyIII actins among *S. purpuratus* actins, particularly CyIIIb.

The *L. pictus* cytoskeletal actin LpC4 has eight replacements of amino acids in the third exon compared to CyI actin. There are four changes of alanine to serine at codons 260, 265, 272 and 319. The replacement of alanine for serine at codon position 265 is a common feature for all cytoskeletal actin genes of *L. pictus* and for the CyIIIa and CyIIIb cytoplasmic actin genes of *S. purpuratus*. The serines at positions 272 and 319, and the replacements of theronine for isoleucine at codon 250 and tyrosine for phenylalanine at codon 262, are unique for the LpC4 actin among echinoderm actins characterized.

The third exon of the LpC3 gene encodes five amino acid differences compared to the CyI actin. Three changes are unique to this echinoid actin: phenylalanine for aspartic acid at codon 292, phenylalanine for leucine at codon 320 and valine for isoleucine at codon 327. The replacement of an aspartic acid with asparagine at the codon 211 also occurs in the *L. pictus* muscle actin.

Compared with the CyI actin, the muscle actin LpM has 13 amino acid replacements, SpM has 16, while PoM has 10. Muscle actin genes of echinoderms are thought to have evolved by duplication from a cytoskeletal-like ancestral actin gene (Kovilur *et al.*, 1993). The muscle actin PoM of the sea star has retained many features of

cytoskeletal actins, while the *S. purpuratus* muscle actin SpM is the most divergent from the cytoskeletal actins.

3.3.4 Comparison of the 3' noncoding regions of actin genes

The 3' noncoding sequences of the LpC1 and LpC2 actin genes were compared to the *S. purpuratus* CyI and CyIIa/b actin genes. Because of the incomplete sequence data available for the *S. purpuratus* actin genes, only 138 nt immediately following the termination codon were aligned (Fig. 3.1). Pairwise comparisons show that LpC1 and CyI are most similar (81.9%) while LpC2 and CyI are also quite similar (77.5%). LpC1 and LpC2 are 62.4% and 58.2% similar to CyIIa, respectively. The first 40-50 nt sequence proximal to the stop codon is more conserved than the remaining 3' noncoding regions of the LpC1 or LpC2 genes to the CyIIb, CyIIIa and CyIIIb actin genes show an average similarity only 53.4%, of marginal significance. As shown in Figure 2.5, the alignment of the entire 3' noncoding sequences of LpC1 and LpC2 reveals a 89% similarity. The alignment of the 3' noncoding regions of the muscle actin genes SpM and LpM is shown in Figure 3.2 and showed a 73.6% similarity. Both sequences are incomplete; 280 bps of sequence were compared.

3.3.5 Molecular phylogenetic analysis

A phylogenetic tree was predicted by applying the neighbor-joining distance matrix algorithm (Saitou and Nei, 1987) to the nucleotide sequences of third exons of echinoderm actin genes (Fig. 3.3, a). A muscle actin gene of the ascidian *S. clava* (ScM) was used as an outgroup. In this analysis all the cytoskeletal actin genes of *L. pictus* form a clade, while the cytoskeletal actin genes of *Strongylocentrotus* (including *S. purpuratus* and *S. franciscanus*) were all grouped in a different clade; this branching

Fig. 3.1 Comparison of the 3' noncoding regions of the *L. pictus* cytoskeletal actin genes LpC1 and LpC2 to the *S. purpuratus* cytoskeletal actin genes CyI, CyIIa and CyIIb. 138 nt immediately following the termination codon of each actin genes were aligned. Stars (*) indicate the nucleotides conserved in all sequences. Only differences are shown. Dots indicate identical nucleotides in comparison with the LpC1 sequence. Dashes are gaps introduced to optimize the alignment.

```
LpC1 3'   ACAACTCATTTT---GTTTTACTT-CTGTTGAGCA-CAATAACGAACTCC
LpC2 3'   .......G.C..TTC.........-..........-....A...------
CyI 3'    .......GC...T--...C...A.-..........-....A........A
CyIIa 3'  ....A..GC...T--...C...A.-..........-....A........A
CyIIb 3'  .......GC...T--.G.GA..AAA..C....A..T....A-------..A
          **** **   **   * *  **   ** **** ** *****

          TGGGGTGGGTTAAGTGGGTCACCTTTATCAAA-TGGGAAAATTCTGGAAC
          --........A.................T...A...-............
          ..........-..A......G....-.C.GTG-...CT..C.CT.T.G..
          ......--------....--GTT...--TGTGA--AC...C.C.C-...T
          -A..--------..AC.A-C..TGA.C-....-.A.T.C...G..T--A.GT
            **           *        *             ** *

          AGTTGGGTTGATCA----GGAATTATCTAGCTA------TC-CATCCC--
          ....A...-.....----.............-------..-G...T.--
          ...C....CT..T.----.A............------..-....T.--
          .-.CAAC..TGCT.T-AC.AG..G..TC--..T------.T-TGC.T.GT
          .--..AC---.C..TTGC..G-CA..GC...CGAAAAAC..G.GCTTTCT
          *            *    *     **    *         *

          CGATAA--T---------
          .......--.AGTC-CA--
          T.....--.A--------
          ......GG.GATTATAAC
          .A.-..CT.GG--AGGAC
            * **  *
```

Fig. 3.2 Alignment of the 3' noncoding termini of musle actin genes LpM and SpM. 280 nt of LpM and 277 nt of SpM immediately distal to the stop codon are aligned.

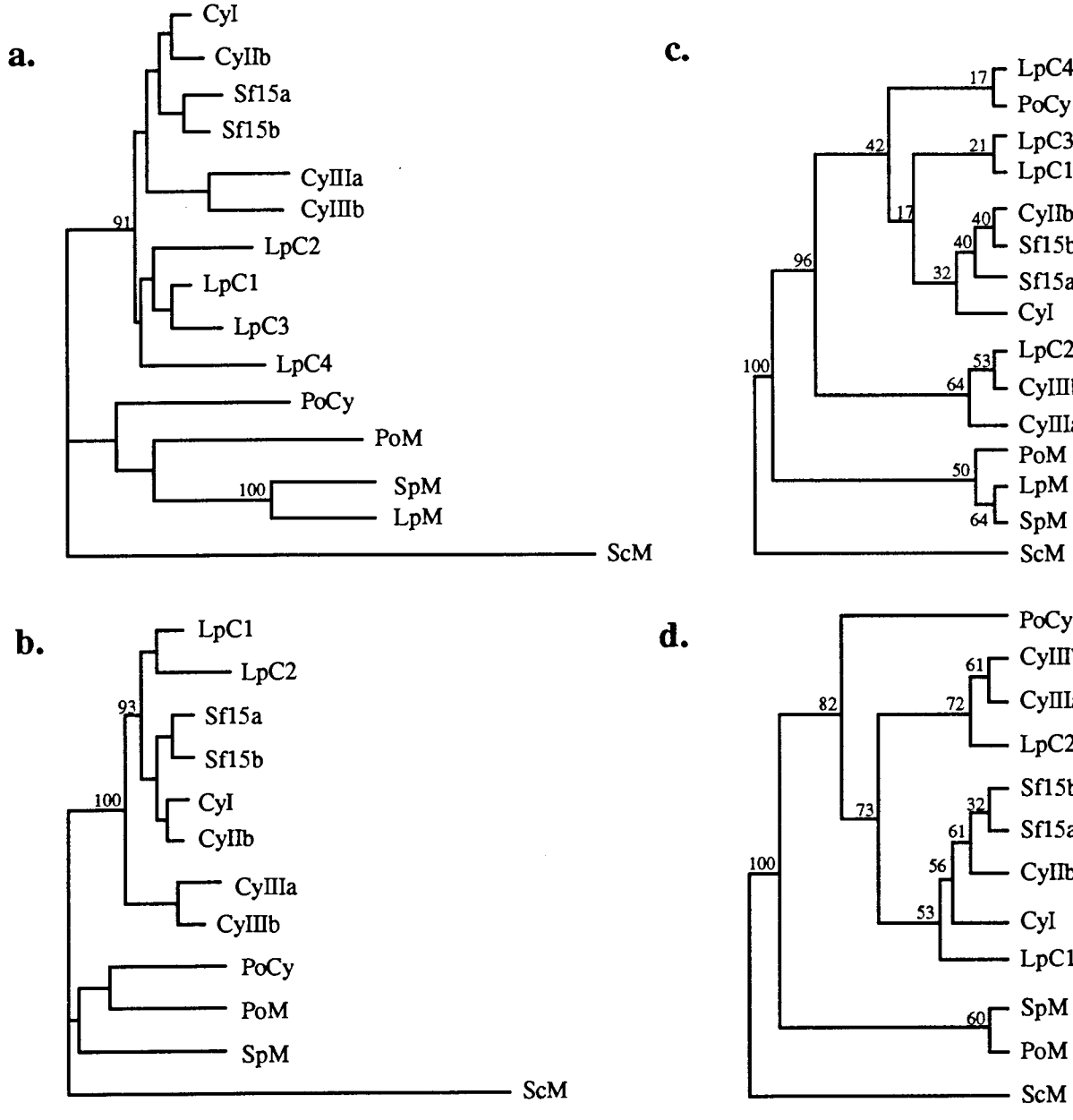
```
LpM     -   TACATCAACGGATAAGGGCTCACTGGTCTAGGAGGGCTGACATTGGCA -48
            :: :::::::::::::::::: ::: :   ::::::::::::::::::
SPM     - CTTATATCAACGGATAAGGGCTCGCTGATT--GGAGGGCTGACATTGGCA -48

        - GTTATTC-TTTGTGNAANCTGTAGTCGGAGCCCTGTAGGNTCTATATTAA -97
          ::: ::  :::::: :: ::::::::::::::::::  :  :::::::: :
        - GTTTTTTCTTTGTGCAATCTGTAGTCGGAGCCCTGATGATTCTATATTTA -98

        - TTTTAGAGTATCATTTATAGAATA-TTGTGACGTCACACTCTTTCTGTCA -146
          :::::::::::: : ::::::: :  :::::::::::  ::::: ::::::
        - TTTTAGAGTATTACTTATAGATTTATTGTGACGTCAACCTCTT-CTGTCA -147

        - TTCANCT-TCCAAGATNACGCCCACGACATGG---GTCGCCCCTCA-GGG -191
          ::    :  ::::: : :: :::   :::    :::  ::::   :::
        - AACATAATTGTAAGATCATGCACACTGTATGACGGGTCATCCCTAGAGGG -197

        - GT-AAGCATTGAGAGCGTCTAACT-AAGGTTGATTATGTTGAATGTCCGA -239
          :  : :: ::::::::::::::     :   : ::    :: :::   :
        - GGTATGCCTTGAGAGCGTCTAACGTCCCG--CAATACCAAGATTGTT--A -243

        - TTTTCTTACATGATGGATCCTCTAAA-TGATGGAAAGTCAAC -280
          : :: :   :::   ::: :: :::  : :::::   ::  :
        - TGTTGT---ATGTCCGATTCTTTAACATCATGGA---TC--C -277
```

pattern was supported by 91% of bootstrap replicates. The same set of nucleotide sequence data was also analyzed by the method of maximum parsimony and resulted in the same branching pattern: the formation of separate *Lytechinus* and *Strongylocentrotus* clades was supported by 73% of bootstrap replicates (data not shown).

Figure 3.3(c) shows the bootstrapped tree derived from the application of the maximum parsimony algorithm to the deduced amino acid amino acid sequences. In this tree the cytoskeletal actin genes of each species did not group together on a branch. The most striking differences from the tree shown in the Figure 3.3(a) were: (1) the LpC1 and LpC3 actins were on a branch supporting the *S. purpuratus* CyI and CyIIb actins and *S. franciscanus* Sf15a and Sf15b actins. (2) CyIIIb actin was more closely related to LpC2 actin than to CyIIIa actin, although the two CyIII actins are encoded by genes which are closely linked and share a very similar 3' noncoding terminus in *S. purpuratus* (Lee et al., 1984); this terminus is not shared with any *L. pictus* actin gene. The formation of a clade including LpC2 and CyIIIa,b actins separated from other sea urchin cytoskeletal actins was supported by 96% of bootstrap replicates (Fig. 3.3, c).

Fig. 3.3(b and d) shows the trees generated by comparison of the nucleotide and amino acid sequences of the entire coding sequences of the actin genes, respectively; the lack of complete sequences for the *L. pictus* LpC3, LpC4, and LpM genes prevented their inclusion. The neighbor-joining tree for the nucleotide sequences (Fig. 3.3, b) placed the *L. pictus* genes LpC1 and LpC2 on a separate branch (supported by 93% of bootstraps) which is more closely related to the CyI-CyII group than to the CyIII family of actin genes. The maximum parsimony tree for deduced amino acid sequences (Fig. 3.3, d) grouped LpC2 with the CyIII actins, and placed LpC1 actin in a clade including of the CyI, CyII, and *S. franciscanus* 15a and 15b actins. The separation of these two clades is supported by 73% of bootstraps.

Fig. 3.3 Molecular phylogenetic analysis of the actin genes of echinoderms. Trees shown in panels a-d were constructed by applying the neighbor-joining distance matrix algorithm to (a) the nucleotide sequence of the third exon, (b) the nucleotide sequence of the entire coding region, and by applying the maximum parsimony algorithm to (c) the deduced amino acid sequence of the third exon,and (d) the deduced amino acid sequence of the entire coding region. The branch lengths shown in panels (a) and (b) are proportional to distance values. In all trees, an ascidian muscle actin gene (ScM) was included as an outgroup. Numbers shown are percentages of 100 bootstrap replicates in which the same internal branch is recovered.


a.
CyI
CyIIb
Sf15a
Sf15b
CyIIIa
CyIIIb
LpC2
LpC1
LpC3
LpC4
PoCy
PoM
SpM
LpM
ScM
91
100
b.
LpC1
LpC2
Sf15a
Sf15b
CyI
CyIIb
CyIIIa
CyIIIb
PoCy
PoM
SpM
ScM
93
100
c.
LpC4
PoCy
LpC3
LpC1
CyIIb
Sf15b
Sf15a
CyI
LpC2
CyIIIb
CyIIIa
PoM
LpM
SpM
ScM
17
42
21
17
40
40
32
96
100
53
64
50
64
d.
PoCy
CyIIIb
CyIIIa
LpC2
Sf15b
Sf15a
CyIIb
CyI
LpC1
SpM
PoM
ScM
61
82
72
73
32
61
56
53
100
60

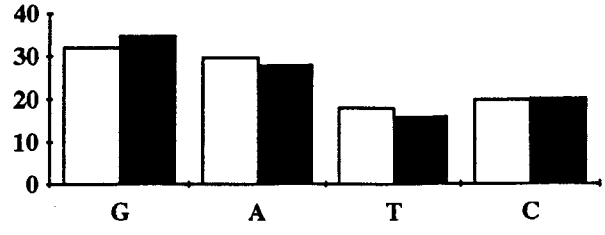
3.4 DISCUSSION

3.4.1 Sequence bias and codon usage in echinoderm genes

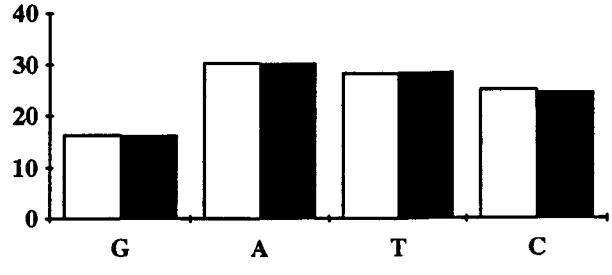
A strong bias for the use of particular synonymous codons has been noted for sea urchin actin genes by Foran *et al.* (1985) and for sea star actin genes by Kovesdi and Smith (1989). Figure 3.4 summarizes the average base compositions of sea urchin actin genes at the first, second, and third positions of each codon. The third positions are highly biased for C. Britten (1993) compared the sequences of actin and other genes from diverse organisms and concluded that highly conserved genes having extreme G+C base compositions at the third sites of codons are constrained in the extent of synonymous base substitutions at these sites. He also found that some synonymous substitutions are never observed, apparently forbidden by selective mechanisms. We have used an empirical equation of Britten (1993) relating the G+C third base compositions of pairwise gene comparisons to predict the fraction of third site substitutions upon saturation. Results of these calculations (data not shown), as well as silent site substitution rates presented in Table 3.1 and 3.2 which have been corrected by the method of Li *et al* (1985), suggested that silent site substitutions have not reached saturation for the sea urchin cytoskeletal actin genes. Therefore, comparison of the nucleotide coding sequences of these echinoid actin genes was expected to provide useful information concerning the molecular phylogeny of these genes, in spite of the long time since the divergence of *L. pictus* and *S. purpuratus*.

Fig. 3.4 Average base composition of sea urchin *S. purpuratus* and *L. pictus* actin genes at first, second, and third codon sites. Open bars, *L. pictus*; closed bars, *S. purpuratus*.

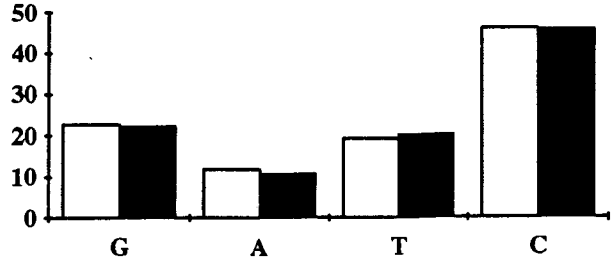
First Codon



Second Codon



Third Codon



3.4.2 Evolution of cytoskeletal actin gene families in sea urchins

Two questions were of most interest to us regarding the cytoskeletal actin gene families of *S. purpuratus* and *L. pictus*. First, how many extant cytoskeletal actin genes were present in the common ancestor at the time of the divergence? Second, which *L. pictus* actin genes are orthologues of *S. purpuratus* actin genes?

Based on sequence comparisons and patterns of expression in embryos, the *L. pictus* actin genes LpC1, LpC2, LpC3 and LpC4 are cytoskeletal, while LpM is a muscle actin gene; there is no evidence that any of them are closely linked. The five expressed cytoskeletal actin genes of *S. purpuratus* have been categorized into three subfamilies (CyI, CyII and CyIII) primarily based on similarities of the 3' noncoding sequences and linkage (Lee *et al.*, 1984). CyI, CyIIa, and CyIIb actin genes are linked and share sequence similarities in their first intron; the 3' noncoding sequences of the CyII genes are very similar to one another and somewhat similar to CyI (Scheller *et al.*, 1981). The CyIIIa and CyIIIb actin genes of *S. purpuratus* are linked and have similar 3' noncoding sequences and similarities in deduced amino acids which distinguish them from other *S. purpuratus* actin genes (Lee *et al.*, 1984; Akhurst *et al.*, 1987; Flytzanis *et al.*, 1989). The cytoskeletal actin genes of *L. pictus* do not easily fit such a categorization. For instance, the third exons of the genes LpC1 and LpC3 are more similar to each other than to other *L. pictus* actin genes, but their 3' noncoding regions are quite different. On the other hand, the 3' noncoding regions of LpC1 and LpC2 are highly (89%) similar, but the two genes encode actins differing by 13 amino acid replacements, more than the difference between several *L. pictus* and *S. purpuratus* actin genes (Table 3.1). The 3' noncoding sequences of the LpC1 and LpC2 actin genes are also very similar to the 3' noncoding sequence of the *S. purpuratus* CyI gene; this is the only 3' terminal sequence of a *S. purpuratus* cytoskeletal actin gene which cross hybridizes with *L. pictus* DNA (Lee *et al.*,

1984). We have confirmed that this is the only 3' noncoding sequence of *S. purpuratus* actin genes extensively shared with *L. pictus* actin genes. Comparison of the 3' noncoding sequences as used to classify *S. purpuratus* actin genes cannot be effectively applied to classify the *L. pictus* actin genes and establish their relationship with actin genes of *S. purpuratus*.

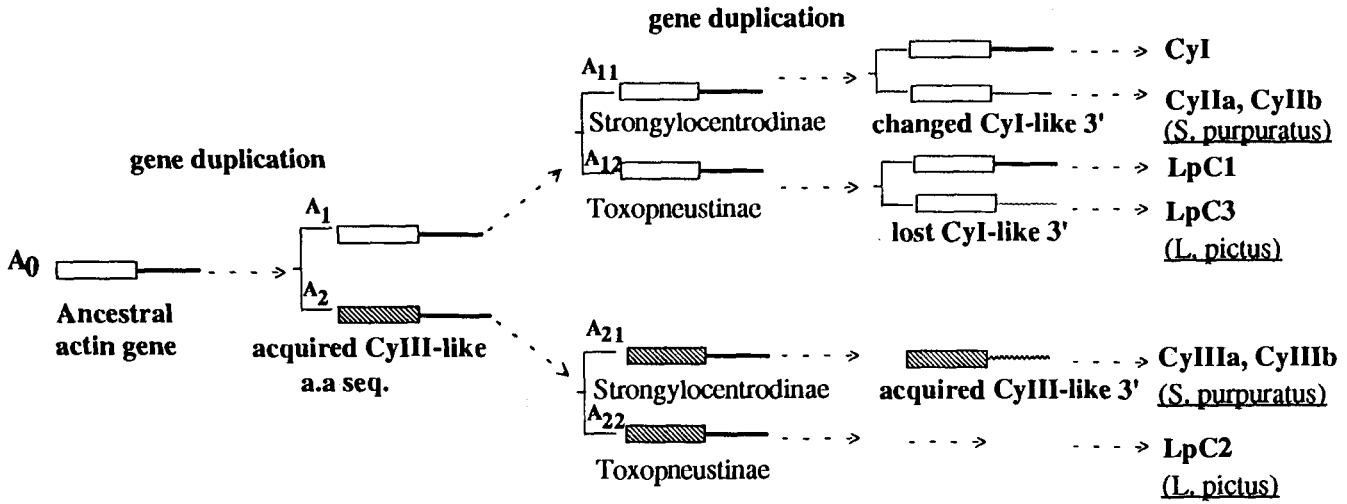
Molecular phylogenetic analysis is a useful tool for study of gene evolution. We have investigated the phylogeny of sea urchin actin genes by application of the neighbor-joining distance matrix and the maximum parsimony algorithms to the nucleotide and amino acid sequences of sea urchin actin genes. The tree built by using the nucleotide sequence data of the third exons showed a separation of all the cytoskeletal actin genes of *L. pictus* from the cytoskeletal actin genes of *Strongylocentrotus* (Fig. 3.3,a). The tree derived from the complete nucleotide coding sequences did not separate the cytoskeletal actin genes of *L. pictus* and *Strongylocentrotus* exclusively, since LpC1 and LpC2 were placed on a branch which is more closely related to the CyI-CyII group and together they are all separated from the CyIII genes (Fig. 3.3,b). Differences in trees generated from the entire coding sequence or the third exon suggest that different domains of the actin genes have evolved differently. In both trees, the *L. pictus* actin genes were more similar in their nucleotide sequence than they are to any *S. purpuratus* actin genes. The simplest interpretation of the tree in Figure 3.3 (a) is that all extant cytoskeletal actin genes of *L. pictus* and *S. purpuratus* arose after the divergence of the two species and were most likely derived from a common ancestral actin gene. In this case, none of the cytoskeletal actin genes of *L. pictus* is orthologous to any of the cytoskeletal actin genes of *S. purpuratus*, and any similarity of expression pattern between cytoskeletal actin genes of the two species would be coincidental or convergent. However, the apparent close relationships observed in the nucleotide sequences of the cytoskeletal actin gene within each sea urchin species might be a consequence of some intraspecific homogenization of

nucleotide sequences. Constraints on allowable nucleotide substitutions within a species and/or gene conversion events may have obscured the phylogenetic relationships. Studies on globin gene families suggest that concerted evolution events may partly or completely erase the evolutionary history of divergence between duplicated genes (Li and Graur 1991).

In contrast to the above conclusions, however, are the trees constructed by the method of maximum parsimony based on the deduced amino acid sequence of the third exon (Fig. 3.3, c) and the entire actin coding region (Fig. 3.3, d). In these trees, the grouping of LpC1 with CyI-CyII and LpC2 with CyIII were supported by 61% (Fig. 3.3, c) and 63% (Fig. 3.3, d) of bootstraps, respectively. Comparison of amino acid differences may be most informative by focusing on rare events which are shared. As described in RESULTS, comparison of the third exons supports a close relationship of LpC1 and CyI actins, which differ by a single nonconservative amino acid substitution at codon 265 (one which is common to all *L. pictus* cytoskeletal actin genes). The LpC2 and CyIIIb genes share an unusual complex of amino acid substitutions at codons 260, 262, 265, and 267. Based on the amino acid sequence of the third exon, LpC2 actin indeed appears to be closely related to CyIIIb, and, to a lesser extent, CyIIIa actins. The LpC2 actin gene has a pattern of expression in embryos similar to that of the CyIIIa and CyIIIb genes, while the LpC1 gene has a pattern of expression most similar to that of the CyIIa gene, as well as the CyI and CyIIb genes (see Chapter IV). These similarities in amino acid sequences of the third exon and spatial patterns of expression are consistent with LpC2 being related to the CyIII subfamily, and LpC1 being related to the CyI/CyII subfamily of actin genes. These relationships were also observed for the entire amino acid sequence (Fig. 3.3, d).

In Figure 3.5 we propose a model, based on similarities of the amino acid sequences (especially those of the third exon) and 3' noncoding sequences, which can account for the origin of the extant cytoskeletal actin genes of *L. pictus* and *S. purpuratus*. We propose that the ancestral cytoskeletal actin gene had a CyI-like third exon and 3' noncoding sequences, since these have been conservatively retained in several echinoid species (Lee *et al.*, 1984); we refer to this gene as A_0. A_0 then duplicated, and one copy (A_1) retained CyI-like amino acid sequences while the other copy (A_2) diverged to become CyIII-like in amino acid sequence. The conserved complex of amino acid substitutions characterizing the CyIII third exon may be the result of selection for regained functionality after initial mutational events resulted in impaired function, and/or the result of selection for specialized function. A_2 acquired a CyIII-like 3' noncoding terminus as A_{21} in an ancestor of *S. purpuratus* after the divergence of the Strongylocentrodinae and Toxopneustinae lineages; this may have been the result of gene conversion or insertion events (although there is no known source of the A_{21} terminus). This CyIII-like A_{21} gene then gave rise by duplication to the extant CyIIIa and CyIIIb genes of *S. purpuratus* (as well as the CyIIIc pseudogene, Lee *et al.*, 1984). In the Toxopneustinae lineage, the A_{22} gene derived from A_2 retained its CyI-like terminus and evolved into LpC2. Alternatively, the A_2 gene acquired its CyIII-like 3' terminus before the Strongylocentrodinae and Toxopneustinae lineages diverged. After divergence, A_2 gave rise to A_{21} in the Strongylocentrodinae lineage, which duplicated to form CyIIIa and CyIIIb actin genes. In the Toxopneustinae lineage A_2 evolved into A_{22}, which subsequently reacquired a CyI-like terminus, probably by gene conversion, and then evolved into LpC2. We favor the first alternative as shown in Figure 3.5 because it is more parsimonious and if CyIIIa termini appeared before the divergence of these species they might be expected to be found in other sea urchin species, but no CyIII-like termini have been reported outside of the Strongylocentrodinae. In our model, the CyI-like A_1 gene evolved into the A_{11} gene in the Strongylocentrodinae lineage, which duplicated

Fig. 3.5. Model for the evolution of the actin gene families of the sea urchins *L. pictus* and *S. purpuratus*. A pathway is presented for the generation of extant CyI-like and CyIII-like genes in contemporary sea urchins from that of a putative ancestral actin gene (A_0). See Discussion for explanation.


gene duplication
gene duplication
A0
Ancestral actin gene
A1
A2
acquired CyIII-like a.a seq.
A11
Strongylocentrodinae
A12
Toxopneustinae
changed CyI-like 3'
lost CyI-like 3'
CyI
CyIIa, CyIIb
(S. purpuratus)
LpC1
LpC3
(L. pictus)
A21
Strongylocentrodinae
A22
Toxopneustinae
acquired CyIII-like 3'
CyIIIa, CyIIIb
(S. purpuratus)
LpC2
(L. pictus)
: CyI-like a.a coding region;
: CyIII-like a.a coding region;
: CyI-like 3' end;
: CyIII-like 3' end.

giving rise to the CyI and CyII genes; the latter duplicated again, after changing much of its 3' terminus, generating CyIIa and CyIIb (as well as the CyIIc psuedogene). In the Toxopneustinae lineage the A_1 gene evolved into A_{12}, which duplicated to form LpC1 and LpC3. The latter subsequently replaced its 3' terminus with one unrelated to CyI, as the CyII genes have done to a lesser extent. The origin of the LpC4 gene is not accounted for by our model. It may either have arisen by duplication after the divergence of the species, or may be derived from a primitive cytoskeletal actin gene now absent in *S. purpuratus*. The latter explanation is consistent with the similarity of LpC4 actin to echinoderm muscle actins and the sea star cytoskeletal actin. The model presented is probably an over simplification, but accounts for the most salient comparative sequence observations.

In summary, the molecular phylogeny of the *Lytechinus* and *Strongylocentrotus* cytoskeletal actin genes cannot be established with certainty. More information about intron placement and sequences and 5' flanking sequences might be useful. The characteristics of genes similar to CyIII/LpC2 in other sea urchins might also be informative. At this time we favor a model in which there were CyI-like and CyIII-like actin genes at the time ancestors of these two species diverged, with more recent duplications generating the remaining extant cytoskeletal actin genes.

3.4.3 The echinoderm muscle actin genes

In the two trees shown in Figure 3.3(c, d) muscle genes of echinoids form a clade separated from all the cytoskeletal actin genes (supported by 100% of bootstraps). We conclude that the ancestral muscle actin gene existed in the common ancestor of echinoderms, and the muscle actin genes of sea urchins *L. pictus* and *S. purpuratus* are orthologous. The sea urchin muscle actin genes have some common features in their

amino acid sequences which are not found in the sea star muscle actin gene, and other features which are shared (see RESULTS and Table 3.3). Compared with the sea urchin muscle actins, the sea star muscle actin shares more amino acids with cytoskeletal actins, probably resembling more closely the cytoskeletal actin gene which served as the common ancestor of echinoderm muscle actin genes (Kovilur *et al.*, 1993). Thus, there appears to be more constraint on the amino acid sequence of muscle actin in sea stars than in urchins. Alternatively, gene conversion between muscle and cytoskeletal actin genes may have resulted in the sequence similarity of the actin genes in asteroids.

Kovilur *et al.* (1993) have provided molecular phylogenetic evidence that a common ancestral actin gene duplicated in an early echinoderm, giving rise to the families of muscle and nonmuscle genes of deuterostomes including vertebrates. The muscle actins of arthropods are distinct from deuterostome muscle actins, and muscle actins have not been identified in molluscs or psuedocoelomates (Mounier *et al* 1992). Thus the muscle actin genes of deuterstomes arose after their divergence from protostomes, apparently by duplication from a cytoskeletal actin gene, while the insect muscle actin genes arose by an independent gene duplication in an early arthropod. After arising by duplication, the vertebrate and insect muscle actin genes evolved rapidly relative to the cytoskeletal actin genes and have sometimes been duplicated (Mounier *et al.* 1992). The cytoskeletal actin genes have also continued to duplicate and diverge during deuterstome evolution. The most rapid rate of divergence of actin genes may occur just after a duplication in which one gene is released from the conservative functional constraint imposed on actins until it acquires an unique function or exclusive pattern of expression.

3.4.4 Constraints on the 3' noncoding sequences of the actin genes

The 3' noncoding sequences of the LpC1 and LpC2 actin genes are very similar. The many small gaps required for alignment indicate that these sequences did not arise recently; thus the 3' noncoding sequences are likely to be selectively constrained. This possibility is supported by the observation that this conserved 3' noncoding sequence is retained by one or more actin genes of all species of sea urchins examined to date (Lee *et al.*, 1984). Slipped-strand mispairing (Levinson and Gutman, 1987; Li and Graur, 1991) has been invoked to explain the insertion/deletion events which occur in DNA regions containing contiguous short repeats, such as the intron and 3' noncoding regions of the linked actin genes of *S. franciscanus* 15A and 15B (Foran *et al.*, 1985), as well as *S. purpuratus* actin genes CyI and CyIIa (Schuler *et al.*, 1983) and CyIIIa and CyIIIb (Durica *et al.*, 1988). The same mechanism may explain the insertion/deletion events in the A-T rich 3' noncoding regions of LpC1 and LpC2 genes. Vertebrate actin genes also have highly conserved, isoform-specific 3' noncoding sequence elements (Ng *et al.*, 1985; Yaffe *et al.*, 1985; Erba *et al.*, 1986, 1988). Several functions have been proposed for the highly conserved noncoding regions sometime observed among members of multigene families, including roles in the stability, localization, or activity of the mRNAs. The nucleic acid sequence responsible for the peripheral localization of the β-actin mRNA resides in the 3' noncoding region (Lloyd and Gunning, 1993). An influence of noncoding regions of the γ-actin mRNA on myoblast morphology has been reported (Lloyd and Gunning, 1993). Based on results of genetic complementation experiments on mammalian muscle cells, Rastinejad and Blau (1993) concluded that expression of certain differentiation-specific 3' noncoding regions of muscle mRNAs inhibits proliferation and promotes differentiation (Rastinejad and Blau, 1993).

While the function of the conserved 3' noncoding sequence of the LpC1 and LpC2 actin genes is unknown, it is unlikely to be related to the spatial patterns of expression of these genes in the embryos, since these are quite distinct for the two genes (see Chapter IV).

CHAPTER IV DIFFERENTIAL EXPRESSION OF THE ACTIN GENE FAMILY OF THE SEA URCHIN *LYTECHINUS PICTUS*

4.1 INTRODUCTION

The temporal (Crain 1981; Lee *et al.* 1986) and spatial (Cox *et al.* 1986) patterns of expression of the sea urchin *S. purpuratus* actin genes indicate they are differentially regulated during embryogenesis. Regulation of *S. purpuratus* actin gene expression has been investigated by analysis of the expression of actin promoter-reporter fusion genes in transgenic embryos (Flytzanis *et al.*, 1987, Hough-Evans *et al.*, 1987, Katula *et al.*, 1987, Franks *et al.*, 1988, Collura and Katula *et al.*, 1993), or by examining patterns of gene expression in hybrid embryos generated between *S. purpuratus* and *L. pictus* (Bullock *et al.*, 1988, Nisson *et al.*, 1989). Due to the lack of hybridization probes and sequence information about the actin genes of L. pictus, these investigations were limited to actin genes of *S. purpuratus*.

Sequence comparisons (Chapter III) indicated the orthologous relationship between the muscle actin genes SpM of *S. purpuratus* and LpM of *L. pictus*. Analyses based on the amino acid sequences of the third protein-coding exon suggested close relationships between the LpC1 gene and the CyI gene, and the LpC2 gene and the CyIII genes.

In this chapter, I describe the temporal and spatial patterns of expression of the *L. pictus* actin gene family. All five members of the gene family are expressed at some stages of embryogenesis.

4.2 MATERIALS AND METHODS

4.2.1 Probes for hybridizations

Probes were either labeled with ^{32}P-dCTP by random priming method for RNA gel blot hybridization experiments, or with digoxygenin-UTP (DIG-UTP) by *in vitro* transcription for whole mount *in situ* hybridization experiments. The strategies for subcloning probes specific to each individual gene were described in Chapter II.

4.2.2 Preparation of cytoplasmic RNA

Total cytoplasmic RNA was prepared by LiCl precipitation from a 4 M urea solution and phenol/chloroform extraction as described by Conlon *et al* (1987). About 10,000-20,000 embryos were collected at desired stages and washed with cold CMFSW once. The embryo pellet was resuspended in 1 ml of cold lysis buffer (50 mM Pipes, pH 6.5; 400 mM NH_4Cl; 12 mM $MgCl_2$; 25 mM EGTA) and transferred to a 1.5 ml microfuge tube. The embryos were collected again by centrifugation in a microfuge (Fisher centrifuge, Model 59) at 1,000g for 20 seconds. The embryos were lysed in 200 µl lysis buffer containing 0.5% Triton X-100 by vigorous aspiration with a P-1000 Pipetman until the homogenized solution turned clear and no more intact embryos could be seen. Nuclei and cell debris were removed by centrifugation in a microfuge at 14,000 rpm (Hermle Z302 K, BHG) for 2 minutes at 0°C. RNA was precipitated from the supernatant with an equal volume of prechilled LiCl-urea solution (4 M LiCl; 8 M urea; 0.5 mM EDTA; 20 mM Tris, pH 7.5) for at least 3 hours at -20°C, collected by centrifugation in a Hermle microfuge (14,000 rpm) for 5 minutes at 0°C, and dissolved in RNA extraction solution (0.1 M NaAc, pH 5; 0.5% SDS; 25 mM EGTA) followed by

extractions once with phenol/chloroform and once with chloroform. RNA was precipitated with 2.5 volume of prechilled ethanol for at least 1 hour at -20°C; The RNA pellet was collected by centrifugation in a Hermle microfuge for 10 minutes at 0°C and dissolved in approximate 30-50 µl of dH_2O. The RNA concentration was determined by the value of absorbence at 260 nm measured in a spectrophotometer (Pharmacia LKB, Ultrospec III). All aqueous solutions used were treated with 0.1% DEP (diethylpyrocarbonate, Sigma) overnight and autoclaved except the LiCl-urea solution which was prepared in autoclaved DEP-treated H_2O.

4.2.3 Preparation of RNA blots

RNA samples, 5 or 10 µg dissolved in dH_2O, were dried by vacuum desiccation (SpeedVac SVC 100), redissolved in 9 µl of 1 × MOPS (3-[N-morpholino]propanesul fonic acid, Sigma) buffer (20 mM MOPS, pH 7; 5 mM NaAc, 1 mM EDTA) containing 50% deionized formamide (BRL, ultrapure) and 6% deionized formaldehyde (Sigma), and denatured by heating at 65°C for 5 minutes and quickly chilled on ice. 1 µl of 10 × RNA loading buffer (Maniatis et al, 1990) containing 1 mg/ml of ethidium bromide was added to each sample. RNA samples were separated by electrophoresis on a 1.2% agarose gel containing 1× MOPS buffer and 3.3% formaldehyde; electrophoresis was carried out in 1 × MOPS with circulation, using a voltage of 4-5 V/cm gel. After electrophoresis, the gel was washed twice with distilled water for 10 minutes each, and blotted onto GeneScreen filter (New England Nuclear) in 10 × SSC overnight. RNA was UV cross-linked to the filter by using a Stratalinker (Stratagene) and baked for 2 hours under vacuum.

4.2.4 Labelling of DNA probes

DNA probes were labeled by random priming reaction using the NEBlot Kit (New England BioLabs), according to the instruction manual. Briefly, 25-50 ng of template DNA was dissolved in 33 µl of nuclease free H_2O, denatured in a boiling water bath for 5 minutes, then quickly placed in ice for 5 minutes. The sample was centrifuged briefly, and combined with 5 µl of 10 × labeling buffer (including random octadeoxyribonucleotides), 6 µl of dNTP mixture, 5 µl α-^{32}P dCTP (3,000 Ci/mmol, 50 µCi) and 1 µl DNA polymerase-Klenow fragment (5 units). The reaction was carried out at 37°C for 1 hour. Labeled DNA product was separated from unincorporated nucleotides by chromatography on a Sephadex G-50 column (Pharmacia) equilibrated with distilled H_2O. The labeling buffer, dNTP mixture and DNA polymerase were provided with the Kit.

4.2.5 Preparation of ribo-probes

In vitro synthesis and labeling of a RNA strand from the recombinant Bluescript plasmid was carried out at 37°C for 1-2 hours in a total volume of 20 µl containing transcription buffer (40 mM Tris-HCl, pH 7.5; 6 mM $MgCl_2$; 2 mM spermidine; 10 mM NaCl), 1 µg DNA template linearized with appropriate restriction enzyme, 2.5 mM each of ATP, GTP, UTP, 12 µM of CTP, 50 µCi (800 Ci/mmol or 3,000Ci/mmol) α-^{32}P-CTP (Amersham) and 5-10 units of appropriate RNA polymerase (T_3 or T_7 RNA polymerase). After performing the in vitro transcription reaction, DNA template was removed by adding RQ1 RNase-free DNase in a concentration of 1 unit/µg DNA and incubating at 37° C for 15 minutes. Unincorporated nucleotides were removed from labeled RNA by chromatography on Sephadex G-50. All solutions involved were prepared with DEP treated and autoclaved distilled water.

4.2.6 Northern blot hybridizations

RNA blot hybridizations with labeled DNA probes were performed in a solution consisting of 6 × SSPE (20× SSPE: 3 M NaCl, 0.2 M NaH_2PO_4, 0.02 M EDTA-Na_2, pH 7.4), 5 × Denhardt's solution (0.1% BSA, 0.1% Ficoll, 0.1% polyvinyl-pyrrolidone), 0.3% SDS and 100 µg/ml denatured salmon sperm DNA. Hybridizations were carried out overnight at the temperatures specified in the text. RNA blot hybridizations with labeled RNA probes were performed in a solution of 50% deionized formamide, 5 × SSPE, 5 × Denhardt's solution, 1% SDS and 100 µg/ml of denatured salmon sperm DNA. Hybridizations were carried out overnight at 65°C or as specified in text. Post-hybridization washes were usually carried out in 2 × SSPE, 0.5% SDS, twice for 30 minutes, at the hybridization temperature. Washes at higher stringency were sometimes performed.

4.2.7 Whole mount *in situ* hybridizations

Whole mount *in situ* hybridizations using digoxygenin-UTP (DIG-UTP) labeled RNA probes were carried out with modifications of the method, developed for *Drosophila* and *Xenopus* embryos (Tautz and Pfeifle, 1989, and Hemmati-Brivanlou *et al.*, 1990, respectively), and adapted by Harkey and Whiteley (1992) and Lepage *et al* (1992) for sea urchin embryos. I have introduced other modifications as described below.

Fixation and Storage of Embryos

Sea urchin embryos were fixed at desired stages in 10 volumes of cold fixative solution (2.5% glutaraldehyde; 0.14 M NaCl, 0.2 M phosphate buffer, pH 7.4) for 2 hours on ice. Embryos were washed twice with 10 volumes wash solution (0.3 M NaCl; 0.2 M phosphate buffer, pH 7.4) at room temperature, and then dehydrated with washes of 10%,

35%, 50% and 70% ethanol. Embryos may be stored indefinitely in 70% ethanol at -20° C.

Digestion and Post-fixation

Approximately 50-60 µl fixed embryos were dehydrated with 5 min washes in 1 ml 80%, 90% and 2 × 100% ethanol, and then suspended in 100% toluene for 10 - 15 min (Note: This step de-lipidates embryos, increasing probe accessibility). After the toluene treatment, embryos were washed once with 100% ethanol and rehydrated with 5 min/ml washes in 90%, 80%, 70%, 50%, 35% and 10% ethanol, and then twice with PBST (0.15 M NaCl, 0.2 M phosphate buffer, pH 7.4, 0.1% Tween-20). Embryos were digested in 300 µl of Proteinase K/PBST solution for 5 - 10 minutes. Concentrations of Proteinase K utilized are species specific: for *L. pictus*, approximately 20 µg/ml for prism and pluteus embryos and 30-40 µg/ml for blastula and gastrula embryos; for *S. purpuratus* embryos, 5 µg/ml was generally applied. The digestion was stopped with 1 ml stop-solution (2 mg/ml glycine in PBST). Embryos were washed once (5 min/ 1 ml) in PBST, then post-fixed for 30 minutes with occasional gentle mixing in 4% freshly prepared formaldehyde in PBST (dissolve 40 mg of para-formaldehyde in 250 µl DEP-dH_2O at 60°C with addition of 2 µl of 1 N NaOH; then add 750 µl PBST to dissolved paraformaldehyde). Embryos were thoroughly washed with PBST five times for 5 minutes each.

Hybridization

PBST was changed gradually to hybridization solution (50% formamide, 5 × SSC, 50 µg/ml heparin, 500 µg/ml yeast tRNA, 0.1% Tween-20) by washing embryos once in 50% hybridization solution/50% PBST and once in 100% hybridization solution. For *S. purpuratus*, embryos were washed once in 30% hybridization solution/70% PBST, once in 60% hybridization solution/40% PBST and once in 100% hybridization solution. Formamide was found to inactivate endogenous alkaline phosphate activity in the gut.

Pre-hybridization was carried out at 45°C to 50°C for 30 minutes. Probes up to 500 bases in length have performed well with no need for nicking to reduce size. Large riboprobes were subjected to hydrolysis at 65 °C for 20 - 40 min in solution of 40 mM sodium bicarbonate, 60 mM sodium carbonate. If lipids had not been extracted with toluene, hybridization probes were reduced to 100-200 nucleotides. Hybridizations are typically carried out overnight at 45°C to 50°C in 50 µl fresh hybridization solution with 1-5 ng/ml probes overnight.

Washes and Probe Detection

Following hybridization, embryos were washed gradually with 60% hybridization solution/PBST, 30% hybridization solution/PBST and 100% PBST for 5 minutes at room temperature, twice in 2 × SSC/0.1% Tween-20 for 10 minutes each at 50°C, and then incubated in 2 × SSC containing 20 µg/ ml RNase A at 37°C for 30 minutes followed by three washes in PBST at room temperature. For A/T rich probes, RNase A digestions were carried out in 6×SSC instead of 2 × SSC, which stabilized the hybrid duplex. In some cases, an additional wash in 0.2 × SSC/0.1% Tween-20 for 10 minutes at 50°C after RNase digestion was carried out. Some riboprobes show spurious binding to pigment cells, which can be eliminated by RNase digestion.

Embryos were suspended in blocking solution (2% normal sheep serum in PBST) for 30 minutes at room temperature, incubated with anti-DIG FAB-AP fragments (Boehringer Mannheim Biochemicals) at 1:1000 dilution with the blocking solution for 1 hour at room temperature, then rinsed three times for 5 minutes each in PBST. Embryos were washed twice in the solution of 100 mM NaCl, 100 mM Tris (pH 9.5), 0.1% Tween-20, twice in the alkaline phosphatase buffer (100 mM NaCl, 100 mM Tris (pH 9.5), 50 mM $MgCl_2$, 1 mM levamisole, 0.1% Tween-20), and incubated in the dark in alkaline phosphatase buffer with 4.5 µl NBT (75 mg/ml nitro blue tetrazolium in 70% DMF) and

3.5 µl BCIP (50 mg/ml 5-bromo-4-chloro-3-indolyl phosphate in 100% DMF) per ml. Colour development usually began after about 15 min. and took between 2 and 4 hours for completion. Stained embryos were semi-permanently stored in 50% glycerol/PBST at 4 C. Photography was performed on an OLYMPUS microscope (AHBS3), using differential interference contrast or bright field optics.

4.3 RESULTS

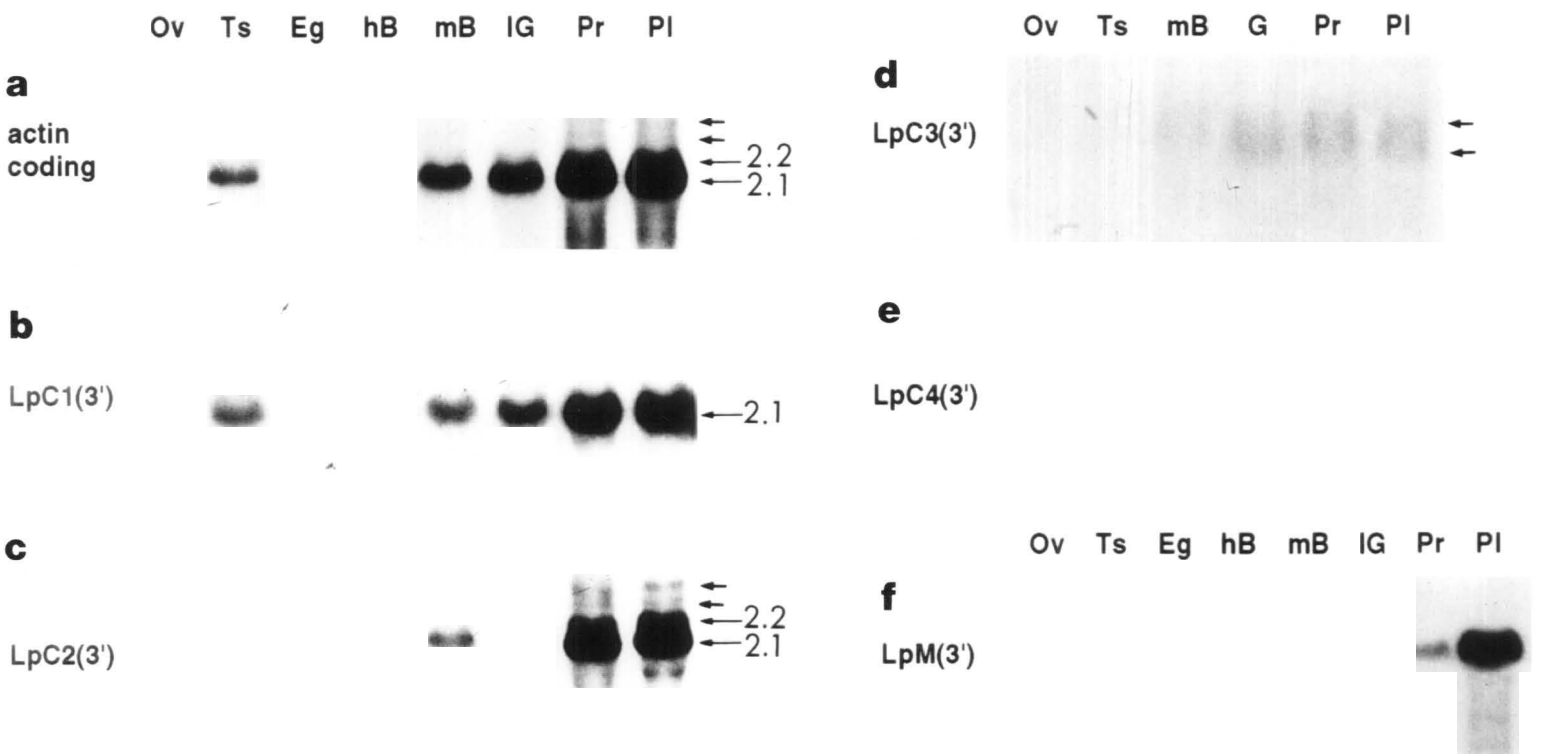
4.3.1 Temporal pattern of gene expression of the actin genes

The temporal patterns of *L. pictus* actin gene expression were investigated by RNA gel blot hybridization with gene-specific 3' noncoding sequences. Total cytoplasmic RNA was extracted from eggs as well as embryos of different stages of embryogenesis.

A Northern blot hybridized with an actin coding sequence probe shows the temporal expression of actin genes during embryogenesis (Fig. 4.1, a). Two major populations of actin transcripts, with estimated sizes of 2.1 and 2.2 Kb, were detected. There is very little actin mRNA stored in the egg or present at hatching. Actin mRNA begins to accumulate at the mesenchyme blastula and the early gastrula stages. The prevalence of actin transcripts is much greater by the late gastrula stage and is maintained at the pluteus stage. At least two other transcripts also hybridized but showed much weaker signal (see below).

RNA gel blot hybridization with gene-specific probes of the five actin genes (LpC1, LpC2, LpC3, LpC4 and LpM) indicate that all members of the actin gene family are expressed during the embryogenesis of *L. pictus* (Fig. 4.1, b-f). The LpC1 actin gene gives rise to a 2.1 Kb actin mRNA (Fig. 4.1, b). The LpC1 mRNA was hardly detectable

Fig. 4.1 Temporal patterns of actin transcript accumulation in *L. pictus* embryos and in ovary and testis. RNA gel blots were hybridized with the hybridization probes that recognize the RNA transcripts of (a) all types of actin genes; (b) the LpC1 actin gene; (c) the LpC2 actin gene; (d) the LpC3 actin gene; (e) the LpC4 actin gene; and (f) the LpM actin gene. Total RNA from ovary (Ov) and testis (Ts), and cytoplasmic RNA from eggs (Eg), and embryos at the stage of hatched blastula (hB), mesenchyme blastula (mB), late gastrula (lG), prism (Pr) and pluteus (Pl) were used.


Ov Ts Eg hB mB IG Pr Pl
a
actin coding
2.2
2.1
b
LpC1(3')
2.1
c
LpC2(3')
2.2
2.1
Ov Ts mB G Pr Pl
d
LpC3(3')
e
LpC4(3')
Ov Ts Eg hB mB IG Pr Pl
f
LpM(3')

in the RNA sample extracted from eggs and hatched blastula embryos, but begins to accumulate in the mesenchyme blastula embryos, increasing until the pluteus stage. A hybridization probe corresponding to the 3' noncoding terminus of the LpC2 actin gene, hybridizes to a 2.2 Kb actin transcript (Fig. 4.1, c). This probe cross hybridizes to the 2.1 Kb LpC1 mRNA, as well as two other transcripts with higher molecular weights (indicated by arrow heads in Fig. 4.1, c) under the hybridization conditions utilized. The use of more stringent wash conditions resulted in loss of hybridization to both actin transcripts. Small amounts of LpC2 transcripts were detected in eggs and hatched blastula embryos. In contrast to LpC1 mRNA, the LpC2 mRNA accumulates later, beginning at the gastrula stage instead of the mesenchyme blastula stage (Fig. 4.1 c). Panels b and c of Figure 4.1 reveal that the LpC1 and LpC2 mRNAs are the two most abundant actin mRNAs; under the conditions used the relative intensities provide an indication of the relative prevalence of these two transcripts. Transcripts of the other two cytoskeletal actin genes, LpC3 and LpC4, show similar temporal patterns of accumulation, but are always much less abundant than LpC1 and LpC2 mRNAs (Fig. 4.1, d, e). Neither LpC3 nor LpC4 transcripts were detected until the mesenchyme blastula stage (about 20 hours post fertilization), when a rapid increase of the LpC3 and LpC4 transcripts was observed (data not shown for the hybridizations with RNA of stages earlier than the mesenchyme blastula). The prevalence of the LpC3 and LpC4 transcripts slightly increases during gastrulation and is maintained through the prefeeding pluteus stage. The LpC3 probe, corresponding to the 3' noncoding region of the gene, hybridized to two transcripts of different size (indicated by arrow heads in Fig. 4.1, d). Since the sequence of the probe is quite different from probes corresponding to the other actin genes, it is not likely that there is a cross hybridization of this probe with other actin transcripts. It is possible that the LpC3 transcript is differentially processed to yield the transcripts of different size. Transcripts expressed by the single muscle actin gene LpM became detectable at the

gastrula stage (Fig. 4.1, f). The prevalence of the LpM transcript increases dramatically at the pluteus stage, when the esophageal contractile muscle band forms.

Expression of all five actin genes was also investigated in adult ovary and testis tissues. Total RNA extracted from ovary and testis tissues was used. RNA gel blots were hybridized with a coding sequence probe (Fig. 4.1, a), and five gene-specific probes (Fig. 4.1, b to f). Transcripts expressed from the cytoskeletal LpC1 and the muscle LpM actin genes were detected in ovary and testis. In both cases, transcripts were more abundant (relative to total RNA) in the testis cells than ovary cells. A similar result has been observed for the *S. purpuratus* CyI actin gene (Shott *et al.*, 1984). Transcripts corresponding to the LpC2, LpC3 and LpC4 actin genes were not detected in the ovary and testis RNA samples.

4.3.2 Spatial pattern of gene expression of the actin gene family

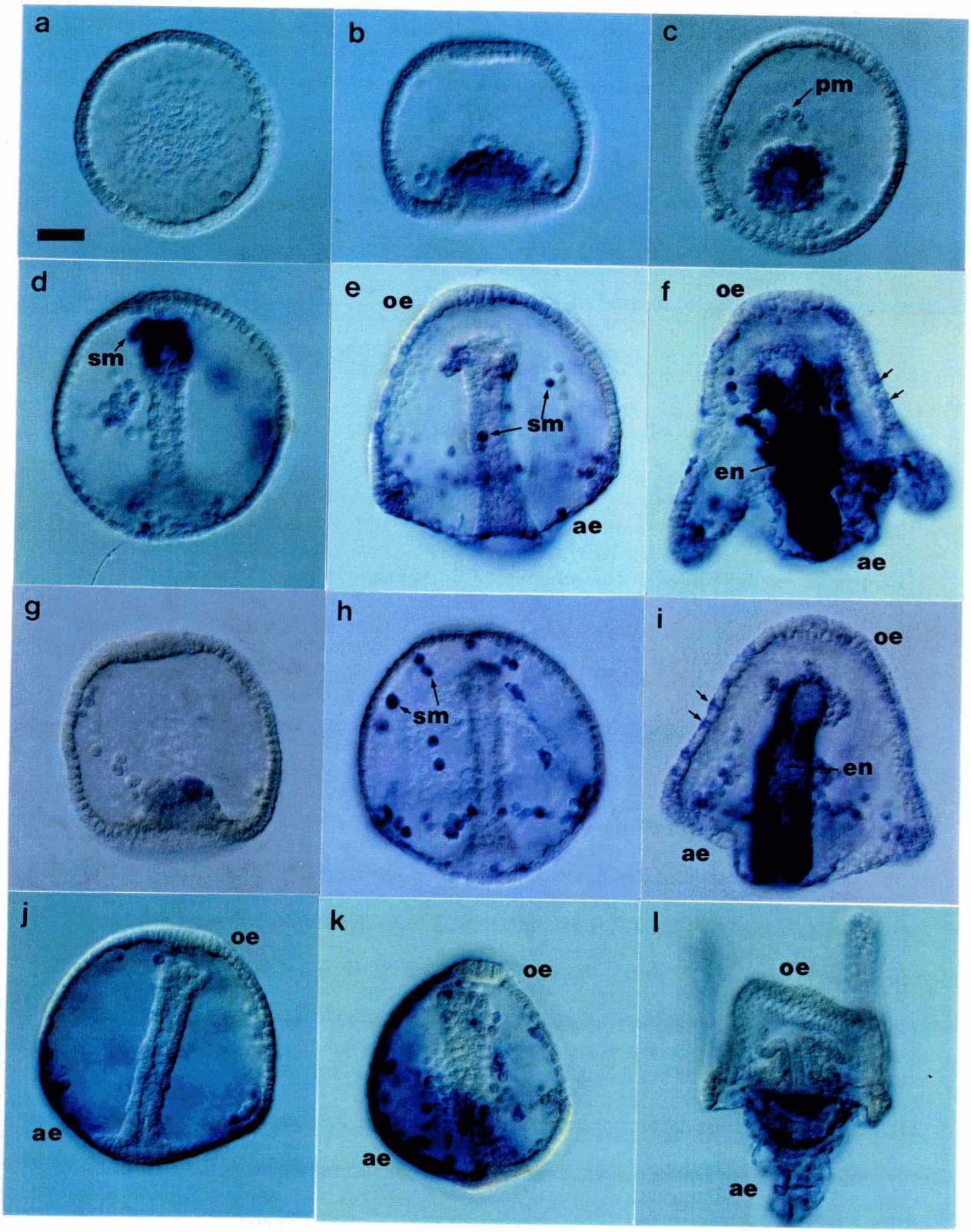
The overall spatial distribution of actin mRNAs was examined by whole mount *in situ* hybridization with a DIG-labeled actin coding sequence, which hybridizes to all actin mRNAs at the stringency used. No actin mRNA was detected in the hatched blastula embryos (Fig. 4.2, a). Actin mRNA becomes easily detectable at the mesenchyme blastula stage, after the ingression of the primary mesenchyme cells, when a disk of cells in the middle of the vegetal plate is stained (Fig. 4.2, b, c). Actin mRNAs are confined to several cells at the tip of the invaginating archenteron at the early gastrula stage (data not shown), and are continuously expressed in these prospective secondary mesenchyme cells as the archenteron elongates (Fig. 4.2, d). Toward the end of gastrulation, these cells appear to ingress from the tip of the archenteron and migrate along the ectodermal wall (Fig. 4.2, e). No detectable actin mRNAs was found in the ingressed primary mesenchyme cells before the completion of gastrulation. Cells surrounding the vegetal

plate of the early gastrula embryo were lightly stained; some of them might be ectoderm cells (Fig. 4.2, b). The staining in these cells is transient since it quickly disappears during the gastrulation. By the completion of gastrulation, there is a large increase in the prevalence of actin mRNAs in the aboral ectodermal cells. Meanwhile, staining of most parts of the archenteron became detectable, but at relatively low levels (Fig. 4.2, e). Some mesenchyme cells, including primary and secondary mesenchyme cells, are stained as well in the late gastrula. At the pluteus stages, the aboral ectodermal cells, the differentiated gut and some mesenchyme cells were heavily stained (Fig. 4.2, f). Actin mRNAs are also detected in the oral ectoderm cells of late gastrula, prism and pluteus embryos, but at very low levels. The overall pattern of total actin gene expression corresponds to the combined pattern of expression of each member of the actin gene family analyzed below.

4.3.3 The changing spatial pattern of expression of the LpC1 actin gene

The LpC1 mRNA became detectable in several cells at the center of the vegetal plate at the mesenchyme blastula stage (data not shown). During gastrulation, the LpC1 actin gene is exclusively expressed in the cells located at the tip of the elongating archenteron, the prospective secondary mesenchyme cells (Fig. 4.2, g). Later in gastrulation, these stained cells are released from the tip of the archenteron. Some, maybe all, of them migrate along the ectodermal wall, then penetrate and are retained within the ectodermal epithelium (Fig. 4.2, h); these cells are presumably the pigment-forming secondary mesenchyme cells. The LpC1 actin gene is also expressed in cells of other spatial territories (Fig. 4.2, h). Endodermal cells, especially those nearest the blastopore, begin to accumulate the LpC1 mRNAs during gastrulation. A few primary mesenchyme cells were heavily stained in gastrulae. Some heavily stained secondary mesenchyme cells were not embedded in the ectodermal wall, but were scattered the blastocoel. These cells probably belong to a subpopulation of secondary mesenchyme cells, such as the basal

Fig. 4.2 Spatial distribution of mRNAs of total actin, LpC1 and LpC2 in embryos. Total actin mRNA was detected with a cross-hybridizing coding sequence probe from the LpC1 gene. Embryos were examined by whole mount *in situ* hybridization method at (a) hatched blastula, (b, c) mesenchyme blastula, (d) late gastrula, (e) prism, and (f) pluteus stages. Panels g-i show the spatial patterns of the LpC1 mRNA accumulation; panels j-l show the spatial patterns of the LpC2 mRNA accumulation. (g) early gastrula; (h, j) late gastrula; (i) and (l), early and late pluteus, respectively; (k) prism. pm, primary mesenchyme cells; sm, secondary mesenchyme cells; en, endoderm; ae, aboral ectoderm; oe, oral ectoderm. Scale bar: 10 μm.


a
b
c
pm
d
sm
e
oe
sm
ae
f
oe
en
ae
g
h
sm
i
oe
en
ae
j
oe
ae
k
oe
ae
l
oe
ae

mesenchyme cells. It is not clear whether these cells are included among those which were heavily stained at the tip of the archenteron before their migration; alternatively, these cells may accumulate the LpC1 mRNA shortly after their ingression from the archenteron. RNA gel blot hybridizations, carried out with a LpC1 gene-specific probe, detected a dramatic increase in the prevalence of the LpC1 actin transcript at the prism and pluteus stages. The analysis of the spatial distribution of the LpC1 mRNA indicates most of this increase is due to accumulation in the differentiated gut, especially in the stomach and intestine (Fig. 4.2, i). Low levels of staining of ectodermal cells were also observed in late gastrulae and thereafter (Fig. 4.2, g,h).

4.3.4 The actin gene LpC2 is predominantly expressed in the aboral ectoderm

After LpC1 mRNA, LpC2 mRNA is the most abundant actin mRNA species expressed during the embryogenesis. Whole mount *in situ* hybridizations using a 3' noncoding sequence probe corresponding to LpC2 indicate the LpC2 gene is preferentially expressed in the aboral ectoderm at the gastrula and pluteus stages (Fig. 4.2, j,k). A few mesenchyme cells were lightly stained in late gastrula and prism embryos (Fig. 4.2, j,k). No significant staining was detected at the stages earlier than mid-gastrula (data not shown). The LpC2 3' probe also reacted with some endodermal cells of embryos at the pluteus stage; these cells were confined to the stomach and intestine (Fig. 4.3, l). Since this probe cross-hybridizes with LpC1 transcripts on RNA gel blot hybridizations, it is possible that the staining of cells other than the aboral ectoderm corresponds to transcripts of the LpC1 gene instead of the LpC2 gene. However, the observation that, under the hybridization stringency we applied, the 3' probe of LpC2 did not stain the presumptive secondary mesenchyme cells at the tip of the archenteron which specifically express the LpC1 gene, suggests the possibility that LpC2 mRNAs are indeed present in some mesenchyme and endoderm cells. The expression of the LpC2 gene is predominantly in

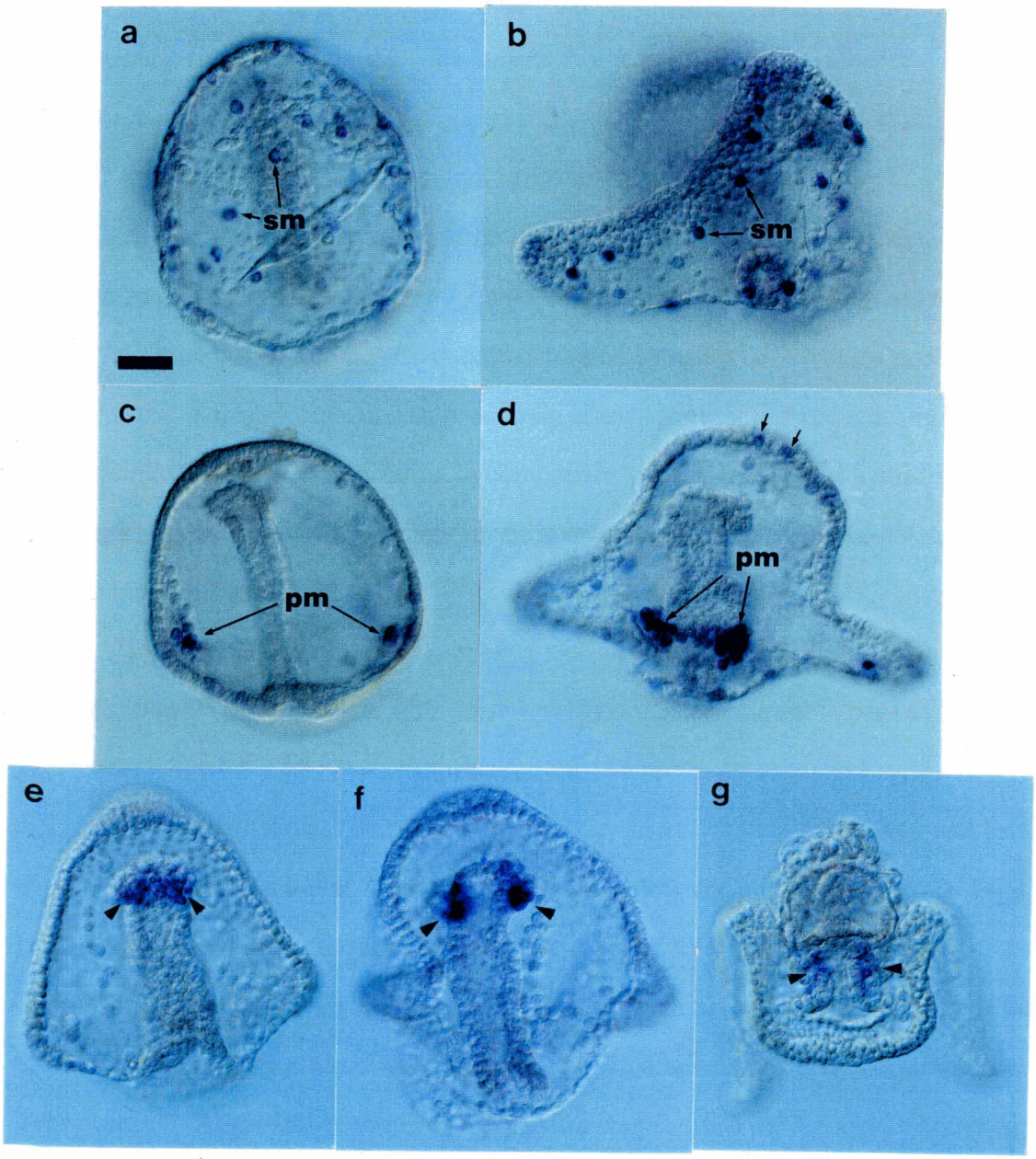
the aboral ectodermal cells. We observed that the hybridization duplex formed between the target mRNA and the DIG-labeled RNA probe is less stable than the duplex formed by using radioactive labeled RNA probe; the DIG chemical group seems to interfere with the stability of the hybridization duplex. Therefore, the apparently less stringent conditions applied for whole mount *in situ* hybridization may distinguish the DIG-labeled LpC1 and LpC2 probes, while the radioactively labeled probes cross hybridized under the more stringent conditions used for Northern blot hybridizations (Fig. 4.1, b, c).

4.3.5 The spatial patterns of gene expression of the cytoskeletal actin genes LpC3 and LpC4

The LpC3 and LpC4 transcripts are much less abundant than the LpC1 and LpC2 transcripts (Fig. 4.1, d,e). Therefore, the histochemical staining step must be carried out for a longer time, usually 5 to 6 hours; by contrast, $1^1/2$ to 2 hours of staining is usually sufficient for the *in situ* hybridizations with the LpC1 and LpC2 gene-specific probes. Under conditions similar to those normally used for the LpC1 and LpC2 probes, no significant staining was observed in the embryos hybridized with a DIG-labeled probe corresponding to 3' noncoding sequence of the LpC3 actin gene. Longer staining usually resulted in nonspecific staining in the early blastula embryos (data not shown). Some mesenchyme cells, presumably secondary, are lightly stained in late gastrula(Fig. 4.3, a) and pluteus (Fig. 4.3, b) stage embryos after 5 to 6 hours of staining with no serious nonspecific staining. Nonspecific staining is often observed for some DIG-labeled RNA probes applied to early blastula embryos; increased stringency of hybridization did not improve this problem but resulted in poor structural preservation.

The spatial pattern of expression of the LpC4 actin gene is very distinctive in comparison with other sea urchin actin genes. It is preferentially expressed in the primary

Fig. 4.3 Localization of LpC3 and LpC4 cytoskeletal actin mRNAs and LpM muscle actin mRNA. LpC3 mRNA was detected in the secondary mesenchyme cells (sm) in the embryos of late gastrula stage (a) and plueus stage (b). LpC4 mRNA accumulated in a few aggregated primary mesenchyme cells (pm) of the embryo at late gastrula stage (c), and in more primary mesenchyme cells as well as some secondary mesenchyme cells (pointed by small arrows) of the embryo at pluteus stage. LpM mRNA exclusively distributed in muscle cells and their precursors in embryos at prism (e) and pluteus (f, g) stages. Scale bar: 10 μm.


a
sm
b
sm
c
pm
d
pm
e
f
g

mesenchyme cells in *L. pictus* embryos at gastrula and pluteus stages (Fig. 4.3, c, d). LpC4 actin transcripts were first detected in several clusters of primary mesenchyme cells ventrolaterally located at the positions where the two aboral arms will later protrude. More primary mesenchyme cells were stained in plutei. Some cells embedded in the ectodermal wall, presumably the pigment cells, were also lightly stained in the late gastrula embryo (Fig. 4.3, c). Although LpC4 mRNA can be detected by RNA gel blot hybridization at the mesenchyme blastula stage (Figure 4.1, e), it has not been detected in the mesenchyme blastula embryos by whole mount *in situ* hybridization (data not shown). As discussed above, hybridization duplexes formed by using DIG-labeled probes appear to be less stable than those using radioactively labeled probes; this reduces the sensitivity of whole mount *in situ* hybridization in comparison with RNA gel blot hybridization. LpC4 mRNAs may be present at low levels in many cells at mesenchyme blastula stage, making them difficult to detect.

4.3.6 Expression of the muscle actin gene LpM is restricted to muscle cells and their precursors

On the RNA gel blot, LpM mRNA was first detected at the late gastrula stage (Fig. 4.1, f). However, this transcript was not detected by whole mount *in situ* hybridization until the late gastrula stage or the early prism stage. Failure to detect the transcript at earlier stages probably reflects the limited sensitivity of whole mount *in situ* hybridization. At the late gastrula or early prism stages, LpM mRNAs were detected in several cells at the tip of the gut (Fig. 4.3, e). In contrast to precursors of the pigment-forming cells which are also located at the tip of the archenteron during gastrulation and specifically express the LpC1 actin gene, the cells expressing the LpM actin gene belong to another subpopulation of secondary mesenchyme cells, the precursors of esophageal muscle cells. Only about 15 cells at the tip of the gut were stained for LpM transcripts.

They are arranged at the edge of the evaginating coelomic pouch on either side of the tip of the gut. Cells at the center of the gut tip, consisting of derivatives of the small micromeres which also participate in the formation of the coelomic pouch, were not stained (Fig. 4.3, e). At the pluteus stage, the gut differentiates into an esophagus, a stomach, and intestine. The muscle cells move vegetally and inwardly from the outside edges of the coelomic sacs to the site of esophagus, and form an esophageal contractile band surrounding the esophagus. LpM mRNA is always detected in these muscle cells (Fig. 4.3, f). Several esophageal cells beneath the contractile band also shown some activity of the muscle actin gene LpM (Fig. 4.3, f). The esophageal contractile band is stained in feeding larvae (Fig. 4.3, g).

4.4 DISCUSSION

Comparison of the temporal and spatial patterns of gene expression between the actin genes of *L. pictus* and *S. purpuratus* raises interesting questions concerning evolution of the actin gene regulatory systems in sea urchins, as well as regulation of orthologous genes in different species.

LpC1 mRNA is the most abundant cytoskeletal actin message in *L. pictus* embryos, and is the only cytoskeletal actin mRNA which has been detected in adult ovary and testis tissues. The other actin message found in these two adult tissues is transcribed from the LpM muscle actin gene. In comparison with actin genes expressed during *S. purpuratus* embryogenesis, we found that the expression of the LpC1 gene combines features of expression of the CyI/CyIIb and CyIIa actin genes in *S. purpuratus*. Like CyIIa, LpC1 is expressed exclusively in the presumptive secondary mesenchyme cells at the tip of the archenteron in gastrulae, and maintains its activity in these cells for a period

of time immediately following their migration from the archenteron (Cox *et al.*, 1986). The LpC1 gene is subsequently activated in ectoderm cells at very low levels, and in the gut at high levels later in the development. These latter activities of LpC1 are similar to the CyI/CyIIb actin genes, which have an indistinguishable pattern of expression at these stages (Cox et al., 1986). On the other hand, the CyI gene was reported expressed in all cells of the early blastula, in the presumptive primary mesenchyme cells prior to and immediately following their ingression, and in the oral ectoderm cells of the pluteus larva (Cox *et al.*, 1986); these features, however, are not found in the expression pattern of the LpC1 gene in *L. pictus* embryos. The differences described above imply that the spatial regulation of the LpC1 gene does not exactly resemble any of the actin genes in *S. purpuratus*. This is intriguing because the sequences of the LpC1 and CyI genes are very similar to each other: five amino acid replacements in the deduced amino acid sequences; 96.3% nucleotide similarity in the coding regions, and 81.9 % similarity in the 3' noncoding regions. A molecular phylogenetic analysis indicated that the LpC1 actin gene of *L. pictus* is more closely related to the CyI actin gene than any other actin gene of *S. purpuratus* and they may be orthologous (Chapter III). It has been reported recently that an actin gene orthologous to CyI in the sea urchin *Tripenustes gratilla*, the TgCyI actin gene, also displays a pattern of expression substantially different from the CyI actin gene, combining features of both CyI and CyIIIa/b (Wang *et al.*, 1993). A common feature of these three actin genes (CyI, LpC1 and TgCyI) is they are all expressed sequentially in several different spatial territories during development.

The actin genes LpC2 of *L. pictus* and CyIIIa/b of *S. purpuratus* display similar but distinguishable patterns of expression. They are all predominantly expressed in aboral ectodermal cells. The prevalence of LpC2 transcripts increase at the late gastrula stage. This temporal pattern of LpC2 gene expression is similar to CyIIIb expression; by contrast, CyIIIa mRNAs begin to accumulate prior to hatching (Shott *et al.*, 1984; Cox *et*

al., 1986). Neither LpC2 nor CyIIIb is expressed in adult ovary and testis tissues (Shott *et al.*, 1984, and this study), distinguishing them from the CyIIIa gene. Sequence comparisons based on the deduced amino acid sequences of the third exon of actin genes of *L. pictus* and *S. purpuratus* indicate that the LpC2 and CyIIIb genes are more closely related to one another than they are to CyIIIa. However, comparisons of the entire nucleotide coding sequences or complete amino acid sequences of these three genes indicate the two CyIII genes are more similar to each other than they are to the LpC2 gene (Chapter III). The 3' noncoding region of the LpC2 gene is 77.5% similar to CyI gene of *S. purpuratus*, but there is no obvious sequence similarity to the 3' noncoding terminus of the CyIII genes.

Molecular phylogenetic analysis of sea urchin actin genes suggests that CyI-like genes most resemble to the common ancestral cytoskeletal actin gene in echinoids (Chapter III). We suspect that, at the time when there was only one ancestral cytoskeletal actin gene, it was expressed in all cell types. This ancestral actin gene duplicated and gave rise to all the cytoskeletal actin genes in modern sea urchin species. The duplicated actin genes presumably acquired new regulatory sequence elements. Meanwhile, genes encoding transcription factors acting in *trans* to the *cis* acting regulatory elements also diverged as the species diverged. Consequently, actin genes became differentially regulated by sets of *trans*-acting factors which are qualitatively and/or quantitatively different in different types of cells, and/or in different sea urchin species. We thus suspect that the pattern of CyI expression in all cells of the early blastula of *S. purpuratus* (Cox *et al.*, 1986) is primitive. The more complex, restricted and differentially regulated expression patterns of extant cytoskeletal actin genes (including CyI in older embryos) among different sea urchin species are the consequence of evolutionary divergence of the gene regulatory systems.

Although the spatial patterns of expression of CyI and TgCyI are very distinctive, the 5' flanking sequences of CyI and TgCyI are highly similar (Wang *et al.*, 1993). Therefore, it was suggested that the distinctive spatial patterns of the CyI and TgCyI gene expressions are attributable to differences in *trans*-acting factors between the two sea urchin species (Wang *et al.*, 1993). *Tripneustes* and *Lytechinus* are phylogenetically more closely related to each other than either is to *Strongylocentrotus*. It is thus reasonable to speculate that the regulatory systems, including *cis*-acting promoter elements and *trans*-acting factors, may be similar for the TgCyI and LpC1 actin genes. However, the spatial patterns of TgCyI and LpC1 mRNA accumulation during embryogenesis are distinctive; in particular, the early and high-level activity of the TgCyI actin gene in the aboral ectoderm has not been observed for the LpC1 gene in *L. pictus* embryos.

It has been proposed that gene regulatory differences among species are achieved through changes in the concentration of regulatory factors and/or changes in the DNA binding affinity of theses factors (Cavener, 1992). It is possible that differences among the *trans* acting regulatory systems for the expression of the CyI, LpC1, and TgCyI genes are more quantitative than qualitative. The fact that nuclear extracts from *L. pictus* embryos include proteins which bind to all known *cis* elements of the *S. purpuratus* CyI promoter supports this possibility (Ganster et al., 1992). On the other hand, genes with apparently similar expression patterns can be subject to quite different regulatory systems. This is true for the two CyIII genes: although they display similar aboral ectoderm specific expression, CyIIIa and CyIIIb have quite distinct 5' regulatory sequences (Theze *et al.*, 1990; Niemeyer and Flytzanis, 1993). The spatial pattern of CyIIIa expression is negatively regulated by several DNA:protein interaction elements, such as P7II and P3A (Coffman and Davidson, 1992, review). Previous observations of the ectopic expression of CyIIIa promoter fusion genes in *L. variegatus* embryos (Franks *et al.*, 1988), and the correct expression of the CyIIIa gene in hybrid embryos generated between *S. purpuratus*

and *L. pictus*, suggest that the negative *trans*-acting factors are products of zygotic gene activity and are missing in *L. variegatus* and *L. pictus* embryos. There have been no reported investigations of the activity of the CyIIIb promoter in *L. pictus* embryos. It will be interesting to establish whether CyIIIb and LpC2 genes share similar 5' regulatory sequence elements; whether promoter fusion genes of these genes are regulated correctly in the other species; and whether these genes are expressed correctly in hybrid embryos. Results from these investigations should provide insight into gene regulatory mechanisms and how they have evolved.

The gradual accumulation of LpC4 mRNA in primary mesenchyme cells in the late gastrula and pluteus embryos is unique among all sea urchin actin genes examined to date. The accumulation of the message is initiated in a few primary mesenchyme cells of the ventrolateral clusters, the sites of initiation of the formation of triradiate spicule rudiments and elongation of these skeletal spicules (Ettensohn and Ingersoll, 1992). In the pluteus embryo, more primary mesenchyme cells contain LpC4 actin mRNAs; these cells form the syncytial cables along which the spicules are formed. The activity of the LpC4 actin gene may be a marker of primary mesenchyme cells as begin to actively engage in skeletogenesis.

The expression of the single muscle actin gene LpM is spatially restricted to the differentiated muscle and their precursors, similar to the expression of the orthologous SpM actin gene of *S. purpuratus*. However, unlike the SpM gene which becomes active at the early pluteus stage (Shott *et al.*, 1984), the LpM actin gene is expressed in the late gastrula embryos of *L. pictus*.

In summary, all five members of the *L. pictus* actin gene family are differentially expressed during development of the embryo. None of the members of the *L. pictus* actin

gene family shows a pattern of expression identical to any actin gene of *S. purpuratus*. The evolutionary processes responsible for these differences are obscure, but comparative experimental investigations of gene regulation should help to illuminate them.

SUMMARY OF CONCLUSIONS

The objectives of this thesis research were to isolate and characterize recombinant DNA clones corresponding to actin genes of the sea urchin *Lytechinus pictus*, to investigate the evolutionary history of sea urchin actin gene families, and to determine the expression of the *L. pictus* actin genes during embryogenesis. The following observations and conclusions have been obtained from this research:

1. Several genomic DNA and cDNA clones corresponding to actin genes of *L. pictus* were recovered by screening genomic and complementary DNA libraries. Two cDNA clones, which corresponded to the LpC1 and LpC2 cytoskeletal actin genes were fully sequenced. The 3' noncoding region and the third actin encoding exon of several actin genes other than LpC1 and LpC2 genes recovered from a genomic library were partially sequenced. Three additional actin genes were identified and designated as LpC3, LpC4 and LpM.

2. The *L. pictus* actin gene family has five members: a single muscle actin gene LpM, and four cytoskeletal actin genes LpC1-LpC4.

3. No linkage of actin genes was detected in the genome of *L. pictus*.

4. The nucleotide sequences of the 3' noncoding region of the LpC1 and LpC2 genes had a similarity of 89%, and were conserved in some cytoskeletal actin genes of several other echinoids. The 3' noncoding sequences of LpC3 and LpC4 showed no significant similarity compared to the LpC1 and LpC2 actin genes, or to the actin genes of *S. purpuratus*.

5. A cluster of amino acids, (Serine)-Threonine-(Leucine)-Serine-Isoleucine located between codon 258 to 267, are found conserved in LpC2 actin of *L. pictus*, and CyIIIa and CyIIIb of *S. purpuratus*.

6. The third codon positions of the *L. pictus* actin genes are highly biased for C, similar to the *S. purpuratus* actin genes.

7. The molecular phylogeny of actin genes of echinoids was analyzed. Comparison of deduced amino acid sequences indicated a close relationship between the LpC1 and the CyI/CyII subfamilies of *S. purpuratus* actin genes, and between the LpC2 and the CyIII subfamily of *S. purpuratus* actin genes. The LpM and SpM muscle actin genes surveyed are orthologous. In contrast, nucleotide sequence comparisons suggested that the extant *L. pictus* cytoskeletal actin genes arose since the two species diverged from a common ancestor, but there appeared to have been some homogenization of sequences within the species, casting doubt on this interpretation.

8. A model was proposed in which two divergent cytoskeletal actin genes of the common ancestral sea urchin gave rise by duplication to the extant cytoskeletal actin genes, some of which have changed 3' noncoding sequences while others have maintained a highly conserved terminus.

9. All the *L. pictus* actin genes are expressed during the embryogenesis, each with a distinct pattern of expression. The LpC1 and LpC2 messenger RNAs were most abundant; LpC1 mRNA accumulated in several spatial territories, including secondary mesenchyme cells, endoderm, and ectoderm, while LpC2 mRNA accumulated predominantly in the aboral ectoderm. The LpC3 and LpC4 mRNAs were much less

abundant in comparison with LpC1 and LpC2 mRNAs. LpC3 mRNA was detected in secondary mesenchyme cells in late gastrulae and plutei; LpC4 mRNA was expressed in the primary mesenchyme cells beginning late in gastrulation. The expression of the LpM muscle actin gene was spatially restricted to the differentiated muscle cells of the esophagus and their precursors.

In summary, I have investigated gene number, gene structure, and gene expression of the sea urchin *L. pictus* actin gene family, and actin gene evolution of echinoids as well. Several actin genes of *L. pictus* were cloned and characterized for the first time. The information obtained from this research will enrich our knowledge of the structure, expression, and evolution of actin genes. The differential expression patterns of actin genes examined in this study will be useful for further investigations of mechanisms of territorial gene regulation in normal sea urchin embryos and in interspecies hybrids as well. The findings that the 3' noncoding sequences of LpC1 and LpC2, and a cluster of amino acids in the third exon of the LpC2 actin and the CyIII actins of *S. purpuratus,* were highly conserved during gene evolution raise interesting questions concerning the functional significance of the conserved 3' noncoding sequences and this cluster of amino acids.

REFERENCES

Akhurst, R. J., F. J. Calzone, J. J. Lee, R. J. Britten, and E. H. Davidson (1987) Structure and organization of the CyIII actin gene subfamily of the sea urchin, *Strongylocentrotus purpuratus*. J. Mol. Biol. 194:193-203

Albertson, D. G. (1985) Mapping muscle protein genes by in situ hybridization using biotin-labeled probes. EMBO J. 4:2493-2498

Beach, R. L. and W. R. Jeffery (1992) Multiple actin genes encoding the same *a*-muscle isoform are expressed during ascidian development. Dev. Biol. 151:55-66

Benson, D., D. J. Lipman, and J. Ostell (1993) Genebank. Nucleic Acids Res. 21:2963-2965

Blin, N. and D. W. Stafford (1976) Isolation of highmolecular-weight DNA. Nucleic Acids Res. 3:2303

Brandhorst, B. P., M. Filion, P. E. Nisson, and W. R. Crain (1991) Restricted expression of the *Lytechinus pictus* Spec1 gene homologue in reciprocal hybrid embryos with *Strongylocentrotus purpuratus*. Dev. Biol. 144:405-411

Brandhorst, B. P. and W. H. Klein (1992) Territorial specification and control of gene expression in the sea urchin embryo. Sem. Dev. Biol. 3:175-186

Brennan, S. M. (1990) Transcription of endogenous and injected cytoskeletal actin genes during early embryonic development in *Xenopus laevis*. Differentiation 44:111-121

Britten, R. J. (1993) Forbidden synonymous substitution in coding regions. Mol. Biol. Evol. 10:205-220

Bullock, B. P., P. E. Nisson, and W. R. Crain, Jr. (1988) The timing of expression of four actin genes and an RNA Polymerase III-transcribed repeated sequence is correct in hybrid embryos of the sea urchin species *Strongylocentrotus purpuratus* and *Lytechinus pictus*. Dev. Biol. 130:335-347

Burn, T. C., J. O. Vigoreaux, S. L. Tobin (1989) Alternative 5C actin transcripts are localized in different patterns during *Drosophila* embryogenesis. Dev. Biol. 131:345-355

Cabot, E. L. and A. T. Beckenbach (1989) Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comput. Appl. Biosci. 5:233-234

Calzone, F. J., N. Theze, P. Thiebaud, R. L. Hill, R. J. Britten, and E. H. Davidson (1988) Developmental appearance of factors that bind specifically to *cis*-regulatory sequences of a gene expressed in the sea urchin embryo. Genes & Dev. 2:1074-1088

Cameron, R. A., B. R. Hough-Evans, R. J. Britten, and E. H. Davidson (1987) Lineage and fate of each blastomere of the eight-cell sea urchin embryo. Genes & Dev. 1:75-85

Cameron, R. A. and E. H. Davidson (1991) Cell type specification during sea urchin development. Trends in Genetics. 7:212-218

Cavener, D. R. (1992) Transgenic animal studies on the evolution of genetic regulatory circuitries. BioEssays 14:237-244

Chung, Y. T. and E. B. Keller (1990) Regulatory elements mediating transcription from the *Drosophila melanogaster* actin 5C proximal promoter. Mol. Cell. Biol. 10:206-216

Chung, Y. T. and E. B. Keller (1990) Positive and negative regulatory elements mediating transcription from the *Drosophila melanogaster* actin 5C distal promoter. Mol. Cell. Biol. 10:6172-6180

Chung, Y. T. and E. B. Keller (1991) The TATA-dependent and TATA-independent promoters of the *Drosophila melanogaster* actin 5C-encoding gene. Gene 106:237-241

Chen, P. S. and F. Baltzer (1975) Morphology and biochemistry of diploid and androgenetic haploid (merogonic) hybrids. In The Sea Urchin Embryo: Biochemistry and Morphogenesis. ed. G. Czihak, pp.424-472. Springer-Verlag, Berlin.

Coffman, J. A. and E. H. Davidson (1992) Expression of spatially regulated genes in the sea urchin embryo. Current Opinion in Genetics & Development 2:260-268

Collins, J. H. and M. Elzinga (1975) The primary structure of actin from rabbit skeletal muscle. J. Biol. Chem. 250:5915-5920

Collura, R. and K. S. Katula (1992) Spatial pattern of expression of CyI actin-β-galactosidase fusion genes injected into sea urchin eggs. Dev. Growth & Differ. 34:635-647

Conlon, R. A., F. Tufaro, and B. P. Brandhorst (1987) Post-transcriptional restriction of gene expression in sea urchin interspecies hybrid embryos. Genes & Dev. 1:337-346

Cooper, A. D. and W. R. Crain, Jr. (1982) Complete nucleotide sequence of a sea urchin actin gene. Nucleic Acids Res. 10:4081-4092

Courchesene-Smith, C. L. and S. L. Tobin (1989) Tissue-specific expression of the 79B actin gene during *Drosophila* development. Dev. Biol. 133:313-321

Cox, K. H., L. M. Angerer, J. J. Lee, E. H. Davidson, and R. C. Angerer (1986) Cell lineage-specific programs of expression of multiple actin genes during sea urchin embryogenesis. Mol. Biol. 188:159-172

Crain, W. R., Jr. (1988) Regulation of actin gene expression during sea urchin development. Dev. Biol. 5:189-208

Crain, W. R., Jr., M. F. Boshar, A. D. Cooper, D. S. Durica, A. Nagy, and D. Steffen (1987) The sequence of a sea urchin muscle actin gene suggests a gene conversion with a cytoskeletal actin gene. J. Mol. Evol. 25:37-45

Crain, W. R., Jr. and F. D. Bushman (1983) Transcripts of paternal and maternal actin gene alleles are present in interspecific sea urchin embryo hybrids. Dev. Biol. 100:190-196

Crain, W. R., Jr., D. S. Durica, and K. V. Doren (1981) Actin gene expression in developing sea urchin embryos. Mol. Cell. Biol. 1:711-720

Crawford, B. D. and R. D. Burke (1994) YIGSR domain of laminin binds surface receptors of mesenchyme and stimulates migration during gastrulation in sea urchin. submitted

Cross, G. S., C. Wilson, H. P. Erba, and H. R. Woodland (1988) Cytoskeletal actin gene families of *Xenopus borealis* and *Xenopus laevis*. J. Mol. Evol. 27:17-28

Czihak, G. (1975) The sea urchin embryo. Springer, Berlin Heidelberg New York

Davidson, E. H., T. L. Thomas, R. H. Scheller and R. J. Britten (1982) The sea urchin actin genes, and a speculation on the evolutionary significance of small gene families. Pp.177-191 in G. A. Dover and R. B. Flavell, eds. Genome Evolution. Systematics Association, Academic Press, New York

Davidson, E. H. (1989) Lineage-specific gene expression and the regulative capacities of the sea urchin embryo: a proposed mechanism. Development 105:421-445

Durica, D. S., J. A. Schloss, and W. R. Crain, Jr. (1980) Organization of actin gene sequences in the sea urchin: molecular cloning of an intron-containing DNA sequence coding for a cytoplasmic actin. PNAS USA 77:5683-5687

Durica, D. S., D. Garza, M. A. Restrepo, and M. M. Hryniewicz (1988) DNA sequence analysis and structural relationships among the cytoskeletal actin genes of the sea urchin *Strongylocentrotus purpuratus*. J. Mol. Evol. 28:72-86

Efstradiatis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connell, R. A. Spritz, J. De Riel, B. G. Forget, S. M. Weissman, J. L. Slightom, A. E. Blechl, O. Smithies, F. E. Baralle, C. C. Shoulders and N. J. Proudfoot (1980) The structure and evolution of the human β-globin gene family. Cell 21:653-668

Engel, J. N., P. W. Gunning, and L. Kedes (1981) Isolation and characterization of human actin genes. PNAS USA 78:4674-4678

Erba, H. P., R. Eddy, T. Shows, L. Kedes, and P. Gunning (1988) Structure, chromosome location, and expression of the human gamma-actin gene: differential evolution, location, and expression of the cytoskeletal beta- and gamma-actin genes. Mol. Cell. Biol. 8:1775-1789

Erba, H. P., P. Gunning, and L. Kedes (1986) Nucleotide sequences of the human γ cytoskeletal actin mRNA: anomalous evolution of vertebrate non-muscle actin genes. Nucleic Acids Res. 14:5275-5294

Ettensohn, C. A. and E. P. Ingersol (1992) Morphogenesis of the sea urchin embryo. Pp. 189-262 in E. F. Rossomando and S. Alexander, eds. Morphogenesis. Wiley-Liss, New York

Felsenstein, J. (1989) PHYLIP-phylogeny inference package (version 3.2). Cladistics 5:164-166

Files, J. G., S. Carr, and D. Hirsh (1983) Actin gene family of *Caenorhabditis elegans*. J. Mol. Biol. 164:355-375

Finnegan, D. J., G. M. Rubin, M. W. Young, and D. S. Hogness (1978) Repeated gene families in *Drosophila melanogaster*. Cold Spring Harbor Symposia on Quantitative Biology 42 Pt. 2:1053-1063

Firtel, R. A., R. Timm, A. R. Kimmel, and M. McKeown (1979) Unusual nucleotide sequences at the 5' end of actin genes in *Dictyostelium discoideum*. Proc. Natl. Acad. Sci. USA. 76:6206-6210

Flytzanis, C. N., E. A. Bogosian, and C .C. Niemeyer (1989) Expression and structure of the CyIIIb actin gene of the sea urchin *Strongylocentrotus purpuratus*. Mol. Reprod. & Dev. 1:208-218

Foran, D. R., P. J. Johnson, and G. P. Moore (1985) Evolution of two actin genes in the sea urchin *Strongylocentrotus franciscanus*. J. Mol. Evol. 22:108-116

Fornwald, J. A., G. Kuncio, I. Peng, and C. P. Ordahl (1982) The complete nucleotide sequence of the chick α-actin gene and its evolutionary relationship to the actin gene family. Nucleic Acids Res. 10:3861-3876

Franks, R. R., R. Anderson, J. G. Moore, B. R. Hough-Evans, R. J. Britten, and E. H. Davidson (1990) Competitive titration in living sea urchin embryos of regulatory factors required for expression of the CyIIIa actin gene. Development 110:31-40

Franks, R. R., B. R. Hough-Evans, R. J. Britten, and E. H. Davidson (1988) Spatially deranged though temporally correct expression of *Strongylocentrotus purpuratus* actin gene fusion in transgenic embryos of a different sea urchin family. Genes & Dev. 2:1-12

Franks, R. R., B. R. Hough-Evans, R. J. Britten, and E. H. Davidson (1988) Direct introduction of cloned DNA into the sea urchin zygote nucleus, and fate of injected DNA. Development 102:287-299

Fyrberg, E. A., B. J. Bond, N. D. Hershey, K. S. Mixter, and N. Davidson (1981) The actin genes of *Drosophila*: protein coding regions are highly conserved but intron positions are not. Cell 24:107-116

Fyrberg, E. A., J. W. Mahaffey, B. J. Bond, and N. Davidson (1983) Transcripts of the six *Drosophila* actin genes accumulate in a stage- and tissue-specific manner. Cell 33:115-123

Gallwitz, D. and R. Seidel (1980) Molecular cloning of the actin gene from yeast *Saccharomyces cerevisiae*. Nucleic Acids Res. 8:1043-1059

Ganster, R., H. Paul, and K. S. Katula (1992) Analysis of the DNA binding proteins interacting with specific upstream sequences of the *S. purpuratus* CyI actin gene. Mol. Reprod. & Dev. 33:392-406

Garcia, R., B. Paz-Aliaga, S. G. Ernst, and W. R. Crain, Jr. (1984) Three sea urchin actin genes show different patterns of expression: muscle specific, embryo specific, and constitutive. Mol. Cell. Biol. 4:840-845

Giudice, G. (1973) Developmental biology of the sea urchin embryo. Academic Press, London, New York, p 469

Giudice, G. (1986) Chapter 2: Embryonic Morphogenesis. Pp. 36-43 in The Sea urchin embryo. A developmental biological system. Springer-Verlag Berlin Heidelberg, Germany.

Gunning, P., P. Ponte, H. Blau, and L. Kedes (1983) α-skeletal and α-cardiac actin genes are coexpressed in adult human skeletal muscle and heart. Mol. Cell. Biol. 3:1985-1995

Gunning, P., P. Ponte, H. Okayama, J. Engel, H. Blau, and L. Kedes (1983) Isolation and characterization of full-length cDNA clones for human alpha-, beta-, and gamma-actin mRNAs: skeletal but not cytoplasmic actins have an amino-terminal cysteine that is subsequently removed. Mol. Cell. Biol. 3:787-795

Gunning, P. T. Mohun, S-Y. Ng, P. Ponte, and L. Kedes (1984) Evolution of the human sarcomeric-actin genes: evidence for units of selection within the 3' untranslated regions of the mRNAs. J. Mol. Evol. 20:202-214

Gustafson, T. and H. Kinnander (1956) Micro aquaria for time-lapse cinematographic studies of morphogenesis in swimming larvae and observations on gastrulation. Exp. Cell Res. 11:36-51

Hambly, B. D., J. A. Barden, M. Miki, C. G. dos Remedios (1986) Structural and functional domains on actin. Bioessays 4:124-128

Hardin, J. and D. R. McClay (1990) Target recognition by the archenteron during sea urchin gastrulation. Dev. Biol. 142:86-102

Harkey, M. A., H. R. Whiteley, and A. H. Whiteley (1992) Differential expression of the msp130 gene among skeletal lineage cells in the sea urchin embryos: a three dimensional *in situ* hybridization analysis. Mechanisms of Development 37:173-184

Hemmati-Brivanlou, A., D. Frank, M. E. Bolce, B. D. Brown, H. L. Sive and R. M. Harland (1990) Localization of specific mRNAs in *Xenopus* embryos by whole-mount *in situ* hybridization. Development 110:325-330

Hightower, R. C. and R. B. Meagher (1986) The molecular evolution of actin. Genetics 114:315-332

Hiromi, Y., and Y. Hotta (1985) Actin gene mutations in *Drosophila*: Heat shock activation in the indirect flight muscles. EMBO J. 4:1681-1687

Holmes, K. C., D. Popp, W. Gebhard, W. Kabsch (1990). Atomic model of the actin filament. Nature 347:44-49

Hörstadius, S. (1973) Experimental embryology of echinoderms. Clarendon Press, Oxford.

Ishikawa, H., R. Bischoff, and H. Holtzer (1969) Formation of arrowhead complexes with heavy meromyosin in a variety of cell types. J. Cell Biol. 43:312-328

Johnson, P. J., D. R. Foran, and G. P. Moore (1983) Organization and evolution of the actin gene family in sea urchins. Mol. Cell. Biol. 3:1824-1833

Kabsch, W., J. Vandekerckhove (1992) Structure and function of actin. Annual Review of Biophysics & Biomolecular Structure 21:49-76

Katula, K. S., B. R. Hough-Evans, R. J. Brittern, and E. H. Davidson (1987) Ontogenic expression of a CyI actin fusion gene injected into sea urchin eggs. Development 101:437-447

Kedes, L., H. Blau, P. Gunning, P. Ponte, C-P Chiu, W. Bains, J. Engel, and S. Guttman (1984) Molecular genetics of human myogenesis. Exp. Biol. Med. 9:202-210

Kindle, K. L. and R. A. Firtel (1978) Identification and analysis of *Dictyostelium* actin genes, a family of moderately repeated genes. Cell 15:763-778

Kost, T. A., N. Theodorakis, and S. H. Hughes (1983) The nucleotide sequence of the chick cytoplasmic β-actin gene. Nucl. Acids Res. 11:8287-8301

Kovesdi, I. and M. J. Smith (1985) Actin gene number in the sea star *Pisaster ochraceus*. Can. J. Biochem. Cell Biol. 63:1145-1151

Kovilur, S., J. W. Jacobson, R. L. Beach, W. R. Jeffery, and C. R. Tomlinson (1993) Evolution of the chordate muscle actin gene. J. Mol. Evol. 36:361-368

Kowbel, D. J. and M. J. Smith (1989) The genomic nucleotide sequences of two differentially expressed actin-coding genes from the sea star *Pisaster ochraceus*. Gene 77:297-308

Krause, M., M. Wild, B. Rosenzweig, and D. Hirsh. 1989 Wild-type and mutant actin genes in *Caenorhabditis elegans*. J. Mol. Biol. 208:381-392

Lee, J. J., R. J. Shott, S. J. Rose, T. L. Thomas, R. J. Britten, and E. H. Davidson (1984) Sea urchin actin gene subtypes. Gene number, linkage and evolution. J. Mol. Biol. 172:149-176

Lee, J. J., F. J. Calzone, R. J. Britten, R. C. Angerer, and E. H. Davidson (1986) Activation of sea urchin actin genes during embryogenesis: measurement of transcript Accumulation from five different genes in *Strongylocentrotus purpuratus*. J. Mol. Biol. 188:173-183

Lepage, T., C. Sardet, and C. Gache (1992) Hatching enzyme gene spatial expression. Dev. Biol. 150:23-32

Levinson, G. and G. A. Gutman (1987) Slipped-strand misparing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. 4:203-221

Li, W-H., D. Graur (1991) Fundamentals of molecular evolution. Sinauer Associates, Sunderland, Mass.

Li, W-H., C-I. Wu, and C-C. Luo (1985) A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol. Biol. Evol. 2:150-174

Lloyd, G. W. (1983) In The cytoskeleton in plant growth and development. pp.3-29. Academic Press, London

Lloyd, C. and P. Gunning (1993) Noncoding region of the γ-actin gene influence the impact of the gene on myoblast morphology. J. Cell. Biol. 121:73-82

Mahaffey, J. W., M. D. Coutu, E. A. Fyrberg, and W. Inwood (1985) The flightless *Drosophila* mutant raised has two distinct genetic lesions affecting accumulation of myofibrillar proteins in flight muscles. Cell 40:101-110

McElroy, D., M. Rothenberg, K. S. Reece, R. Wu (1990) Characterization of the rice (*Oryza sativa*) actin gene family. Plant Mol. Biol. 15:257-268

McKeown, M., K. P. Hirth, C. Edwards, and R. A. Firtel (1982) Examination of the regulation of the actin multigene family in *Dictyostelium discoideum.* Prog. Clinic. & Biol. Res. 85 Pt A:51-78

McKeown, M., W. C. Taylor, K. L. Kindle, R. A. Firtel, W. Bender, and N. Davidson (1978) Multiple, heterogeneous actin genes in *Dictyostelium.* Cell 15:789-800

Minor, J. E., J. J. Lee, R. J. Akhurst, P. S. Leahy, R. J. Britten, and E. H. Davidson (1987) Sea urchin actin gene linkages determined by genetic segregation. Dev. Biol. 122:291-295

Miwa, T., Y. Manabe, K. Kurokawa, S. Kamada, N. Kanda, G. Bruns (1991) Structure, chromosome location, and expression of the human smooth muscle (enteric type) gamma-actin gene: evolution of six human actin genes. Mol. & Cell. Biol. 11:3296-3306

Mohun, T. J. and N. Garrett (1987) An amphibian cytoskeletal-type actin gene is expressed exclusively in muscle tissue. Development 101:393-402

Mohun, T., N. Garrett, and R. Treisman (1987) *Xenopus* cytoskeletal actin and human c-fos gene promoters share a conserved protein-binding site. EMBO J. 6:667-673

Morgan, T. H. (1927) The development of species hybrids. In Experimental embryology, pp. 608-657. Columbia University Press: New York

Mounier, N., M. Gouy, D. Mouchiroud, and J. C. Prudhomme (1992) Insect muscle actins differ distinctly from invertebrate and vertebrate cytoplasmic actins. J. Mol. Evol. 34:406-415

Nellen, W., C. Donath, M. Moos, and D. Gallwitz (1981) The nucleotide sequences of the actin genes from *Saccharomyces carlsbergensis* and *Saccharomyces cerevisiae* are identical except for their introns. J. Mol. & Appl. Genetics 1:239-244

Ng, R., H. Domdey, G. Larson, J. J. Rossi, and J. Abelson (1985) A test for intron function in the yeast actin gene. Nature 314:183-184

Niemeyer, C. C. and C. N. Flytzanis (1993) Upstream elements involved in the embryonic regulation of the sea urchin CyIIIb actin gene: temporal and spatial specific interactions at a single *cis*-acting element. Dev. Biol. 156:293-302

Nisson, P. E., L. E. Dike, and W. R. Crain, Jr. (1989) Three *Strongylocentrotus purpuratus* actin genes show correct cell-specific expression in hybrid embryos of *S. purpuratus* and *Lytechinus pictus*. Development 105:407-413

Nudel, U., R. Zakut, M. Shani, S. Neuman, Z. Levy, and D. Yaffe (1983) The nucleotide sequence of the rat cytoplasmic beta-actin gene. Nucleic Acids Res. 11:1759-1771

Ordahl, C. P. and T. A. Cooper (1983) Strong homology in promoter and 3'-untranslated regions of chick and rat alpha-actin genes. Nature 303:348-349

Ponte, P., P. Gunning, H. Blau, and L. Kedes (1983) Human actin genes are single copy for alpha-skeletal and alpha-cardiac actin but multicopy for beta- and gamma-cytoskeletal genes: 3' untranslated regions are isotype specific but are conserved in evolution. Mol. & Cell. Biol. 3:1783-1791

Pulak, R. A., and P. Anderson (1988) Structure of spontaneous deletions in *Caenorhabditis elegans*. Mol. Cell. Biol. 8:3748-3754

Rastinejad, F. and H. M. Blau (1993) Genetic complementation reveals a novel regulatory role for 3' untranslated regions in growth and differentiation. Cell 72:903-917

Saitou, N. and M. Nei (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406-425

Sanchez, F., S. L. Tobin, U. Rdest, E. Zulauf, and B. McCarthy (1983) Two *Drosophila* actin genes in detail. Gene structure, protein structure and transcription during development. J. Mol. Biol. 163: 533-551

Schuler, M. A. and E. B. Keller (1981) The chromosomal arrangement of two linked actin genes in the sea urchin *S. purpuratus*. Nucl. Acids Res. 9:591-604

Schliwa, M. (1981) Protein associated with cytoplasmic actin. Cell 25:587-590

Schwob, E. and R. P. Martin (1992) New yeast actin-like gene required late in the cell cycle. Nature 355:179-182

Shah, D. M., R. C. Hightower, and R. B. Meagher (1983) Genes encoding actin in higher plants: intron positions are highly conserved but the coding sequences are not. J. Mol. & Appl. Genetics 2:111-126

Shott, R. J., J. J. Lee, R. J. Britten, and E. H. Davidson (1984) Differential expression of the actin gene family of *Strongylocentrotus purpuratus*. Dev. Biol. 101:295-306

Smith, A. B. (1988) Phylogenetic relationship, divergence times, and rates of molecular evolution for camarodont sea urchins. Mol. Biol. Evol. 5(4):345-365

Sprinkle, J. and P. M. Kier (1987) Chapter 18: Phylum Echinodermata. Pp. 505-611 in R. S. Boardman, A. H. Cheetham, A. J. Rowell, eds. Fossil Invertebrates. Blackwell Scientific Publications

Stossel, T. P. (1984) Contribution of actin to the structure of the cytoplasmic matrix. J. Cell. Biol. 99:15-21

Tautz, D., and C. Pfeifle (1989) A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. Chromosoma 98:81-85

Taylor, A., H. P. Erba, G. E. Muscat, and L. Kedes (1988) Nucleotide sequence and expression of the human skeletal alpha-actin gene: evolution of functional regulatory domains. Genomics 3:323-336

Tellam, R. L., D. J. Morton, F. M. Clarke (1989) A common theme in the amino acid sequences of actin and many actin-binding proteins? Trends in Biochemical Sciences 14:130-133

Theze, N., F. J. Calzone, P. Thiebaud, R. L. Hill, R. J. Britten, and E. H. Davidson (1990) Sequences of the CyIIIa actin gene regulatory domain bound specifically by sea urchin embryo nuclear proteins. Mol. Reprod. & Dev. 25:110-122

Thiebaud, P., M. Goodstein, F. Calzone, N. Theze, R. J. Britten, and E. H. Davidson (1990) Intersecting batteries of differentially expressed genes in the early sea urchin embryos. Genes & Dev. 4:1999-2010

Tobin, S. L., P. J. Cook, and T. C. Burn (1990) Transcripts of individual *Drosophila* actin genes are differentially distributed during embryogenesis. Dev. Genet. 11:15-26

Vandekerckhove, J. and K. Weber (1984) Chordate muscle actins differ distinctly from invertebrate muscle actins. J. Mol. Biol. 152:413-426

Vandekerckhove, J. and K. Weber (1978) At least six different actins are expressed in a higher mammal: an analysis based on the amino acid sequence of the amino-terminal tryptic peptide. J. Mol. Biol. 126:783-802

Vandkerckhove, J. and K. Weber (1981) Actin typing on total cellular extracts: a highly sensitive protein-chemical procedure able to distinguish different actins. Differentiation 14:123-133

Wang, A. V. T., L. M. Angerer, G. J. Dolecki, R. Lum, G. V. L. Wang, R. Carlos, R. C. Angerer and T. Humphreys (1994) Unique pattern of embryonic expression of the sea urchin CyI actin in *Tripneustes gratilla.* personal communication

Wilt, F. H. and S. C. Benson (1988) Development of the endoskeletal spicule of the sea urchin embryo. Pp. 203-227 in J. E. Varner, ed. Self-Assembling Architecture. 46th Symp. Soc. Dev. Biol. Alan R. Liss. New York

Xiang, M., A. P. Bedard, G. M. Wessel, M. Filion, B. P. Brandhorst, and W. H. Klein (1988) Tandem duplication and divergence of a sea urchin protein belonging to the troponin C superfamily. J. Biol. Chem. 263:17173-17180

Yaffe, D., U. Nudel, Y. Mayer, and S. Neuman (1985) Highly conserved sequences in the 3' untranslated region of mRNAs coding for homologous proteins in distantly related species. Nucleic Acids Res. 13:3723-3737

Zakut, R., M. Shani, D. Givol, S. Neuman, D. Yaffe, and U. Nudel (1982) Nucleotide sequence of the rat skeletal muscle actin gene. Nature 298:857-859