

# Egg-Laying Hormone Genes of *Aplysia*: Evolution of the ELH Gene Family

John R. Nambu and Richard H. Scheller

Department of Biological Sciences, Stanford University, Stanford, California 94305

Evolution of the egg-laying hormone (ELH) gene family was examined in the genus *Aplysia* using genomic Southern blotting, gene cloning, and immunocytochemical techniques to identify and characterize homologous sequences. Most of the species examined have fewer than the 4-5 ELH-related genes present in the *A. californica* genome (Mahon et al., 1985; Scheller et al., 1983). In *A. parvula* there are 2 ELH genes, and unlike *A. californica*, no sequences were found to encode the A or B peptides. The 2 *A. parvula* ELH genes share at least 90% DNA sequence homology, while the homology between the *A. parvula* and *A. californica* ELH genes is 71%. The structural organization of the *A. parvula* ELH precursor is quite similar to the ELH precursor of *A. californica*, with all but one of the potential proteolytic cleavage sites conserved. The overall amino acid homology between the *A. parvula* and *A. californica* ELH precursors is 66%; however, the alpha and beta bag cell peptides, as well as ELH, are more highly conserved, suggesting that these peptides have important physiological and behavioral roles within both *Aplysia* species. Immunocytochemical studies indicate that the *A. parvula* ELH genes are expressed in 2 bag-cell-like clusters of about 40 neurons each in the abdominal ganglion. There does not appear to be an atrial gland in *A. parvula*; however, ELH-immunoreactive peripheral neurons and their processes are observed along the perimeter of the large hermaphroditic duct.

The egg-laying behavior of the marine gastropod *Aplysia californica* constitutes a fixed action pattern consisting of several stereotypic components, including cessation of locomotion, inhibition of feeding, increased respiratory pumping, head waving, and egg deposition (Arch and Smock, 1977). The neuroendocrine basis for this behavior pattern has been widely studied, and several bioactive peptides, including the closely related A and B peptides and the egg-laying hormone (ELH), have been implicated in mediating the egg-laying response (Chiu et al., 1979; Heller et al., 1980). ELH is produced principally by 2 clusters of 800 electrically interconnected neurons, the bag cells, located at the junction of the pleural-abdominal connectives

and the abdominal ganglion, although it is also synthesized in a small network of neurons throughout the CNS (Chiu and Strumwasser, 1984; McAllister et al., 1983). ELH alters the electrical activities of several neurons in the abdominal, buccal, and pedal ganglia and also elicits the release of eggs from isolated ovotestis tissue (Mayeri et al., 1985; Ram, 1983a; Rothman et al., 1983a; Stuart and Strumwasser, 1980). The A and B peptides are synthesized by the atrial gland, an exocrine organ that forms a continuous lumen with the large hermaphroditic duct at the base of the gonopore. These peptides will induce activation of the bag cell neurons, a prerequisite for the egg-laying response (Heller et al., 1980; Pinsker and Dudek, 1977).

Molecular genetic studies have revealed that a small multigene family encodes the precursors for ELH and the A and B peptides in *A. californica* (Scheller et al., 1982, 1983). This gene family consists of 4-5 discrete members, and DNA sequence analysis indicates that the genes encoding ELH and the A and B peptides all share approximately 90% sequence homology. Divergence within the coding regions of the 3 genes has resulted in distinct polyprotein precursors from which different sets of peptides are proteolytically cleaved. The ELH precursor contains several peptides in addition to ELH, including the acidic peptide, and 3 small, related peptides—the alpha, beta, and gamma bag cell peptides.

The ELH gene family probably arose through initial intragenic duplications that produced a polyprotein gene and subsequent intergenic duplications involving larger regions of the genome that resulted in the formation of a gene family (Mahon and Scheller, 1983; Scheller et al., 1983). The different members of this gene family have since diverged in structure and function, now displaying temporal and tissue specific patterns of expression. Multigene families are common units of genomic organization that may serve to provide high biosynthetic levels of specific products such as rRNAs or histones, or, as is probably the case with the ELH gene family, permit related genes to exhibit novel functions or patterns of expression without a concomitant loss of the original gene function (Hood et al., 1975; Long and Dawid, 1980).

The *Aplysia* genus is comprised of some 35 species grouped into 5 distinct subgenera (Table 1). There exists a wide range of anatomical differences, and specific behavioral programs such as swimming, inking, or burrowing are expressed in some species but not others. The phylogenetic history of the *Aplysia* genus is not well established; however, the richness in species diversity suggests that there may exist corresponding molecular and behavioral variability with respect to egg-laying processes. In order to study the evolution of the ELH gene family, we utilized genomic DNA blot analysis to examine the conservation of the ELH gene family in 10 different *Aplysia* species, including representatives from 4 out of the 5 different subgenera (Table 1). A more detailed analysis of gene and protein evolution was accomplished through the isolation and structural characterization of the ELH-related genes from *A. parvula*.

Received Oct. 7, 1985; revised Dec. 12, 1985; accepted Dec. 30, 1985.

We wish to thank Marina Picciotto, Thane Kreiner, and Ron Taussig for technical assistance with various aspects of this project. We are very grateful to the many colleagues throughout the world who provided us with live animals or tissue sources, including C. Baxter, J. Blankenship, C. Cazaux, M. Moulins, J. Ram, A. Susswein, and E. Tobach. We are especially indebted to Dr. M. Switzer-Dunlap for continued support throughout the project. We also appreciate the critical reading of this manuscript by A. C. Mahon, P. A. Nambu, and W. Sossin, and we thank S. Painter for providing us with preprints of her work prior to publication. J.R.N. is an NSF predoctoral fellow and R.H.S. is a Klingenstein Fellow in the Neurosciences, a McKnight Foundation Scholar, and the recipient of a Presidential Young Investigator Award from the NSF. This work is supported by grants from the NIH.

Correspondence should be addressed to Dr. Richard H. Scheller at the above address.

Copyright © 1986 Society for Neuroscience 0270-6474/86/072026-11\$02.00/0

**Table 1.** The 5 different subgenera of the *Aplysia* genus and the corresponding species

Subgenus	Species	
<i>Pruvotaplysia</i>	<i>parvula</i>	
	<i>punctata</i>	
<i>Neoaplysia</i>	<i>californica</i>	
<i>Varria</i>	<i>brasiliiana</i>	<i>maculata</i>
	<i>cervina</i>	<i>morio</i>
	<i>cornigera</i>	<i>oculifera</i>
	<i>cronullae</i>	<i>pulmonica</i>
	<i>dactylomela</i>	<i>rehderi</i>
	<i>denisoni</i>	<i>reticulata</i>
	<i>extraordinaria</i>	<i>robertsi</i>
	<i>fasciata</i>	<i>sagamiana</i>
	<i>gigantea</i>	<i>sowerbyi</i>
	<i>gracilis</i>	<i>syneyensis</i>
	<i>inca</i>	<i>willcoxi</i>
	<i>keradreni</i>	<i>winneba</i>
	<i>kurodai</i>	
<i>Aplysia</i>	<i>cedrosensis</i>	
	<i>depilans</i>	
	<i>dura</i>	
	<i>juliana</i>	
	<i>nigra</i>	
	<i>vaccaria</i>	
<i>Phycophyla</i>	<i>euchlora</i>	

Specimens examined for ELH homologous sequences by genomic Southern blot analysis are boldface. Data from Kandel (1979).

## Materials and Methods

### DNA extraction

Frozen whole body, muscle, ovotestis, or sperm was utilized as a tissue source to extract genomic DNA from the different *Aplysia* specimens as described by Maniatis et al. (1982). Tissues from the various *Aplysia* species were supplied by the following persons: *A. californica*, M. Morris/Sea Life Supply; *A. brasiliiana*, J. Blankenship; *A. dactylomela*, *A. juliana*, and *A. parvula*, M. Switzer-Dunlap; *A. oculifera*, *A. fasciata*, and *A. depilans*, J. Ram and A. Susswein; *A. punctata*, C. Cazaux and M. Moulins; *A. vaccaria*, Pacific Biomarine.

### Genome blots

Genomic Southern blots were performed as described in Maniatis et al. (1982) using a hybridization and wash temperature of 60°C. Hybridization probes were derived from either *A. californica* ELH cDNA or *A. parvula* ELH genomic DNA clones.

### Isolation of *A. parvula* ELH clones

A genomic DNA library was constructed from the DNA of a single *A. parvula* specimen using the lambda phage cloning vector  $\lambda$ J1 (Mullins et al., 1984) as outlined in Maniatis et al. (1982). Approximately 500,000 individual recombinants were screened from the unamplified library ( $2.4 \times 10^6$  total recombinants) by the technique of Benton and Davis (1977) using the same *A. californica* ELH cDNA probe used in the genome blot experiments.

### Characterization of *A. parvula* ELH clones

The *A. parvula* ELH clones were characterized by restriction endonuclease mapping and blotting procedures. Hind III fragments found to contain the regions of ELH homology were subcloned into the plasmid vector SP 65 (Melton et al., 1984) and transformed into *E. coli* strain MM 294. Recombinant plasmid DNA was prepared according to the

method of Maniatis et al. (1982) and used as substrate for DNA sequence analysis.

### DNA sequencing

*A. parvula* ELH subclones 1 and 2 were subjected to direct nucleotide sequence analysis using the chemical modification and cleavage procedure of Maxam and Gilbert (1980). The products were separated on 0.35 mm thick 20 and 6% polyacrylamide gels (Sanger and Coulson, 1977) and analyzed by autoradiography. Sequencing strategies are presented in Figure 3.

### Immunocytochemistry

Frozen tissue sections of *A. parvula* abdominal ganglia and large hermaphroditic duct were stained with rabbit anti-ELH, rabbit anti-A peptide, and preimmune rabbit sera as described in Kreiner et al. (1984). (We thank Rudy Kaldany and Wayne Sossin for use of the ELH and A peptide antisera.)

### Amino acid and nucleotide homology plots

Comparisons of the amino acid and nucleotide homologies between the *A. californica* and *A. parvula* ELH precursors, as presented in Figure 4, were determined as follows. For the amino acid homology plot, beginning at position no. 5, a run of 10 consecutive residues (nos. 1–10) was compared between the 2 precursors and the percentage homology (the number of conserved amino acids  $\times 10$ ) determined. The determination was then repeated for the next 10 consecutive residues (from nos. 2–11) to define the homology value at residue no. 6. This procedure was repeated until a point 5 residues from the C-terminal end of the precursors, and then the points were graphed, with the percentage homology plotted along the y-axis and the position along the precursor serving as the x-axis. For the nucleotide homology plot, a similar strategy was used, except that at each amino acid position, the corresponding 30 consecutive nucleotides (coding for each of the 10 consecutive amino acids) were compared. The percentage nucleotide homology was thus determined for each amino acid position and plotted similarly to the amino acid homology plot. Insertions–deletions between the 2 precursor sequences were ignored for this calculation.

### Analysis of *A. californica* and *A. parvula* ELH DNA sequences

The percentage nucleotide sequence homology between various combinations of ELH-related genes (Table 2) was determined by comparing the shared sequences between 2 ELH-related genes. Insertions–deletions between sequences were omitted from calculations of homology within the coding regions but were treated as mismatches in 5' and 3' flanking regions. The 5' and 3' untranslated regions were defined by sequences found on *A. californica* cDNA clones encoding ELH or A peptide precursors as previously described by Mahon et al. (1985). Analysis of replacement and silent site substitutions were performed as described in Perler et al. (1980).

## Results

### ELH gene family in the genus *Aplysia*

Genomic Southern blot hybridization was utilized to characterize ELH-related sequences in 10 different *Aplysia* species (Table 1). Genomic DNA from these specimens was digested with 3 different restriction enzymes—Eco RI, Hind III, and Pst I. Eco RI and Hind III do not cleave within the *A. californica* ELH cDNA fragment used as probe; however, there is a single Pst I recognition site within the ELH cDNA probe sequence, and this restriction enzyme site is conserved within the corresponding region of the A and B peptide genes (Mahon et al., 1985; Scheller et al., 1983). ELH homologous sequences were detected in all *Aplysia* species examined, and the pattern of hybridizing restriction enzyme fragments from 8 of these animals is presented in Figure 1. Two conclusions can be drawn from these data. First, there is variability in the signal strength produced by different specimens. Thus, *A. brasiliiana* and *A. fasciata* (not shown) both gave very strong signals, implying a high level of DNA sequence homology with the *A. californica* ELH probe, while other specimens, such as *A. parvula*, yielded

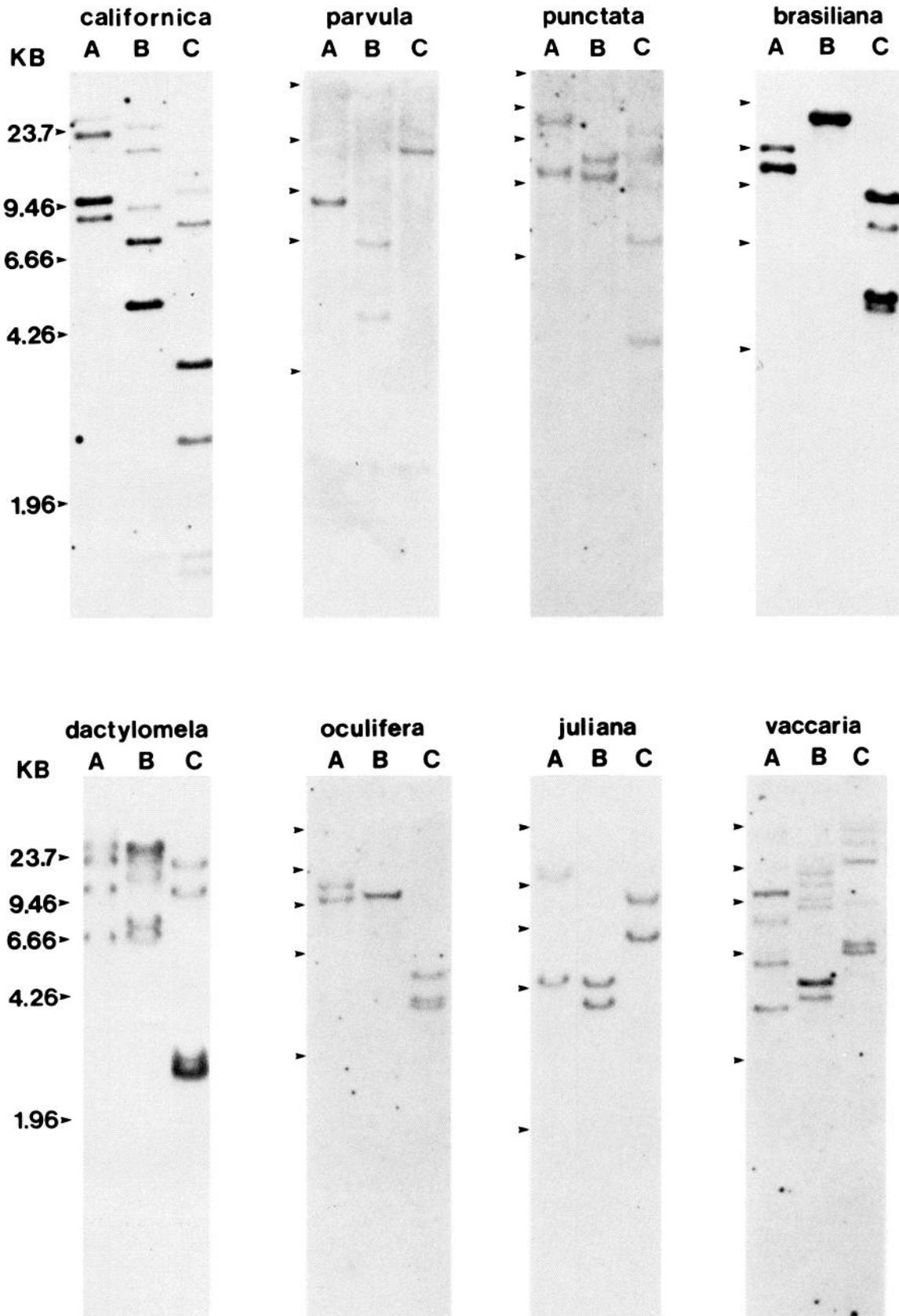
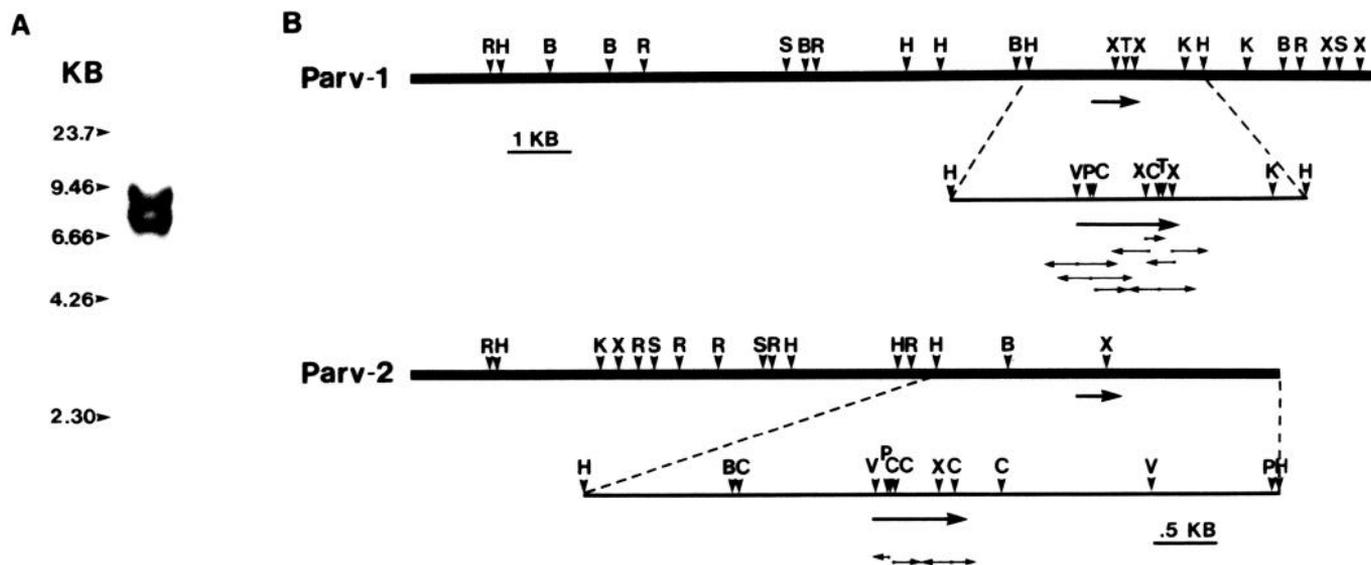


Figure 1. Genomic Southern blot analysis of the ELH genes in 8 different *Aplysia* species. Genomic DNA, 10  $\mu$ g, from a single specimen of each *Aplysia* species was digested to completion with the restriction enzymes Eco RI (A), Hind III (B), or Pst I (C), electrophoresed on 0.7% agarose gels, and transferred to nitrocellulose filters. A *Hinc* II/Eco RI restriction fragment, encompassing the coding region of the *A. californica* ELH cDNA clone J (Mahon et al., 1985), was labeled with  $^{32}$ P via nick-translation and used as a probe in hybridization analysis for ELH homologous sequences. The resulting Southern blot pattern is presented. Positions of molecular weight size standards are given at left and by arrows. KB, kilobase pairs.



**Figure 2.** *A*, Genomic Southern blot analysis of *A. parvula* DNA using cloned *A. parvula* ELH probe. Ten micrograms of the same *A. parvula* DNA from which the recombinant library was constructed were digested to completion with Eco RI, electrophoresed on 0.7% agarose gel, and transferred to nitrocellulose. The blot was hybridized to a nick-translated  $^{32}$ P-labeled Hinc II/Hinc II restriction fragment derived from parv-1 subclone. This fragment is wholly contained within the coding region of the *A. parvula* ELH gene. The resulting Southern blot pattern is shown. *B*, Restriction enzyme maps and organization of the *A. parvula* ELH genomic clones and subclones. Positions of the various restriction enzyme cleavage sites were deduced using single and double digests of phage clone and plasmid subclone DNAs. The position of the ELH genes contained in parv-1 and parv-2 clones as well as the direction of transcription was determined via Southern blot and DNA sequence analysis and is indicated by large arrows. The strategy used for DNA sequencing is indicated by small arrows. Restriction enzyme designations are as follows: B, Bgl II; C, Hinc II; H, Hind III; K, Kpn I; P, Pst I; S, Sal I; T, Sst I; V, Pvu II; R, Eco RI; X, Xho I.

much weaker signals, implying lower levels of DNA homology. Second, there is variability in the pattern of hybridizing bands in the DNA blots from the different *Aplysia* species. In addition to *A. californica*, both *A. dactylorella* and *A. vaccaria* yielded patterns consisting of 4–8 bands, suggesting that these species each has 4–5 ELH-related genes. However, most species, including *A. parvula* and *A. brasiliana*, appear to have only 1–2 ELH homologous sequences, as only 1–4 bands were observed in their DNA blot patterns.

#### Identification of the *A. parvula* ELH genes

To better characterize the representation of the ELH gene family within the genus *Aplysia*, ELH-related genes were isolated from an *A. parvula* genomic DNA library. In order to facilitate analysis of ELH gene copy number and reduce the possibility of allelic variability in gene structure, the library was constructed using whole-body DNA from a single *A. parvula* specimen. Seventeen clones with ELH homology were detected, and several of these were structurally characterized. Two different classes of clones were identified as containing sequences corresponding to the 2 restriction fragments with ELH homology detected in the genome blot study (Figs. 1, 2). Two representative clones, parv-1 and parv-2, were mapped in detail, and regions containing the ELH homologous sequences were subcloned into a plasmid vector for further analysis (Fig. 2). The restriction enzyme maps of these clones are distinct, demonstrating that in *A. parvula* there are 2 nonallelic ELH genes per haploid genome.

To verify that the 2 isolated genes encompass the complete ELH gene family in *A. parvula*, a genome blot was performed using an aliquot of the same *A. parvula* DNA used to construct the genomic library. Two strongly hybridizing bands were observed in the products of an Eco RI restriction enzyme digest using a probe derived from the coding region of the *A. parvula* ELH gene (Fig. 2). An identical, though less intense pattern was observed using a heterologous *A. californica* ELH cDNA clone

as the probe (data not shown). Thus, there are no additional ELH-related sequences detected when using the homologous *A. parvula* ELH probe. The blot pattern differs slightly from that presented for *A. parvula* in Figure 1 because of a polymorphic difference in the size of the Eco RI bands containing the ELH genes. Similar restriction enzyme site polymorphisms have been previously described in the *A. californica* ELH gene family (Mahon and Scheller, 1983; Scheller et al., 1982).

#### Analysis of ELH gene sequences

To characterize the *A. parvula* ELH genes further, regions of the 2 subclones containing ELH homology were subjected to DNA sequence analysis. The DNA sequence of these regions and the corresponding amino acid sequence of the predicted proteins are presented in Figure 3. The *A. californica* ELH gene and precursor sequences are included for comparison. The 2 *A. parvula* genes are nearly identical and both encode precursor proteins that are clearly similar to the ELH precursor in *A. californica*. Surprisingly, no genes encoding precursors for the A and B peptides were found in the *A. parvula* genome.

The structural organization of the *A. parvula* ELH genes is similar to the *A. californica* ELH gene family members in that the entire coding region appears to exist on a single exon (Mahon et al., 1985; Scheller et al., 1983). An analysis of the DNA sequence conservation between ELH-related genes in *A. parvula* and *A. californica* is presented in Table 2. DNA homology is categorically higher among ELH gene family members within a species than between the interspecific ELH genes. Within the coding region, the highest levels of homology are those observed between the A and B peptide genes and the parv-1 and parv-2 ELH genes; the lowest level is between the *A. californica* and *A. parvula* ELH genes. The 5' flanking sequences of the *A. californica* and *A. parvula* ELH genes are not conserved, and the highest homology between these genes occurs in the 3' untranslated region around the AUUAAA polyadenylation addition



Table 2. Sequence homologies

Gene	% DNA			Codon substitutions						
	5'	Codon	3'	% AA	% 1	% 2	% 3	% Silent : % Replacement		
p-1/p-2	NA	95.1 (34)	87.6	96.5 (8)	20.6	5.9	73.5	20.60	(10.7)	1.93
p-1/c-E	44.1	71.1 (229)	75.3	65.9 (90)	29.6	18.9	51.5	134.3	(5.8)	23.23
c-E/c-A	95.5*	87.9 (66)	100	75.5 (44)	35.3	39.2	25.5	9.20	(0.8)	11.77
c-A/c-B	NA	96.4 (13)	100	91.1 (11)	23.1	53.8	23.1	3.65	(0.9)	4.04

Comparison of the DNA sequence homology between members of the *A. californica* ELH gene family and the *A. parvula* ELH genes. *A. californica* gene sequences are from genomic clones ELH-1 and ELH-18 (Scheller et al., 1983). Comparisons include: p-1/p-2 = parv-1 vs parv-2, ELH genes, p-1/c-E = parv-1 vs ELH-1, ELH genes, c-E/c-A = ELH-1 vs ELH-18, ELH and A peptide genes (\*, comparison with 5' untranslated region of B peptide gene), c-A vs c-B = ELH-18 vs ELH-1, A and B peptide genes. The percentage nucleotide homology is listed for coding regions as well as 5' and 3' untranslated regions where applicable. The number of DNA substitutions ( ) in the coding region is given. The percentage amino acid homology and number of amino acid substitutions ( ) is also presented for 4 combinations of ELH-related genes. Analysis of the distribution of DNA substitutions within the coding regions at the 3 codon positions is presented as the percentage of total substitutions occurring at each position. The percentage of the total number of silent sites that have been substituted and the percentage of the total number of replacement sites that have been substituted are also given. The ratio of the percentage of silent sites substituted to that of replacement sites substituted (% Silent: % Replacement) is indicated ( ).

signal (Proudfoot and Brownlee, 1976). In contrast, members of the *A. californica* ELH gene family exhibit extremely high conservation of both 5' and 3' untranslated regions.

Nucleotide substitutions within the coding region of several ELH-related genes were analyzed with respect to codon positions and silent versus replacement sites using the method of Perler et al. (1980) (Table 2). The relative distribution of DNA substitutions varies between different combinations of ELH-related genes. Most of the substitutions between the parv-1 and parv-2 ELH genes occur in third base positions and there is a 10-fold greater substitution rate at silent sites than at replacement sites. Between the parv-1 ELH sequence and the *A. californica* ELH sequence, third base substitutions again predominate, as does a 5-fold higher silent site substitution rate. The majority of the DNA substitutions between ELH genes thus do not result in amino acid changes. This pattern suggests that these genes are diverging independently and that selectively neutral substitutions (Kimura, 1983) have been fixed. However, among members of the *A. californica* ELH gene family, roughly equal rates of silent and replacement site substitutions are observed, along with a corresponding decrease in the proportion of third base changes (Table 2). Thus, as originally noted by Scheller et al. (1983), even though there are very few DNA substitutions between the A and B peptide genes, almost all of them lead to amino acid changes. Most of the substitutions between the *A. californica* ELH and A peptide genes similarly alter the amino acid sequences. This could be the result of strong positive selection acting on a few substitutions that have occurred in very recently diverged genes, with slower fixation of silent and untranslated substitutions. However, this substitution pattern is not generally observed (Miyata et al., 1980) and may indicate that these genes have not diverged independently.

#### Characterization of the *A. parvula* ELH precursors

The 2 *A. parvula* ELH precursors share 96.5% amino acid homology within the sequenced regions. The overall amino acid

homology between the *A. parvula* and the *A. californica* ELH precursors is 65.9%; however, different regions of the precursors exhibit widely varying levels of conservation. Schematic representations of the structures of the *A. parvula* ELH precursor, as well as the *A. californica* ELH and A peptide precursors, are presented in Figure 4. Additionally, the amino acid homology along the lengths of the *A. californica* and *A. parvula* ELH precursors, calculated as described in Materials and Methods, is also presented.

The ELH precursor of *A. parvula*, as well as every other precursor encoded by the ELH gene family characterized to date, has a hydrophobic signal sequence at the amino (N)-terminus which contains 2 conserved cysteine residues. Although the amino acid homology of these regions in the *A. parvula* and *A. californica* ELH precursors is only approximately 50%, the hydrophobic nature is clearly maintained. Proceeding in the carboxy (C)-terminal direction, the *A. parvula* ELH precursors both contain most of the 80 amino acid region of the *A. californica* ELH precursor that encompasses several candidate peptides, including the beta, gamma, and delta bag cell peptides (BCPs). This region is not present in either the A or B peptide precursors and therefore was probably deleted from an ancestral ELH gene prior to divergence of the A or B peptide genes. The *A. californica* pentapeptide, beta BCP, and both pairs of Lys-Arg cleavage sites are 100% conserved in the *A. parvula* ELH precursors, while the related pentapeptide, gamma BCP, has diverged at 2 positions. After this region, the amino acid homology decreases dramatically. In *A. californica*, an unrelated octapeptide sequence, delta BCP, then follows; however, in the *A. parvula* precursor, this region, including the Arg-Arg pair at the C-terminus of delta BCP, is completely divergent.

The amino acid homology increases over the next 20 residues directly preceding the alpha BCP. Alpha BCP has been shown to have a number of physiological activities on the CNS of *A. californica* (Rothman et al., 1983b), and this peptide, as well as its flanking cleavage sites, is 100% conserved between the 2

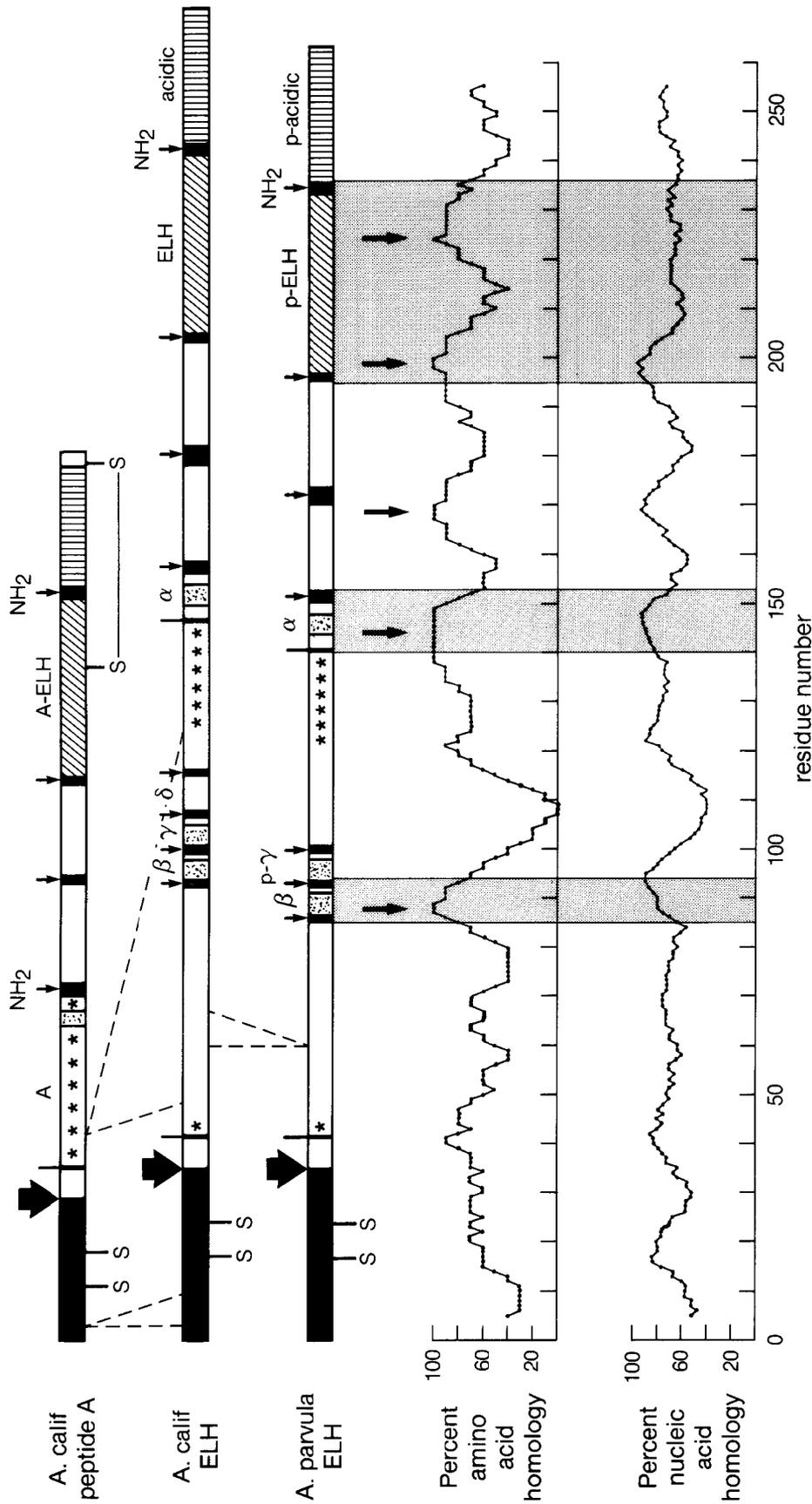


Figure 4. Polyprotein precursors containing ELH and the A peptide from *A. californica* and ELH from *A. parvula* presented schematically along with homology plots comparing the 2 different ELH precursors. The positions of identified peptides—including A peptide (+); ELH (▨); alpha, beta, gamma, and delta bag cell peptides (▩); and the acidic peptide (▧) are indicated. The signal sequence (■) and its presumed cleavage site (♣) are given. Other known or potential proteolytic cleavage sites occur at single arginine (●) or multiple basic residues (⊖). Amino acid deletions with respect to the *A. californica* ELH precursor are shown by broken lines; cysteine residues (S), disulfide bridges (S-S), and carboxy terminal amidation sites (NH<sub>2</sub>) are also indicated. Below are homology plots for amino acid and DNA sequences of *A. parvula* and *A. californica* ELH precursor proteins and genes. These plots were generated as described in Materials and Methods; y-axis, percentage homology; x-axis, amino acid position along precursor. Arrows are placed in the center of the 5 regions where the amino acid sequence homology reaches 100%. The shaded areas (▨) indicate the positions of sequences and cleavage sites for beta and alpha BCP and ELH.

## EGG-LAYING HORMONES

1    **I** **S** **I** **N** **Q** **D** **L** **K** **A** **I** **T** **D** **M** **L** **L** **T** **E** **Q** **I** **R** **E** **R** **Q** **R** **Y** **L** **A** **D** **L** **R** **Q** **R** **L** **L** **E** **K** -NH<sub>2</sub>  
 2) **I** **S** **I** **N** **Q** **D** **L** **K** **A** **I** **T** **D** **M** **L** **L** **T** **E** **Q** **I** **Q** **A** **R** **R** **R** **C** **L** **D** **A** **L** **R** **Q** **R** **L** **L** **D** **L** -NH<sub>2</sub>  
 3) **I** **S** **I** **N** **Q** **D** **L** **K** **A** **I** **A** **D** **M** **L** **I** **V** **E** **Q** **K** **Q** **E** **R** **E** **K** **Y** **L** **A** **D** **L** **R** **Q** **R** **L** **L** **N** **K** -NH<sub>2</sub>  
 4) **L** **S** **I** **T** **N** **D** **L** **R** **A** **I** **A** **D** **S** **Y** **L** **Y** **D** **Q** **H** **W** **L** **R** **E** **R** **Q** **E** **N** **L** **R** **R** **R** **F** **L** **E** **L** -NH<sub>2</sub>

Figure 5. Amino acid sequences in standard 1-letter code for ELH regions of (1) *A. californica* ELH precursor, (2) ELH region of *A. californica* A peptide precursor, (3) ELH region of *A. parvula* ELH precursor, and (4) ELH sequence from *Lymnaea stagnalis* (Joose et al., 1985). Residues that are conserved among all 4 peptides are designated by boldface and asterisk. Amino acid homology between the different ELH peptides is as follows: (1)/(2) = 78%, (1)/(3) = 78%, (1)/(4) = 44%, (2)/(3) = 67%, (2)/(4) = 42%, (3)/(4) = 42%.

species. Moving further in the C-terminal direction, the region between alpha BCP and ELH contains an Arg-Arg-Lys-Arg potential cleavage site, which is completely maintained between the precursors, suggesting that this region could be functionally important in the processing of the *A. californica* and *A. parvula* ELH precursors. ELH is flanked at its N-terminus by a Lys-Arg pair in both precursors and exhibits 78% homology, with nearly all of the mismatches occurring in the middle third of the peptide (Fig. 5). The Gly-Lys-Arg C-terminus of ELH is also conserved between the 2 species and, thus, it is probable that like *A. californica* ELH, *A. parvula* ELH is amidated at its C-terminus. Following ELH, the acidic peptide (AP) is present in both precursors. This peptide has no known physiological activities, although it may serve to facilitate packaging and/or transport of basic peptides like ELH through ionic interactions. The amino acid homology of the AP is only 56%; however, the highly acidic nature of this molecule is maintained.

#### Morphology of ELH-expressing tissues in *A. parvula*

Immunocytochemical techniques were employed to identify cells expressing ELH gene products in *A. parvula*. Sections of the abdominal ganglion and large hermaphroditic duct were assayed for immunoreactivity using rabbit antiserum directed against *A. californica* ELH and A peptide. In the abdominal ganglion, 2 ELH-immunoreactive clusters of approximately 40 neurons each were observed along the rostral margin (Fig. 7D). These cells most likely correspond to the bag cells, although in *A. californica* the bag cell neurons are far more numerous, at about 400/hemiganglion, and are usually located more rostrally in distinct clusters on the pleural-abdominal connectives. In addition, there were what appeared to be ectopic bag cells (Chiu and Strumwasser, 1981) within the ganglion that also showed ELH immunoreactivity. The sizes of both these cells and those in the bag cell clusters are 6–7  $\mu$ m. This is considerably smaller than the 40  $\mu$ m bag cell neurons of *A. californica*; however, there is up to a 1000-fold difference in mass between adult specimens of the 2 species. In *A. parvula*, neuronal processes with ELH immunoreactivity were observed in the pleural-abdominal connectives and connective tissue sheath surrounding the abdominal ganglion, a pattern of arborization similar to that observed in *A. californica* (Chiu and Strumwasser, 1981; McAllister et al., 1983). The overall organization of the *A. parvula* CNS suggests that it is perhaps the most primitive species of *Aplysia* (Eales, 1960).

Visual inspection of the large hermaphroditic duct under the phase-contrast microscope suggests that no morphologically distinct atrial gland is present in *A. parvula*. However, a number of ELH-immunoreactive neurons and their processes are observed along the entire length of the duct (Fig. 7A–C). It is interesting that the size and shape of these cells is quite similar to the bag cell neurons, suggesting that they could arise from a

common population of precursor cells during development. These ELH-expressing peripheral neurons may act in a homologous manner to the specialized epithelial cells in the atrial gland and hermaphroditic duct of other *Aplysia* species (Painter et al., 1985) and secrete ELH-related gene products into the lumen of the duct. In *A. parvula*, no immunoreactive cells were found using the anti-A peptide serum, consistent with the nucleotide sequence data. The lack of both A and B peptide genes and an atrial gland in *A. parvula* suggests that these genes and tissue may have evolved concurrently and that the atrial gland may be a relatively recent structural addition to the *Aplysia* reproductive system.

#### Discussion

##### Phylogenetic conservation of ELH

The ELH gene family varies in size and composition between individual species of the genus *Aplysia*. Most of the *Aplysia* species examined appear to have only 2 ELH-related genes, and in *A. parvula* there exist 2 highly similar genes that encode ELH-like polyprotein precursors. No genes encoding A or B peptide-like precursors were identified in this species, a finding that supports the notion that the ELH gene is the ancestral prototype to the gene family. As the A and B peptide genes probably arose after divergence of the lines leading to the *A. californica* and *A. parvula* species, it is not clear whether other *Aplysia* species also possess only ELH genes, implying that they have diverged on the same line as *A. parvula*, or whether one or more of their genes will be similar to the A and B peptide genes. The genomic Southern blotting data suggest that *A. brasiliensis* has 2 linked ELH-related genes. The presence of atrial gland-like epithelial tissue in the large hermaphroditic duct (Painter et al., 1985) suggests that this species may have a single ELH-like and a single A or B peptide-like gene.

Our results indicate that 3 species—*A. californica*, *A. dactylo-*  
*omela*, and *A. vaccaria*—belonging to 3 different subgenera (*Neoaplysia*, *Varrisa*, and *Aplysia*) each have 4–5 members of the ELH gene family, while other species within the same subgenera have only 2 (Table 1, Fig. 1). This implies either that multiple, independent gene duplication or deletion events have occurred, or that the present classification scheme does not reflect the ancestral history of the ELH gene family. Interestingly, 2 species—*A. brasiliensis* and *A. fasciata*, which may be synonymous (Susswein and Ram, personal communication)—did in fact yield DNA blot patterns that were nearly identical in pattern and intensity.

Conservation of ELH in more distantly related molluscs has been investigated by Ram using an egg-laying bioassay to assess conserved bioactivity (Ram, 1982). *Aplysia* abdominal ganglion extracts can elicit intra- and interspecific egg laying and are also active across genus boundaries (Kupfermann, 1970; Ram, 1982).

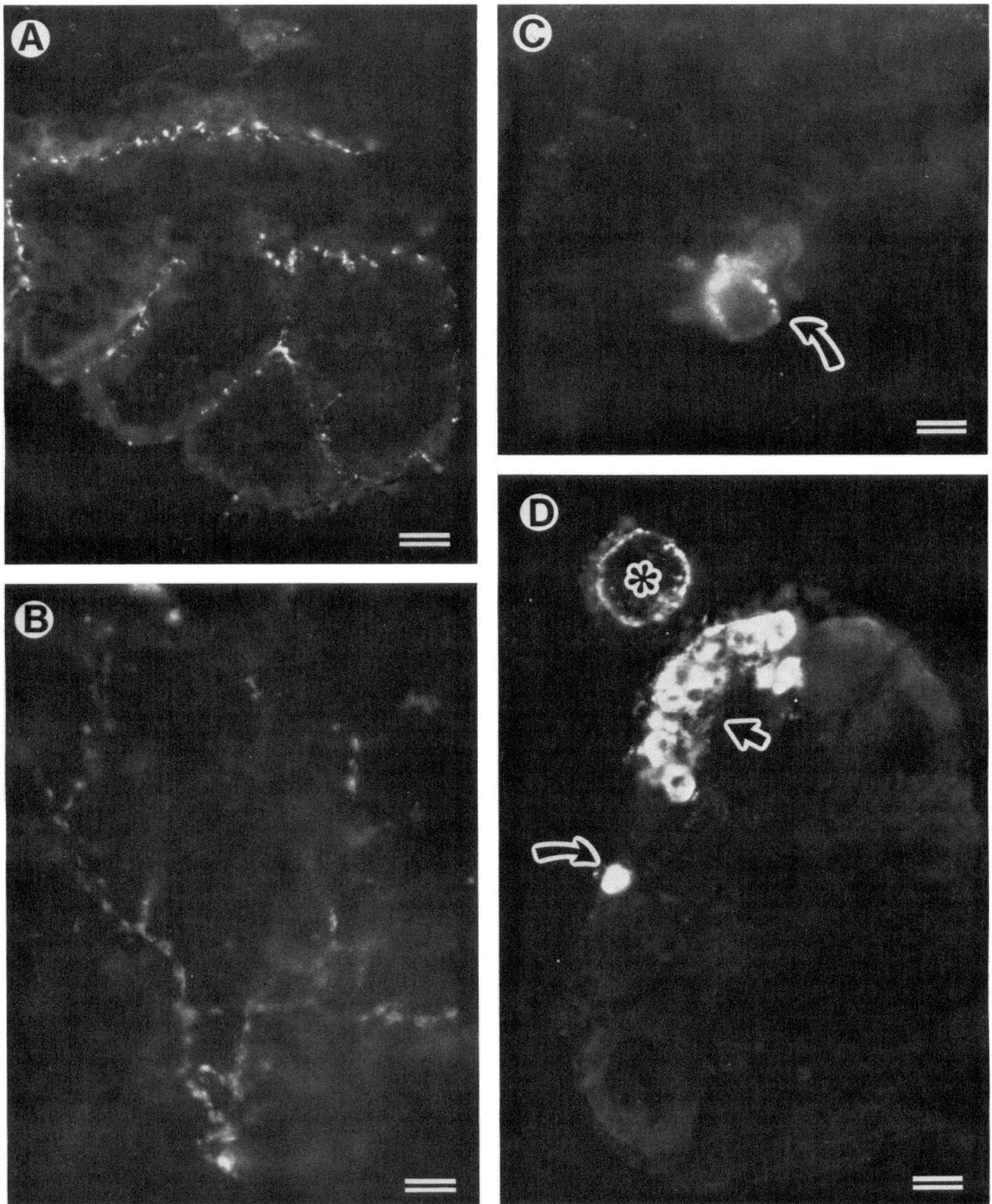


Figure 6. ELH-immunoreactive cells and processes in the abdominal ganglion and large hermaphroditic duct of *A. parvula*. Frozen sections of *A. parvula* tissues were reacted to rabbit antiserum against *A. californica* ELH, and immunoreactivity was detected using a fluorescein conjugated goat anti-rabbit IgG. *A*, Cross section of the *A. parvula* large hermaphroditic duct. Positively staining processes are seen around the perimeter of the duct. Bar, 10  $\mu$ m. *B*, Higher magnification of ELH-immunoreactive processes within the large hermaphroditic duct. Bar, 3  $\mu$ m. *C*, ELH-immunoreactive cell body along the perimeter of the large hermaphroditic duct is indicated by arrow. Bar, 3  $\mu$ m. *D*, Cross section through the

In addition, peptides with egg-laying activity in several other gastropods from different subclasses all have similar molecular weights as *Aplysia* ELH (Ram, 1983b). Recently, it has been demonstrated that the ovulation hormone of *Lymnaea stagnalis* is clearly related to *Aplysia* ELH, with 44% amino acid homology (Joosse et al., 1985). Thus, ELH is conserved across a wide range of the molluscan phylum. The amino acid sequence of the *Lymnaea* peptide, which functions in an analogous capacity as *Aplysia* ELH (Joosse et al., 1985), is presented in Figure 5 along with the amino acid sequences of *A. parvula* ELH and the ELH regions of the *A. californica* ELH and A peptide precursors. There are 13 identical residues in all 4 peptides; 7 of these occur in the N-terminal third of the molecule and 4 more are found in the C-terminal third, suggesting that these conserved regions of ELH might serve as binding sites in receptor recognition.

Based on evidence from the fossil record, it is believed that the pulmonate and opisthobranch gastropod subclasses diverged from each other about 350 million years ago (Moore and Pitrat, 1960). As this represents the minimum amount of time since divergence of the *Aplysia* and *Lymnaea* lines, the point of speciation between *A. californica* and *A. parvula* can be estimated. This is possible because in many instances the rate of amino acid substitutions in a specific protein has been constant with respect to time, thereby defining a "molecular clock" (Wilson et al., 1977). Approximately 20 amino acid substitutions have occurred between the *Lymnaea* and *Aplysia* ELH-like peptides, predicting a substitution rate of 1 amino acid/17.5 million years. The 8 substitutions between *A. parvula* and *A. californica* ELH then implies a speciation time of  $8 \times 17.5 = 140$  million years ago. This suggests that while ELH first evolved at least 350 million years ago, the A and B peptides probably arose less than 140 million years ago.

#### Evolution of the ELH gene family

Because all of the *Aplysia* species examined appear to have at least 2 ELH-related genes, it seems likely that duplication of the original ELH gene occurred before species divergence within the *Aplysia* genus. The existence of 2 ELH genes in the common ancestor of *A. californica* and *A. parvula*, however, implies that the divergence of ELH gene sequences between these species should be no greater than the divergence between the related genes within either species. Yet this is exactly the opposite of what is observed, as the interspecies divergence is much greater than the intraspecies divergence (Table 2). This paradoxical arrangement has been described in many other multigene families and implies that related genes within a species are not diverging independently, but instead are evolving in a "concerted" fashion, which serves to decrease intraspecific gene divergence (Arnheim, 1983).

Gene-correction events are probably responsible for both the high levels of intraspecific conservation between ELH-related genes and the large proportion of replacement substitutions among members of the *A. californica* ELH gene family. Thus, the higher-than-expected DNA sequence homology between the parv-1 and parv-2 ELH genes and the *A. californica* ELH and A or B peptide genes is likely to be the result of gene-correction events that have occurred within the last 140 million years. The large percentage of replacement substitutions and conservation of flanking regions suggests that similar processes have acted on the A and B peptide genes.

We do not yet know whether, as in *A. californica*, the highly homologous ELH genes in *A. parvula* are differentially expressed

in the abdominal ganglion and large hermaphroditic duct. In *A. californica*, members of the ELH gene family are expressed either in neural or reproductive tissues, although the A peptide gene is transcribed in both tissue types (Shyamala et al., 1985). It would be of some interest to determine at what point the tissue-specific regulation of the ELH gene family first arose. The representation of ELH-related genes among the different *Aplysia* species suggests that such specificity arose after duplication of the ancestral ELH gene, but before or concurrent with divergence of the A and B peptide genes.

#### Structure-function relationships of the *A. parvula* and *A. californica* ELH precursors and peptides

The *A. parvula* and *A. californica* ELH precursors are very similar in overall organization, implying that biosynthetic and processing pathways are conserved. The conservation of molecular structure usually denotes a corresponding conservation of function, and 2 peptides, ELH and alpha BCP, that have been established as bioactive molecules in *A. californica*, are both highly conserved in the *A. parvula* ELH precursor (Fig. 3). High sequence conservation of a third peptide, beta BCP, strongly suggests that this peptide also has important physiological functions in the 2 species. Alternately, the region surrounding the *A. californica* delta BCP is extremely divergent between the 2 precursors and also contains the only nonconserved dibasic pair. It is likely that this region either has no significant function in *A. californica* or *A. parvula*, or that it has been subjected to very different selection pressures in the 2 species.

The variability in gene number and type in the ELH gene family suggests that a corresponding variability in the egg-laying behavioral processes of different *Aplysia* species may also exist. In particular, products liberated from the A and B peptide precursors in *A. californica* may govern aspects of the egg-laying behavioral repertoire that are not expressed in *A. parvula*. (Interestingly, however, observations of mating and egg laying by *A. parvula* specimens maintained in the laboratory suggest more similarities than differences in the reproductive behavior of this species and *A. californica*. For example, both species mate in groups, simultaneously donating and receiving sperm. Furthermore, inhibition of locomotion and coiling of the egg mass are also shared behavioral components.) The A and B peptides, released from either the atrial gland or neurons in the head ganglia could serve to activate the bag cells and thereby initiate egg laying (Heller et al., 1980; Shyamala et al., 1985). In addition, it has been found that extracts from the large hermaphroditic duct can increase the frequency of mating, suggesting that atrial gland peptides may act as pheromones (Susswein and Benny, 1985). However, the finding that *A. parvula* does not possess genes coding for the A or B peptides implies that these molecules may not be required for either activation of the bag cell clusters *in vivo* or synchronization of mating behaviors. The precise physiological roles of these peptides are yet to be determined.

Egg laying is a vital component of the reproductive capacity of *Aplysia*, and the ELH gene products are thus intimately tied to biological success. Products that are involved in the regulation of reproduction could also have been involved in reproductive isolation events that led to speciation. It is hoped that the findings reported here will be useful in furthering our understanding of how the various peptide products of the ELH gene family arose and serve to govern egg laying in *Aplysia*. In addition, as egg laying is governed by both neuronal and endocrine processes, perhaps by analyzing the evolution of the peptide transmitters

←

abdominal ganglion and a pleural-abdominal connective. ELH-immunoreactive bag cell clusters and an ectopic bag cell are indicated by arrows. Processes within the connective sheath of the pleural-abdominal connective with ELH immunoreactivity are indicated by the asterisk.

used by these 2 systems we will gain insight into the evolution of their respective roles in governing this fixed action behavior pattern.

## References

- Arch, S., and T. Smock (1977) Egg-laying behavior in *Aplysia californica*. *Behav. Biol.* 19: 45-54.
- Arnheim, N. (1983) Concerted evolution of multigene families. In *Evolution of Genes and Proteins*, M. Nei and R. K. Koehn, eds., pp. 38-61, Sinauer Associates, Sunderland, MA.
- Benton, W. D., and R. W. Davis (1977) Screening gt recombinant clones by hybridization to single plaques *in situ*. *Science* 196: 180-182.
- Chiu, A. Y., and F. Strumwasser (1981) An immunohistochemical study of the neuropeptidergic bag cells of *Aplysia*. *J. Neurosci.* 1: 812-826.
- Chiu, A. Y., M. W. Hunkapiller, E. Heller, D. K. Stuart, L. E. Hood, and F. Strumwasser (1979) Neuropeptide egg-laying hormone of *Aplysia californica*: Purification and primary structure. *Proc. Natl. Acad. Sci. USA* 76: 6656-6660.
- Eales, N. B. (1960) Revision of the world species of *Aplysia* (Gastropoda, Opisthobranchia). *Bull. Br. Mus. (Nat. Hist.) Zool.* 5: 276-404.
- Heller, E., L. K. Kaczmarek, M. W. Hunkapiller, L. E. Hood, and F. Strumwasser (1980) Purification and primary structure of two neuroactive peptides that cause bag cell afterdischarge and egg-laying in *Aplysia*. *Proc. Natl. Acad. Sci. USA* 77: 2328-2332.
- Hood, L., J. H. Campbell, and S. C. R. Elgin (1975) The organization, expression and evolution of antibody genes and other multigene families. *Annu. Rev. Genet.* 9: 305-353.
- Joesse, J., R. H. M. Ebberink, W. P. M. Geraerts, R. F. Jansen, and A. Ter Maat (1985) The caudo-dorsal cells of *Lymnaea stagnalis*: Primary structure of the ovulation hormone and biosynthesis and role of multiple peptides released during egg laying. In *Neurosecretion and the Biology of Neuropeptides*, H. Kobayashi et al., eds., pp. 50-59, Japan Sci. Soc. Press, Tokyo.
- Kandel, E. R. (1979) *Behavioral Biology of Aplysia*, Freeman, San Francisco.
- Kimura, M. (1983) *The Neutral Theory of Molecular Evolution*, Cambridge U. P., Cambridge, UK.
- Kreiner, T., J. B. Rothbard, G. K. Schoolnik, and R. H. Scheller (1984) Antibodies to synthetic peptides defined by cDNA cloning reveal a network of peptidergic neurons in *Aplysia*. *J. Neurosci.* 4: 2581-2589.
- Kupfermann, I. (1970) Stimulation of egg-laying by extracts of neuroendocrine cells (bag cells) of abdominal ganglion of *Aplysia*. *J. Neurophysiol.* 33: 877-881.
- Long, E. O., and I. B. Dawid (1980) Repeated genes in eucaryotes. *Annu. Rev. Biochem.* 49: 727-764.
- Mahon, A. C., and R. H. Scheller (1983) The molecular basis of a neuroendocrine fixed action pattern: Egg laying in *Aplysia*. *Cold Spring Harbor Symp. Quant. Biol.* 48: 405-412.
- Mahon, A. C., J. R. Nambu, R. Taussig, M. Shyamala, A. Roach, and R. H. Scheller (1985) Structure and expression of the egg-laying hormone gene family in *Aplysia*. *J. Neurosci.* 5: 1872-1880.
- Maniatis, T., E. F. Fritsch, and J. Sambrook (1982) *Molecular Cloning*, Cold Spring Harbor Press, Cold Spring Harbor, NY.
- Maxam, A. M., and W. Gilbert (1980) Sequencing end-labeled DNA with base specific chemical cleavages. *Methods Enzymol.* 65: 499-560.
- Mayeri, E., B. S. Rothman, P. H. Brownell, W. D. Branton, and L. Padgett (1985) Nonsynaptic characteristics of neurotransmission mediated by egg-laying hormone in the abdominal ganglion of *Aplysia*. *J. Neurosci.* 5: 2060-2077.
- McAllister, L. B., R. H. Scheller, E. R. Kandel, and R. Axel (1983) *In situ* hybridization to study the origin and fate of identified neurons. *Science* 222: 800-808.
- Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green (1984) Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* 12: 7035-7056.
- Miyata, T., T. Yasunaga, and T. Nishida (1980) Nucleotide sequence divergence and functional constraint in mRNA evolution. *Proc. Natl. Acad. Sci. USA* 77: 7328-7332.
- Moore, R. C., and C. W. Pitrat (1960) *Treatise on Invertebrate Paleontology, Part I, Mollusca I*, Geological Soc. Am. and University of Kansas Press, Kansas City, KS.
- Mullins, J. I., D. S. Brody, R. C. Binari, and S. M. Cotter (1984) Viral transduction of c-myc gene in naturally occurring feline leukemias. *Nature* 308: 856-858.
- Painter, S. D., V. K. Kalman, G. T. Nagle, R. A. Zuckerman, and J. E. Blankenship (in press) The anatomy and functional morphology of the large hermaphroditic duct of three species of *Aplysia*, with special reference to the atrial gland. *J. Morphol.*
- Perler, F., A. Efstratiadis, P. Lomedico, W. Gilbert, R. Kolodner, and J. Dodgson (1980) The evolution of genes: The chicken preproinsulin gene. *Cell* 20: 555-566.
- Pinsker, H. M., and F. E. Dudek (1977) Bag cell control of egg-laying in freely behaving *Aplysia*. *Science* 197: 490-493.
- Proudfoot, N. J., and G. G. Brownlee (1976) 3' Non-coding region sequences in eukaryotic messenger RNA. *Nature* 263: 211-214.
- Ram, J. L. (1982) Interspecific induction of egg-laying in Hawaiian aplysids. *Gen. Comp. Endocrinol.* 47: 522-528.
- Ram, J. L. (1983a) Neuropeptide activation of an identifiable buccal ganglion motoneuron in *Aplysia*. *Brain Res.* 288: 177-186.
- Ram, J. L. (1983b) Gastropod egg-laying hormones. In *Molluscan Neuro-endocrinology*, J. Lever and H. H. Boer, eds., pp. 94-100, North-Holland, Amsterdam.
- Rothman, B. S., G. Weir, and F. F. Dudek (1983a) Direct action of egg laying hormone on ovotestis of *Aplysia*. *Gen. Comp. Endocrinol.* 52: 134-141.
- Rothman, B. S., E. Mayeri, R. O. Brown, P.-M. Yuan, and J. E. Shively (1983b) Primary structure and neuronal effects of alpha-bag cell peptide, a second candidate transmitter encoded by a single gene in bag cell neurons of *Aplysia*. *Proc. Natl. Acad. Sci. USA* 80: 5753-5757.
- Sanger, F., and A. R. Coulson (1977) The use of thin acrylamide gels for DNA sequencing. *FEBS Lett.* 87: 107-110.
- Scheller, R. H., J. F. Jackson, L. B. McAllister, J. H. Schwartz, E. R. Kandel, and R. Axel (1982) A family of genes that codes for ELH, a neuropeptide eliciting a stereotyped pattern of behavior in *Aplysia*. *Cell* 28: 707-719.
- Scheller, R. H., J. F. Jackson, L. B. McAllister, B. S. Rothman, E. Mayeri, and R. Axel (1983) A single gene encodes multiple neuropeptides mediating a stereotyped behavior. *Cell* 35: 7-22.
- Shyamala, M., J. R. Nambu, and R. H. Scheller (in press) Expression of the ELH gene family in the head ganglia of *Aplysia*. *Brain Res.*
- Stuart, D. K., and F. Strumwasser (1980) Neuronal sites of action of a neurosecretory peptide, egg-laying hormone, in *Aplysia californica*. *J. Neurophysiol.* 43: 499-519.
- Susswein, A. J., and Benny, M. (1985) Sexual behavior in *Aplysia fasciata* induced by homogenates of the distal large hermaphroditic duct. *Neuroscience Lett.* 59: 325-330.
- Wilson, A. C., S. S. Carlson, and T. J. White (1977) Biochemical evolution. *Annu. Rev. Biochem.* 46: 573-639.