
Jessica C. Kissinger, Jang-Hee Hahn, and Rudolf A. Raff

Department of Biology and Indiana Molecular Biology Institute, Indiana University, Bloomington

Camarodont sea urchins possess a rapidly evolving actin gene family whose members are expressed in distinct cell lineages in a developmentally regulated fashion. Evolutionary changes in the actin gene family of echinoids include alterations in number of family members, site of expression, and gene linkage, and a dichotomy between rapidly and slowly evolving isof orm-specific 3′ untranslated regions. We present sequence comparisons and an analysis of the actin gene family in two congeneric sea urchins that develop in radically different modes, Heliocidaris erythrogramma and H. tuberculata. The sequences of several actin genes from the related species Lytechinus variegatus are also presented. We compare the features of the Heliocidaris and Lytechinus actin genes to those of the the actin gene families of other closely related sea urchins and discuss the nature of the evolutionary changes among sea urchin actins and their relationship to developmental mode.

Introduction

There are two primary modes of development in sea urchins. The ancestral, indirect development, has been extensively studied and is the mode of development of most sea urchin species. Indirect development proceeds from a relatively small egg (generally, 60–120 μm in diameter) to a feeding pluteus larva which can feed for up to several months before settling and metamorphosing into a juvenile. The other mode of development starts from a larger egg (300–2,000 μm in diameter) and directly proceeds to a juvenile sea urchin following gastrulation (Wray and Raff 1991). For example, in Heliocidaris erythrogramma, which has a 430-μm egg, no feeding occurs during juvenile development, and metamorphosis occurs approximately 4 days postfertilization. As part of an investigation of the molecular evolutionary changes that underlie this radical shift in early development, we have defined the evolution and expression of the actin gene family in H. erythrogramma, a direct-developing sea urchin, and in its indirect-developing relatives.


Sea urchin actin genes are summarized in figures 1 and 2. The Strongylocentrotus purpuratus actin gene family was cloned and characterized by a number of researchers (Cooper and Crain 1982; Durica et al. 1988; Flytzanis, Bogosian, and Niemeyer 1989; Lee et al. 1984; Akhurst et al. 1987). Strongylocentrotus purpuratus has six functional actin genes and two pseudogenes (fig. 1). Five of the six functional actin genes express cytoplasmic isoforms that were divided into three iso types, CyI, CyII and CyIII (Lee et al. 1984). Both the CyII and CyIII types are represented by two genes each, designated a and b. The CyIIa and CyIIb 3′ untranslated regions (UTRs) are more closely related to each other (62%) than to any other 3′ UTR, as are the CyIIa and CyIIIb 3′ UTRs at 90% (Durica et al. 1988; Hahn 1991). The remaining actin isoform is expressed in muscle and is designated M. One pseudogene has been characterized, CyIIIc (Scheller et al. 1981; Lee et al. 1984), and another has been hypothesized in the CyII class (Lee et al. 1984; Akhurst et al. 1987). Two linked actin genes have been isolated from the congener S. franciscanus (Foran, Johnson, and Moore 1985), most probably CyIIa and CyIIb. A CyI ortholog has been cloned from a related urchin, Tripneustes gratilla (Wang et al. 1994), and the entire actin gene family has been cloned and characterized from Lytechinus pictus (Fang and Brandhorst 1994). This species has a total of four cytoplasmically expressed actin genes and one muscle gene. All sea urchin actin gene family members cloned thus far are closely related to each other. The cytoplasmically expressed genes have 97%–99% identity when compared to each other and the muscle isoforms share similar identities when compared to each other.

The embryonic expression patterns for each of the actin genes described above are known. Muscle isoforms are expressed exclusively in the muscle tissues of the developing larval pharynx. CyIII expression is restricted to tissues that are or will become aboral ectoderm (Cox et al. 1986; unpublished data). CyI and CyII genes are expressed in a variety of tissue types including ectoderm and endoderm. However, the patterns of their expression differ depending on the urchin species (Cox et al. 1986; Wang et al. 1994; Fang and Brandhorst 1996; unpublished data). The L. pictus actin gene LpC4 is expressed in a subset of primary mesenchyme cells

1 Present address: Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Disease, NIH, Bethesda, Maryland.
2 Present address: Department of Pathology, College of Medicine, Seoul National University, South Korea.

Key words: actin, sea urchins, 3′ untranslated region, Lytechinus, Heliocidaris.

Address for correspondence and reprints: Rudolf A. Raff, Department of Biology and Indiana Molecular Biology Institute, Indiana University, Bloomington, Indiana 47405. E-mail: rraff@bio.indiana.edu.

© 1997 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038
Evolution of Actin in Urchins

FIG. 1.—Inferred relationships among camarodont sea urchins (shown as embryos) and actin family members. Phylogeny and times of divergences are based on Raff et al. (1988); Smith, Boom, and Raff (1990); McMillan, Raff, and Palumbi (1992); and Littlewood and Smith (1995). The actin genes known for each of these species are indicated above the embryos (see text). The *L. pictus* gene designations of Fang and Brandhorst (1994) are given, along with their inferred identities in the nomenclature of Lee et al. (1984) (in parentheses). Evolutionary changes are mapped on the phylogeny: (1) CyI gene duplication and gene conversion among *Lytechinus* CyI 3' UTRs; (2) duplication of CyII and CyIII genes in *Strongylocentrotus*; (3) reduction in numbers of cytoplasmic actin genes expressed in early development in *Heliocidaris*; (4) loss of CyII embryonic expression in *H. tuberculata*; (5) evolution of direct development; (6) loss of CyIII gene expression and conversion to pseudogene in *H. erythrogramma*. Embryonic expression data are for *Heliocidaris* species unpublished.

believed to be involved in skeletogenesis (Fang and Brandhorst 1996).

We undertook this study of actin genes in *H. erythrogramma* because of the rapid and radical evolutionary changes its development underwent in less than 10 Myr. The present study demonstrates that, although actin coding sequences are conserved, there has been a great deal of evolution among the actin genes of closely related sea urchins. No two sea urchin actins are identical at the amino acid level. There are changes in the size of the gene families among camarodont genera. There are changes in the number of actin genes used during development, and their sites of expression (Shott et al. 1984; Cox et al. 1986; Wang et al. 1994; Fang and Brandhorst 1996; unpublished data). All of this has occurred in a 7–35-Myr window. This is in sharp contrast to stability of actin genes over 350 Myr of vertebrate evolution. All mammals use six actin genes, four muscle isoforms and two cytoplasmic. There has been no report of a change in the tissue-specific expression of one of the chordate isoforms. Finally, although some evolutionary change in actin genes appears to be related to change in developmental mode in sea urchins, most of the evolutionary changes appear to be independent of developmental style.

**Materials and Methods**

Genomic DNA Preparation

Genomic DNA was purified from *H. erythrogramma* and *H. tuberculata* sperm. Sperm (100 μl) was diluted with 400 μl of ASW and then mixed with 10 ml of extraction buffer (0.1 mg/ml proteinase K, 0.1% SDS, 50 mM Tris-HCL, 5 mM EDTA; pH 8.0) and incubated overnight at 55°C. The DNA was then repeatedly extracted with phenol and chloroform, dialyzed against TE (10 mM Tris-HCl, 1 mM EDTA; pH 8.0), and stored at 4°C.

Library Construction and Screening

Genomic libraries from *H. erythrogramma* and *H. tuberculata* were constructed in Lambda DASH (Stratagene). Five genomic equivalents of the *H. erythrogramma* and *H. tuberculata* genomic libraries were screened. Hybridization conditions were moderate at 38°C in a solution that was 50% deionized formamide, 5 X Denhardt's, 5 X SSC, 0.1% SDS, and 100 pg/ml denatured salmon sperm.

cDNA Screening and Subtractions

Two unamplified Lambda ZAP cDNA libraries were screened. The first represented pooled mRNA from midgastrula through prism stages of *H. tuberculata*, and the second pooled mRNA from midgastrula through 44-h-old stages of *H. erythrogramma*. An aliquot of 500,000 PFU from each library was plated, and triplicate sets of lifts were made. The first set of lifts from each library was probed with a mixture of gene-specific probes (CyI, CyII, CyIII, and M 3' UTR sequences) under stringent conditions (42°C). The second set of lifts from each library was probed with fourth exon sequence at moderate stringency (38°C) in 50% formamide. Autoradiographs were compared and any plaques that were identified with the actin coding probe that failed to be detected with the isoform-specific probes were further characterized. A lambda gt10 library containing cDNA sequences derived from late (57 h) *H. erythrogramma* larvae was also screened.
3' RACE (Rapid Amplification of cDNA Ends)

RNA yielded a new actin gene that was expressed in gastrula stage RNA (later identified as HtCyIII). Subtractions with unamplified cDNA libraries made with RNA pooled from several developmental stages did not yield any new genes. The few actin clones recovered from the screen were missing 3' UTR sequences. Expression of CyII in *H. tuberculata* was not detected by 3' RACE, by cDNA library screening, or by probing a RNA gel blot containing RNA from successive developmental stages with a HeCyII 3' UTR probe (data not shown). However, a faint signal was detected on genomic Southern blots (fig. 3a). Genomic copies of HtCyII and HeCyII were subsequently cloned by PCR from genomic DNA using a universal actin forward primer for the 5' end of the gene and two different CyII 3' UTR primers designed to be complementary to the HeCyII 3' UTR cDNA sequence. The HtCyII gene contains a few unique amino acid substitutions but otherwise is structurally intact and contains an open reading frame of the appropriate length. The two HeCyII genomic clones are identical at the amino acid level but differ slightly in nucleotide sequence. Both are identical in the second intron and the proximal 85 nucleotides of UTR. The remaining 3' UTR sequence is similar to the HeCyII cDNA clone for another 40 nucleotides, but then diverges.

To determine the number of actin genes in the *Heliocidaris* species and their possible linkage, CHEF and genomic Southern blots were each probed with coding regions and gene-specific probes. Both analyses show four prominent bands when probed with coding sequence from the fourth exon, and two or three faint bands, depending on the restriction enzyme used. Each of the four prominent bands can be identified by hybridization with a gene-specific probe. The remaining bands have not been identified, and their faint hybridization signal suggests that they are divergent (fig. 3).

In both *Heliocidaris* species, the CyI and CyII genes appear to be linked. When large fragments are analyzed by CHEF electrophoresis and probed with 3' UTR-specific probes using two different restriction enzymes, the same bands are detected with CyI and CyII probes (fig. 3b and d). The appearance of two bands on the CHEF blot is the result of allelic variation. Single bands for each isoform are detected on the genomic Southern blot which contained DNA from another individual. The smallest band that contains signals for both isoforms is 300 kb in size. No genomic lambda clone has ever been found to contain both genes, suggesting that they are at least 8 kb apart.

### Sequence Comparisons and Gene Identities

The coding regions of the functional *Heliocidaris* cytoplasmic actin genes are 1131 nt in length and encode putative polypeptides 375 amino acids in length. The genes are highly conserved at the nucleotide level. The majority of the observed changes occur in third positions of the reading frame and in a region of the fourth exon.

The traditional naming standard set for *S. purpuratus* using 3' UTR sequences runs into difficulties in

---

**Table 1**

Identification and Origins of Actin Clones

<table>
<thead>
<tr>
<th>Species</th>
<th>Gene Name</th>
<th>Clone Type(s)</th>
<th>Library Source/Isolation Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Heliocidaris erythrogramma</em></td>
<td>HeCyI</td>
<td>Genomic</td>
<td>λ DASH</td>
</tr>
<tr>
<td></td>
<td>HeCyII</td>
<td>cDNA</td>
<td>λt10, AZAP, 3' RACE</td>
</tr>
<tr>
<td></td>
<td>HeCyII</td>
<td>Genomic</td>
<td>PCR (two nonidentical clones G1 and G2 were isolated with different PCR primers from genomic DNA)</td>
</tr>
<tr>
<td></td>
<td>HeCyII</td>
<td>cDNA</td>
<td>λt10, 3' RACE</td>
</tr>
<tr>
<td></td>
<td>HeCyI</td>
<td>Genomic</td>
<td>λ DASH</td>
</tr>
<tr>
<td><em>Heliocidaris tuberculata</em></td>
<td>HtCyI</td>
<td>Genomic</td>
<td>λ DASH</td>
</tr>
<tr>
<td></td>
<td>HtCyII</td>
<td>cDNA</td>
<td>AZAP, 3' RACE</td>
</tr>
<tr>
<td></td>
<td>HtCyII</td>
<td>Genomic</td>
<td>PCR from genomic DNA</td>
</tr>
<tr>
<td></td>
<td>HtCyII</td>
<td>cDNA</td>
<td>λ DASH</td>
</tr>
<tr>
<td></td>
<td>HtCyII</td>
<td>Genomic</td>
<td>AZAP, 3' RACE</td>
</tr>
<tr>
<td></td>
<td>HtCyII</td>
<td>Genomic</td>
<td>λ DASH</td>
</tr>
<tr>
<td><em>Lytechinus variegatus</em></td>
<td>LvC1(CyI)</td>
<td>Genomic</td>
<td>PCR from genomic EMBL3 clone LV33*</td>
</tr>
<tr>
<td></td>
<td>LvC4</td>
<td>Genomic</td>
<td>PCR from genomic EMBL3 clone LV47*</td>
</tr>
<tr>
<td></td>
<td>LvC2(CyIII)</td>
<td>Genomic</td>
<td>PCR from genomic EMBL3 clone LV52*</td>
</tr>
</tbody>
</table>

**Note:** Exact PCR primers and cloning details are described in the Materials and Methods and Results sections of the paper. Asterisks indicate genomic clones provided by Roger Anderson.

---

**Fig. 2.—**Schematic of sea urchin cytoplasmic actin gene structure and the actin gene probes used in this study. Exons are boxed in. Translated areas are shaded in black. The first intron interrupts the 5' UTR in the cytoplasmic isoforms. The muscle isoforms have two additional introns (Crain et al. 1987). Amino acid numbers are indicated above the beginning of each translated exon. UTR = untranslated region. Species names in parentheses indicate the species of origin for the probe fragment.
Fig. 3.—CHEF linkage and genomic southern analyses. *a*, *Helioicidaris erythrogramma* genomic Southern. The probes, as described in figure 2, are indicated above each lane. *b*, *Helioicidaris erythrogramma* CHEF Southern. N, Not I; S, Sfi I. Probes are indicated above each pair of lanes. Positions and sizes of selected yeast chromosome markers, in kb, are listed on the left. *c* and *d*, as in *a* and *b* except the species being probed is *H. tuberculata*.

other species. The *Helioicidaris* species contain genes that do not have conserved 3′ UTRs previously identified in *S. purpuratus*, and the *Lytechinus* species contain yet an additional novel 3′ UTR, and one actin UTR (a CyI type) is represented twice. There are two other ways by which the actin genes may be classified. The first uses a small 15-amino-acid variable region located in the middle of the fourth cytoplasmic exon. The second method is to compare tissue-specific sites of expression. Neither of these approaches is entirely satisfactory. The number of amino acid differences is small in some cases, and phylogenetic analyses do not always support identifications (see below).

Using the site of expression is complicated by gene duplications, possible evolutionary changes in sites of expression, and issues of whether there is a need for particular isoforms in certain tissues. We have chosen to identify the genes according to their variable amino acids. When possible, other information, such as flanking sequence and introns, has been used (Hahn 1991; Hahn, Kissinger, and Raff 1995). Naming the genes is also problematic. Two naming systems are currently in use.
The first, established by Lee et al. (1984), uses the genus and species designation followed by Cy or M and a Roman numeral to designate the gene type as defined by 3' UTR sequence. The second system appeared after genes were identified whose orthology to existing genes could not be definitively ascertained. This system uses genus and species followed by C or M and an Arabic numeral (Fang and Brandhorst 1994). We have named the *Heliociduris* actin genes according to the *S. purpuratus* model and the *L. variegatus* genes after their *L. pictus* homologs. For clarity, presumed *Lytechinus* gene identities according to the *S. purpuratus* model (when possible) are also provided in parentheses.

Figure 4 shows an alignment of the divergent amino acids in the region from aa204 to aa375 in sea urchin actins. On the basis of relatively few amino acid differences, these genes can be sorted into groups. In general, the gene types appear to be more conserved between species than within species, indicating that the divergence of these genes predates the divergence of the species. The CyIII family members are different from the other cytoplasmically expressed genes in that they share a few amino acids characteristic of muscle actins. The *H. tuberculata* CyIII and the *H. erythrogramma* pseudogene are also similar, but the pseudogene has accumulated a number of changes that are shared with no other gene (a 4-bp insertion (TAAT) that inserts a stop codon in the reading frame) (data not shown). The distinction between CyI and CyII types is the most difficult, since it rests on only a few amino acid positions. The *H. erythrogramma* CyI and CyII isoforms differ by only three amino acids in the available sequence.

The actin gene coding regions are highly conserved, but the 3' UTRs are less so. Figure 5 shows a comparison of all known cytoplasmic sea urchin 3' UTR sequences. UTR sequences from the different actin types are not easily aligned with each other, so each type is presented separately. The CyI 3' UTR appears to be the most widely conserved between species. Sequences from *S. purpuratus* and *T. gratilla* species, which diverged >35 MYA, are easily aligned. Such extreme conservation of an untranslated region suggests that this

<table>
<thead>
<tr>
<th>Actin</th>
<th>3'UTR Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpCyI</td>
<td>CyI</td>
</tr>
<tr>
<td>ReCyI</td>
<td>CyI</td>
</tr>
<tr>
<td>HwCyI</td>
<td>CyI</td>
</tr>
<tr>
<td>TgCyI</td>
<td>CyI</td>
</tr>
<tr>
<td>LvCl</td>
<td>CyI</td>
</tr>
<tr>
<td>LpCl</td>
<td>CyI</td>
</tr>
<tr>
<td>LpC3</td>
<td>C3</td>
</tr>
<tr>
<td>SpCyIL</td>
<td>CyIIa</td>
</tr>
<tr>
<td>SpCyIB</td>
<td>CyIIb</td>
</tr>
<tr>
<td>Sfa-151</td>
<td>CyIIa</td>
</tr>
<tr>
<td>Sfa-152</td>
<td>CyIIb</td>
</tr>
<tr>
<td>ReCyII</td>
<td>CyII</td>
</tr>
<tr>
<td>ReCyII9</td>
<td>CyII</td>
</tr>
<tr>
<td>ReCyII9</td>
<td>CyII</td>
</tr>
<tr>
<td>ReCyII9</td>
<td>CyII</td>
</tr>
<tr>
<td>SpCyII6</td>
<td>CyIIa</td>
</tr>
<tr>
<td>SpCyII6</td>
<td>CyIIb</td>
</tr>
<tr>
<td>HapsCyIII</td>
<td>CyIIb</td>
</tr>
<tr>
<td>HapsCyIII</td>
<td>CyIIb</td>
</tr>
<tr>
<td>LcCyII</td>
<td>CyI</td>
</tr>
<tr>
<td>LpC3</td>
<td>CyI</td>
</tr>
<tr>
<td>LpC4</td>
<td>CyI</td>
</tr>
<tr>
<td>LpC4</td>
<td>CyI</td>
</tr>
<tr>
<td>SpN</td>
<td>Muscle</td>
</tr>
<tr>
<td>ReN</td>
<td>Muscle</td>
</tr>
<tr>
<td>RmN</td>
<td>Muscle</td>
</tr>
<tr>
<td>LpN</td>
<td>Muscle</td>
</tr>
</tbody>
</table>

FIG. 4.—Alignment of variable amino acids in sea urchin actin genes. Gene names are indicated in the left-hand column and the UTR type of each gene is indicated in the right-hand column. Numbers above each amino acid refer to the position of that amino acid in the 375-aa product. The initial methionine is counted as position 1. Conserved amino acids that appear to be indicative of isoform type are boxed and shaded.
3' UTRs may serve a functional role in the regulation or localization of this transcript. The 3' UTR types can be divided into six groups: CyI, CyIIa-CyIIb, CyIIIa-CyIIIb, and a new Lytechinus sequence associated with a novel coding region, a novel cytochrome C (CyIII), and one resembles S. purpuratus.

Each species possesses a gene with a CyI-type 3' UTR. However, Lytechinus species have two genes with CyI-type UTRs. This situation is not analogous to that of the S. purpuratus CyIIa-CyIIb subtypes. In S. purpuratus, the subtypes that are similar in their 3' UTRs are also similar in their inferred amino acid sequences. In Lytechinus, one of the coding regions resembles S. purpuratus CyI and one resembles S. purpuratus CyIII (fig. 4). Some actin genes have coding regions indicative of another type of gene but have 3' UTRs indicative of another, or a novel, isoform. This "switching" is especially obvious in the Lytechinus CyIII type genes, L.v.c2 and L.p.c2, which both have CyI 3' UTRs, and the Heliocidaris CyIII genes, which have unique 3' UTRs.

The HtCyIII and He pseudogene 3' UTRs are 82.7% identical. The S. purpuratus CyIIa/CyIIb 3' UTRs are 74.4% identical, and share 44.4% and 53.4% identity, respectively, with HeCyIII in the proximal 350 nt of the UTR. All of the 3' UTRs are A/T rich, and the 3' UTRs associated with the Strongylocentrotus and Heliocidaris CyIII-type genes have several polythymi-dine stretches 3-9 nt in length (fig. 4).

Each of the Heliocidaris 3' UTRs was analyzed by DOTPLOT analysis against each of the Strongylocentrotus isomers to look for any conserved stretches of nucleotides. There is a 24-nt sequence that is conserved between the HtCyIII and the SDCVIII 3' UTRs. The HtCyIII sequence is 5' TATTACTCCATGTGGT-ACAGCT-----GCATCTATATTGAGCAGTTGAATGAACTCTTTGGGGTGGTGGATTG-------GGTTTTGTTTTTACTCGATGTCCCCAGCAATGGAAGTGATCGAAGGGTACGAC

**Fig. 5.—Alignment of 3' UTR types.** Gaps have been introduced to maximize the alignment. Only the first 90–270 nucleotides are shown because of space considerations. Note that the class of genes referred to as CyIII is broken into two different groups that cannot be aligned with each other.
Phylogenetic Analysis of Actin Gene Family Members

An amino acid and nucleotide data set was constructed that included all of the sea urchin actin genes and two sea star (*Pisaster ochraceus*) (Kowbel and Smith 1989) actin sequences as outgroups. Muscle genes were included in the analysis because it has been proposed that echinoderm muscle actins evolved from their cytoplasmic counterparts (Vandekerckhove and Weber 1984; Kovilur et al. 1993). The data sets were analyzed by parsimony, distance, and maximum-likelihood analyses. In all of the phylogenetic analyses performed, none of the gene types forms a strongly supported monophyletic grouping other than the muscle forms, which do so regardless of whether third positions are included. When third positions are ignored or amino acids are used, the CyIII-type genes either are monophyletic or form a paraphyletic group outside of the CyI/CyII group of genes. However, bootstrap support is low (34% on the amino acid tree, 22% by nucleotide parsimony) (data not shown). An examination of this node on the amino acid parsimony trees shows that four characters unite this node, one with a consistency index (CI) of 1.0, two with a CI of 0.75, and one with a CI of 0.5. Most of the other nodes in the tree are united only by a single character with a CI of 0.5 or less (fig. 5a). Characters with a CI of 1.0 have no homoplasy (convergent evolution, or character duplication in another lineage) and are generally considered strong. CI values below 1.0 are indicative of some level of homoplasy.

The CyI gene comparisons of coding, 3' UTR, and promoter sequences indicate these genes are orthologous (Hahn 1991, Hahn, Kissing, and Raff 1995) (fig. 5), but they do not form a monophyletic group in phylogenetic analyses. *Heliocidaris* species have a slight bias toward cytosine in third positions in all genes and toward GGT as the codon for glycine, but no strong codon bias that would affect the analyses was detected. Gene conversion or recombination between isoforms within a species may be responsible for the aberrant phylogenetic result. A case of gene conversion in sea urchin actin genes has been reported for the *S. purpuratus* muscle actin (Crain et al. 1987).

An examination of the relationship of the CyI 3' UTRs to each other was made using parsimony on the proximal 205 nt (fig. 6b). It is difficult to establish a root for the 3' UTR sequences, but if a root is hypothetically placed where the arrow is located, the results are consistent with what is known about the relationships of the species and suggests that the CyI 3' UTRs attached to the genes classified as CyIIs are monophyletic. *Lytechinus* species have two genes with CyI 3' UTRs, and both were included in the analysis. The two CyI 3' UTRs that are present in *Lytechinus* group together and form the basal branch of CyI genes if a hypothetical root is placed as indicated.

The relationships and true evolutionary identity of the CyII genes are more difficult to evaluate. The number of differences in the coding regions between the CyI and CyII genes is very small. The 3' UTRs can be aligned between *S. purpuratus* and *S. franciscanus*, and weakly so with *Heliocidaris*. The *S. franciscanus* genes SfA-151 and SfA-152 appear to be CyII types based on comparisons of their coding regions, 3' UTRs, and linkage. A comparison of the HeCyII coding sequence with other actins shows that it belongs to the CyI/CyII group. It was classified as a CyII by Hahn (1991) on the basis of similarity in the 3' UTR. HeCyII is 58% similar to *SpCyIIa* in the proximal 3' UTR (Hahn 1991). Additional evidence for the identity of HeCyII comes from the CHEF linkage analyses. HeCyII is linked to HeCyI; however, it is not known if this linkage is ancient or the result of independent duplication events in both the *Strongylocentrotus* and *Heliocidaris* lineages.

![Phylogram of one of 251 trees of length 103 found during 100 random addition replicates with the third positions ignored. One hundred replicate bootstraps were performed on both data sets. Bootstrap values are indicated above each branch. Bullets indicate branches that are not present in the strict consensus tree. Sequences for LpC4, LvC4, LvC2, and LpM are partial. The characters were treated as missing.](image-url)
suggests that the *Heliocidaris* genes may be the result of independent duplication events. The *Heliocidaris* introns are easily aligned with each other, as are the *Strongylocentrotus* introns (fig. 7). Alternatively, there may be a gene conversion event among the *Heliocidaris* *CyI* and *CyII* isoforms.

**Discussion**

The highly regulated expression of the actin genes documented in *S. purpuratus* suggested two interesting hypotheses concerning actin expression in *Heliocidaris*. First, the loss by *H. erythrogramma* of the characteristic squamous epithelium (aboral ectoderm) of the pluteus and *Lytechinus pictus* (unpublished data). Second, the precise tissue-specific expression patterns of the CyIII gene expressed in that cell type might have changed. Indeed, it has become a pseudogene and is not expressed in this study support this suggestion. Our phylogenetic analyses did not, however, identify which of the cytoplasmic genes, one of which is not functional in each species, and one muscle gene for a minimum estimate of four genes. *Lytechinus pictus* has a total of five actin genes, and all are functional (Fang and Brandhorst 1994).

Fang and Brandhorst (1994) suggested that an ancestral cytoplasmic actin gene duplicated and diverged to form the coding regions that are referred to as *CyI* and *CyIII*. Then, by additional gene duplications, divergence, and the addition or switching of 3' UTRs, the extant genes can be accounted for. The data presented in this study support this suggestion. Our phylogenetic analyses did not, however, identify which of the cytoplasmic lineages is the oldest, the *CyI/CyII* or the *CyIII*. Addition of the novel *Lp* and *Lv* C4 sequences did not suggest an origin for these genes and resulted in an unresolved trichotomy with the other two cytoplasmic lineages.

The *CyI* and *CyII* genes almost certainly arose via a gene duplication. The duplication event leading to these two gene types apparently took place before the split of *Strongylocentrotus* and *Heliocidaris*, since the species examined in each of these genera contain these genes (fig. 1). However, independent gene duplications in each lineage cannot be ruled out. Two lines of evidence support the hypothesis that the *CyI/CyII* divergence occurred before the *Strongylocentrotus/Heliocidaris* divergence. First, the proximal region of the 3' UTR of the *HeCyII* gene shares 58% identity with the *SpCyIIa* 3' UTR. Independent duplication events within each lineage would be an extremely unlikely occurrence. Second, the *CyI* genes, which are clearly orthologous by all regions compared (Hahn, Kissinger, and Raff 1995), do not form a monophyletic group in phylogenetic analyses either. This result suggests the possibility that these analyses are subject to fault with the given data set and
running conditions or, more likely, that the phylogenetic analyses are detecting gene conversions within the different genera. For example, one of the HeCyI introns is more easily aligned with HeCyII than it is with other CyI genes, indicating that a conversion event is likely (fig. 7).

Two CyII genes, CyIIa and CyIIb, are present in *S. purpuratus*. They are linked to each other, and to CyI. The duplication event that created CyIIb most likely occurred within the *Strongylocentrotus* lineage, as multiple CyII isoforms are not detected in *Heliocidaris*. Only one CyIII gene has been isolated from *Heliocidaris*, consistent with a duplication event that led to CyIIa and CyIIb within the *Strongylocentrotus* lineage.

In *Strongylocentrotus*, the SpCyIIIb 3' UTR is highly unusual. The proximal 3'-UTR consists of a 78-nt direct repeat of the 3' end of the coding sequence. This is followed by a 105-nt stretch of sequence that is itself directly repeated distally (Flytzanis, Bogosian, and Nie-meyer 1989). None of these features is detected in either *Heliocidaris* CyIII UTR. In each of the camoradont genera which have been examined, the CyIII genes have divergent 3' UTRs. The divergent 3' UTRs are conserved within a genus, but not between genera. Phylogenetic analyses of the coding region at the nucleotide level did not place the CyIII genes in a monophyletic group when all positions were included, but did when amino acids were used. When third positions are excluded the CyIIIs are monophyletic or show a paraphyletic grouping outside of the CyI/CyII cluster. As these genes are all expressed in the same tissue type in the pluteus (Cox et al. 1986; Fang and Brandhorst 1996; unpublished data), we would argue that these genes are orthologous among camoradont sea urchin lineages.

We present sequence for three orthologous genes in *L. variegatus*. The *Lytechinus* gene family is interesting because it contains two genes with CyI 3' UTRs that belong to different gene types based on their coding regions and expression patterns. Additionally, the two remaining gene types have unique 3' UTRs that have not been observed in other echinoid species. One of the *L. pictus* genes appears to be a new cytoplasmic type, as it has a unique coding sequence and a unique pattern of expression (Fang and Brandhorst 1994, 1996) (fig. 5). The available linkage data for *Lytechinus* species (no genomic clone contains two genes) suggests that there may be no close linkage among actin genes in this species. Thus, it is difficult to determine if the two genes with the CyI 3' UTRs arose by tandem duplication followed by divergence in the coding region, or if some other mechanism was involved. Gene conversion or recombination among related actin genes provides mechanisms by which actin 3' UTRs may be evolving.

A general comparison of the actin 3' UTRs in sea urchins reveals two very interesting trends. There are the CyI 3' UTRs which have changed very little between the genera (Hahn, Kissinger, and Raff 1995), and there are the CyIII 3' UTRs, which appear to be different in each genus examined thus far. It has become clear that 3' UTRs play a role in a number of regulatory events, including mRNA localization, signaling for pol-
Thus, several distinct levels of selection appear to be operating on the different actin isoforms.

Because many of the actin genes are linked in sea urchins, and their origins are known, sea urchins provide a useful system for the study of actin evolution. The duplication and divergence events could then lead to the observed differences in expression patterns (Cox et al. 1986; Wang et al. 1994; Fang and Brandhorst 1996; unpublished data). Obviously, this “shuffling” of expression patterns would be less likely to occur if there were functional differences between the actin gene types. It has been argued that there are real functional differences between the chordate gene types because of their high degree of conservation in structure and expression. However, little is currently known about the functional constraints of echinoderm actin genes (Rubenstein 1990). Evolution of the actin gene family is radically different between chordates and echinoids. These genes are thus operating under very different constraints in these two deuterostome taxa.

Acknowledgments

We thank Roger Anderson for making available Lytechinus variegatus actin clones and partial sequences, Eric Davidson for S. purpuratus CyIII UTR DNA, Eric Haag for providing H. erythrogramma and H. tuberculata unamplified cDNA libraries, and Michael Frohman for RACE primers. We also thank Kelley Rogers for critical reading of the manuscript. This study was supported by NIH grant HD21337 to R.A.R. J.C.K. was supported by an NIH genetics training grant.

LITERATURE CITED


CLAUDIA KAPPEN, reviewing editor

Accepted February 28, 1997