Nucleic Acids Research, 1993, Vol. 21, No. 15 3479–3484

The *ets* multigene family is conserved throughout the Metazoa

Bernard M.Degnan+, Sandie M.Degnan+, Takeshi Naganuma§ and Daniel E.Morse*
Marine Biotechnology Center and Department of Biological Sciences, University of California, Santa Barbara, CA 93106, USA

Received April 15, 1993; Revised and Accepted June 11, 1993

ABSTRACT

This study provides the first empirical evidence for the conservation of the *ets* proto-oncogene transcription factor family throughout the Metazoa. Using the polymerase chain reaction with degenerate primers corresponding to conserved sequences within the ETS DNA-binding domain, we have detected *ets* genes in a range of lower metazoans, including sponges, ctenophores, anemones, flatworms and nematodes, and in several higher invertebrate metazoans. Many of these sequences are significantly divergent from the original v-*ets-1* oncogene, although most can be aligned with recently defined groups within the *ets* gene family. Multiple ETS domain sequences were detected in a number of the lower metazoan species, providing evidence for the existence of an *ets* multigene family at the earliest stages of metazoan evolution. In contrast, we were unable to detect any ETS sequences in fungal, plant or several protozoan DNAs. Our findings suggest that the duplication and divergence of *ets* proto-oncogenes responsible for generating the multigene family occurred concomitantly with the development of metazoan animals. In addition, these data corroborate other recent molecular evidence in providing strong support for the monophyletic origin of all multicellular animals, including sponges.

INTRODUCTION

Animal development requires complex intercellular cooperation and coordination for construction of the three-dimensional body plan. Although the mechanisms by which animals of diverse taxa undergo differentiation vary significantly at the external cellular and morphological levels, the molecular mechanisms by which the differentiated patterns of gene expression are achieved are highly conserved among metazoans (1). Evidence for this conservation comes from the large number of evolutionarily conserved protein domains shown to play essential roles in intercellular signalling, ligand binding, signal transduction and the regulation of transcription in animal development. One group of conserved genes are the proto-oncogenes, originally identified as the cellular homologs of sequences within mammalian and avian transforming retroviruses (2). Proto-oncogenes generally encode proteins known to be involved in the cascade of events, initiated by the interaction of cell surface receptors with intercellular signals, that specify the developmental fate of cells (i.e. signal transduction and gene expression) (3). The structural and functional conservation of many proto-oncogene families among several Metazoan phyla attests to their pivotal role in the process of normal animal development.

Of those nuclear proto-oncogenes that have been shown to be evolutionarily conserved, *myc* (4, 5) so far has been detected only in deuterostomes, *rel* (6), *jun* (7), *fos* (7) and *ets* (8, 9) have been identified only from higher metazoans, and *myb* (10, 11) has been detected throughout the Eukaryota. The *ets* family of proto-oncogenes was originally identified through sequence homology to the v-*ets* gene of the avian erythroblastosis virus E26 (12, 13). Members of this family are transcriptionally active in many different cell types in sea urchins (14), *Drosophila* (e.g. 15–17) and a wide variety of vertebrate species (e.g. 18–21), and appear to be one of the most conserved proto-oncogene families known. *Ets* family members are defined by the presence of the highly conserved ETS DNA-binding domain, corresponding to approximately 85 amino acid residues. This domain constitutes a recently recognized DNA-binding structural motif (22, 23) that has no structural homology to other well-characterized DNA-binding motifs; it is essential for binding to purine rich DNA sequences (23). Recent phylogenetic and structural comparisons of ETS domains identified from three triploblastic phyla have suggested that the origin of the *ets* gene family is very ancient, with gene duplication events occurring prior to the separation of arthropods (protostomes) and vertebrates (deuterostomes) (8, 9).

The function of *ets* genes in development remains unclear. In *Drosophila*, six ETS domain-containing genes are expressed in a complex spatial pattern during early embryogenesis and larval metamorphosis (15–17, 24, 25). *E74*, the best characterized of

* To whom correspondence should be addressed

Present addresses: +Department of Zoology, University of Queensland, Brisbane, Queensland 4072, Australia and §Deep Star Program, Japan Marine Science and Technology Center, 2-15 Natsushima-cho, Yokosuka, Kanagawa 237, Japan

Table 1. Summary of metazoan taxa from which ETS sequences were isolated. Major animal groups (Super-phyla) are in bold and underlined, phyla and higher classifications are in bold, and species names are in italics.

Classification and specific name	Common name	DNA source	Clones sequenced	ETS-containing clones
Deuterostome				
Chordate; ascidian				
Styela montereyensis	ascidian	sperm	15	2
Echinoderm				
Patiria miniata	sea star	sperm	8	3
Protostome				
Arthropod; crustacean				
Cancer antennarius	rock crab	hepatopancreas	13	10
Mollusc; gastropod				
Haliotis rufescens	red abalone	sperm	12	5
Mollusc; bivalve				
Hinnites giganteus	rock scallop	sperm	3	3
Annelid				
Chaetopterus variopedatus	polychaete	sperm	3	3
Pseudocoelomate				
Nematode				
unknown species	nematode	whole animal	24	8
Acoelomate				
Platyhelminth				
Notoplana acticola	flatworm	anterior tissue	12	3
Diploblast				
Cnidarian				
Anthropleura elegantissima	sea anemone	gonad	21	2
Ctenophore				
Pleurobranchia bachei	ctenophore	oocyte cDNA	21	5
Poriferan				
Haliclona sp.	sponge	dissociated cells	27	5
Tethya aurantia	sponge	dissociated cells	12	3

the *Drosophila ets* genes, is induced directly by the steroid hormone ecdysone at metamorphosis and is likely to be involved directly in regulating downstream genes essential for the global morphological changes that occur at this stage (15). Insight into the canonical function of the *ets* gene family may benefit from a phylogenetic approach, and from elucidation of the structure and function of the gene and its product in simpler model organisms. To elucidate the origin of the *ets* gene family, and as a first step in isolating these genes from simple model organisms, we have used the polymerase chain reaction (PCR) in efforts to amplify a portion of the ETS domain from a range of lower and higher metazoans, and from fungal, plant and protozoan genomes. Our results suggest that an *ets* multigene family may already have existed at the origin of animal multicellularity.

MATERIALS AND METHODS

Isolation of nucleic acids

Marine invertebrates. Twelve species of marine invertebrates, representing 10 phyla (Table 1) were collected in the vicinity of Santa Barbara, CA and maintained in an ambient temperature flow-through seawater system. Gametes were used as a source of nucleic acid when possible and were isolated either by dissection or induction of spawning. Tissues dissected from *Cancer antennarius, Notoplana acticola* and *Anthropleura elegantissima* were thoroughly washed in sterile seawater and inspected microscopically prior to cell lysis. Unidentified nematodes of similar size and morphology were transferred through three seawater bathes prior to lysis; it is not known if they represented a single species. Total nucleic acids were isolated from the sources listed in Table 1 by homogenization in lysis buffer (50 mM Tris−HCl pH 7.6, 300 mM NaCl, 100 mM EDTA, 2% SDS, 500 μg/ml proteinase K) in a ground-glass homogenizer. For species from which sperm were collected (Table 1), approximately 5 μl of pelleted sperm were resuspended in 100 μl sterile sea water prior to the addition of 400 μl lysis buffer. For remaining species, the ratio of tissue to lysis buffer volume varied with tissue type (Table 1), but in general was approximately 1:20. In all cases, lysates were incubated at 55°C for 15 min and then extracted twice with an equal volume of phenol/chloroform (26). Nucleic acids were precipitated by the addition of 0.1 volume of 3 M NaOAc pH 5.2 and 2 volumes of 100% ethanol, collected by centrifugation, and dissolved in water to concentrations of 50 to 500 ng/μl.

For the ctenophore *Pleurobranchia bachei*, total RNA was isolated from dissected oocytes as described above. cDNA was synthesized from 5 μg maternal total RNA at 37°C for 1 h in a 20μl reaction consisting of 50 mM Tris−HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 5 μM random hexamers, 1 mM dNTPs, 1unit/μl RNasin and 10 units/μl BRL Superscript reverse transcriptase (27). After denaturing the reverse transcriptase at 95°C for 5 min, the cDNA was amplified by PCR as described below.

Because of their propensity to harbor contaminating organisms, sponges were treated somewhat differently. Whole sponges were thoroughly cleaned using high pressure, UV-sterilized seawater. Cell suspensions were prepared by squeezing diced sponge tissue through 16 layers of cheese cloth into UV-sterilized seawater and inspecting dissociated cells microscopically for contaminants. The cells were incubated in sterile seawater with 2 μg/ml rifampicin at 23°C for 16 h (*Tethya aurantia*) or 3 h (*Haliclona* sp.) to allow

reaggregation. Dissociated sponge cells form homospecific cell masses in culture (28), reducing the likelihood of protozoan contamination. A sponge cell aggregate of approximately 100 μm diameter was removed from the culture in 0.5 μl of seawater, added directly to 24.5 μl PCR buffer and amplified as described below. For both species, three different cell aggregates were added separately to three gene amplification reactions.

Other DNAs. Most of the non-metazoan DNAs were donated from other laboratories, including DNAs from the ciliate protozoan *Tetrahymena thermophila* (E.Orias), the dinoflagellate *Alexandrium fundyense* (E.DeLong), the fungus *Aspergillus nidulans* (J.Carbon), and the pea plant (J.Cooper). For the ameoba *Trichoshaerium* sp., DNA was isolated, as described above for sperm, from 10 μl of pelleted cells from an axenic culture donated by M. Polner-Fuller.

PCR amplification

Degenerate PCR primers were designed based on consensus sequences within the conserved ETS domain compiled from sea urchin, *Drosophila* and vertebrate sequences (19). The primers were designed to match most closely with the human *ets-1* and *ets-2* amino acid sequences (19), although they also contain significant similarity to other ETS sequences. Sequences of primers 1 (27-mer) and 2 (30-mer) are 5′ ATMWSNTGGAC-NGGNGAYGGNTGGGAR 3′ and 5′ YTCCCANCCRTCNC-CNGTCCANSWKAT 3′, respectively. Within the ETS domain of *ets-1*, primer 1 corresponds to amino acid positions 24–32, and primer 2 corresponds to amino acid positions 66–75 (19). ETS sequences were amplified from 50 to 500 ng cDNA or genomic DNA in 25 μl reactions containing 10 mM Tris–HCl pH 8.3, 50 mM KCl, 2 mM $MgCl_2$, 1.25 μM primers 1 and 2 and 0.6 units *Taq* polymerase. An initial step of 95°C for 4 min was followed by 30 cycles of 94°C for 45 sec, 50°C for 60 sec, and 72°C for 10 sec. As a positive control to assess the suitability of DNAs for PCR amplification, duplicate reactions containing primers used to amplify the ribosomal DNA internal transcribed spacer region (Degnan and Lavin, unpublished) were performed for all samples of DNA tested. Negative controls run in parallel included the complete reaction mixture without DNA. All PCR reaction products were analyzed by agarose gel electrophoresis (26).

Cloning and sequencing

Following electrophoresis, ETS PCR products of the correct size were excised from a 2% agarose gel, purified and used to seed a second amplification reaction (26). The reamplification products were blunt-ended by the addition of $MgCl_2$ to a final concentration of 5 mM, and 5 units of Klenow polymerase, directly to the PCR reaction and incubation at 23°C for 30 min. Blunt-ended products were electrophoresed in a 2% agarose gel, excised, purified, and ligated to an *Eco*RV-digested pSK+ plasmid (Stratagene) for 12 h at 16°C (26). E.coli SURE (Stratagene) were transformed with ligated plasmid (29). DNA from randomly selected recombinant colonies was sequenced by the dideoxy cycle-sequencing method using BRL reagents and *Taq* polymerase (30).

RESULTS

Identification of metazoan ETS domains

In all metazoan DNAs investigated, primary PCR amplification with the degenerate ETS primers yielded a predominant 159 bp product. In a number of species (*Styela montereyensis*, *C.antennarius*, *A.elegantissima* and *P.bachei*), larger, less prevalent products also were observed (not shown). Our strategy of sequencing only 159 bp PCR products for the present study precluded detection of either intron-containing genes or *SPI*-like ETS domains (31, 32), hence we cannot comment on their existence in the species investigated. Attempts to generate a primary 159 bp ETS PCR product from representative protozoan, fungal and plant DNAs were not successful, even when the annealing temperature was reduced to 40°C. The same non-metazoan DNAs were successfully used as templates to amplify the internal transcribed spacer region of the rDNA unit (not shown), suggesting that failure to amplify ETS sequences using the degenerate ETS primers was due to the lack of suitable target sequences within the genomes of these organisms.

A total of 171 clones were sequenced from the 12 marine invertebrate species; these yielded 16 unique ETS sequences from 52 confirmed ETS-containing clones (Table 1; Figure 1). The 159 bp cloned PCR product generated 99 bp of sequence

```
                    35          45          55          65
                    |           |           |           |
H-ets-1     FKLSDPDEVARRWGKRKNKPKMNYEKLSRGLRY  44
H-ets-2     ---A-------------------------------  45
H-erg       --MT----------E--S--N---D----A---  46
H-fli-1     --MT-------F--E--S--N---D----A---  47
H-elk-1     ---V-AE----L--L----TN---D----A---  48
H-elf-1     ---V-SKA-S-L---H----D----TMG-A---  49
H-SAP-1     ---LQAE----L--I-----N---D----A---  50
M-PEA-3     ---IE-E----L--IQ--R-A---D----S---  21
R-GABPα     ---NQ-EL--QK--Q-----T--------A---  51

D-ets-2     ---T-------------------------------  24
D-ets-3     ---T----------E--S--N---D----A--/  17
D-ets-4     --IE-SVR--KL--R---R-A---D----SI-Q  17
D-elg       ---T---R---L--EK----A--------A---  16
D-ets-6     -R-I----------E--A--N---D----A---  17
D-E74A      ---V-SKA-S-L--MH----D----TMG-A---  15
D-pok       --IV--AGLAKL--IQ--HLS---D-M--A---  52

higher metazoans
ascidian    --------------------------------- (x2)
sea star    --------------------------------- (x3)
rock crab   ---T-------L--L------------------ (x10)
abalone     ---A----------I------------------ (x5)
scallop     --------------I------------------ (x3)
polychaete  --------------I------------------ (x3)

lower metazoans
nematode 1  ----------------------------F---  (x5)
nematode 2  ---V----------E--S--N---D----A---  (x2)
nematode 3  -RVV-H--------N----KT-T-D-----M-F  (x1)
flatworm    -RFK--EK--KK--DM----S-----M------  (x3)
ctenophore  --I-NSV-L--L--V--SN-I--FD----A---  (x5)
anemone 1   ---T----------I------------------  (x1)
anemone 2   ---TN-N----L--LH-----------------  (x1)
sponge Hal  ---I--E-------A------------------  (x5)
sponge Tth1 ---NNSE----M--L----TN---D----A---  (x2)
sponge Tth2 ---I--E-------S------------------  (x1)

consensus   FK.....V...WG..K....M.YE...R..RY
              R       L              FD       F
```

Figure 1. Alignment of derived amino acid ETS domain sequences with members of the *ets* gene family. Amino acid sequences correspond to ETS domain positions 33–65 (19). The basic domain begins at Arg (position 43). Dashes indicate identity with the human *ets-1* sequence. All new sequences are listed in bold and the number to the right of these sequences indicates the number of clones containing the sequence. A reference number follows all previously characterized *ets* family members. A consensus sequence of invariant and conserved amino acids (Asn/Gln, Asp/Glu, Arg/Lys, Ile/Leu/Val, Phe/Tyr, Ser/Thr) is listed at the bottom. Letter codes prefixing known ETS sequences are H, human; M, mouse; R, rat; D, *Drosophila*; and Su, sea urchin. The following *ets* family members are not shown because of 100% identity between positions 33–65 with a listed sequence: *v-ets-1* (12), chicken *ets-1* (53), *Xenopus ets-1A* and *ets-1B* (36) and sea urchin *ets-1* and *ets-2* (14) are identical to H-*ets-1* (44); chicken *ets-2* (54) and *Xenopus ets-2B* (35) are identical to H-*ets-2* (45); sea urchin *erg* (8) is identical to H-*erg* (46); *Xenopus ets-2A* (35) is identical D-*ets-2* (24); and *Drosophila Yan* (55) is equal to D-*pok* (52).

information corresponding to ETS domain amino acid positions 33–65 (Figure 1). This region contains portions of both the hinge domain (positions 33–42) and the basic domain (positions 43–65) as defined by Wang *et al.* (33). All ETS domain derived amino acid sequences possessed a number of highly conserved signature residues in these regions (9, 19), including Phe 33, Lys/Arg 34, Val/Leu 41, Trp 45, Gly 46, Lys 49, Met 54, Tyr/Phe 56, Glu/Asp 57, Arg 61, Arg 64 and Tyr/Phe 65 (Figure 1). It should be noted that members of the divergent *SPI* group of ETS domains (31, 32) have a number of non-conserved residues in these positions, and additional codons within the region under investigation.

Comparison of species-specific ETS sequences revealed a small number of single nucleotide misincorporations. In a comparison of 2,970 bp of ETS sequence from nine different species, 15 misincorporations were found, resulting in an average of one incorrect nucleotide every 198 nucleotides. Prior to cloning, all PCR products were reamplified, resulting in a slightly higher error rate than would be expected if the PCR products were cloned directly. Direct sequencing of the uncloned RT-PCR products was not undertaken, since the original 159 bp product could have represented a mixture of ETS sequences. Therefore, all ETS sequences (Figure 1) that have not been sequenced in at least triplicate may contain a small number of single nucleotide misincorporations. This PCR error rate is similar to that reported previously (34).

From all but one of the higher metazoan species analyzed, a single ETS sequence was obtained at least in triplicate by the PCR amplification method employed, so that PCR errors could easily be identified (Figure 1). The exception was the ascidian *S.monteryensis*, for which only two sequences were obtained. Comparison of the two *S.monteryensis* sequences revealed three neutral nucleotide changes in the third position of three codons: Ala 42, Lys 51 and Leu 59. This nucleotide substitution rate is higher than expected from polymerase errors alone, suggesting that at least some of the differences may be real and that the two ascidian sequences may represent two separate *ets* genes. Two forms of *ets-1* and *ets-2* have been identified in *Xenopus* and each form contains a number of neutral nucleotide changes within the ETS domain (35, 36). This uncertainty could be resolved by further sequencing of more ascidian ETS sequences.

From several of the lower metazoan species analyzed, more than one ETS sequence was confirmed (Figure 1). Three distinct sequences were obtained from the unidentified nematode, two from the anemone *A.elegantissima* and two from the sponge *T.aurantia*. Of the two *T.aurantia* derived amino acid sequences, one (Tth2) was most similar to an ETS sequence obtained from the second species of sponge, *Haliclona* sp. (Figure 1). These results thus extend the detection of *ets* genes to the entire range of metazoans and reveal the presence of an *ets* multigene family in the most primitive multicellular animals.

Comparison of lower and higher metazoan ETS domains

We attempted to classify each of the 16 partial ETS sequences into one of nine groups recently proposed by Laudet *et al.* (9) for the *ets* gene family: *ETS*, *ERG*, *ELG*, *PEA3*, *ELK*, *ELF*, *DETS4*, *POK* and *SPI* (Figure 2). Based upon derived amino acid sequence, ten of the 16 sequences were most similar to *ets-1* and *ets-2* sequences, which define the *ETS* group (Figure 2). Further discrimination between *ets-1* and *ets-2* sequences cannot be achieved without sequence data from distinguishing residues

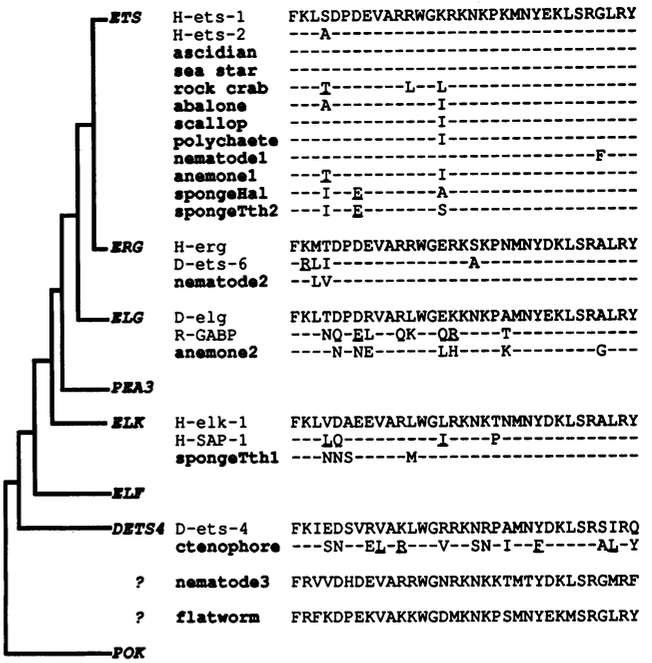

Figure 2. Alignment of ETS derived amino acid sequences with representatives of *ets* family groups. An evolutionary tree of ETS domain sequences is drawn to the right and shows the relationships among *ets* family groups (italics) as in (9). Tree branch lengths are arbitrary and do not represent degree of sequence divergence. New ETS sequences are aligned with representatives of each group as described in Figure 1. Conserved amino acid substitutions are underlined. Flatworm and nematode 3 ETS sequences could not be aligned to any other known ETS domains.

located outside the region of current analysis. The nematode 2 sequence has greatest identity with the human *erg* and *Drosophila ets-6* ETS domains, both members of the *ERG* group, which appears to be most closely related to the *ETS* group (9). The anemone 2 ETS is most similar to members of the *ELG* group, while the sponge Tth1 ETS sequence is likely a member of the more divergent *ELK* group with greatest sequence identity to the human *elk-1* and *SAP-1* genes (Figure 2). The ctenophore ETS may be a member of the *DETS4* group, which is substantially diverged from the *ETS* group to which the PCR primers were originally designed (Figure 2). Because of the high level of sequence divergence found in flatworm and nematode 3, we are unable to associate these sequences with any of the previously defined groups of *ets* genes. Further sequence information is required to determine if these divergent lower metazoan ETS domains can be included in previously defined groups, or if they constitute novel, more divergent groups.

DISCUSSION

Prior to this study, *ets*-related sequences had been identified only from chordates, arthropods and echinoderms (Figure 3). Based on the identification of *ets* genes in both protostomes and deuterostomes, it has been proposed that there existed an *ets* multigene family in primitive metazoans (8, 9). By identifying ETS sequences in a variety of lower metazoan taxa, using degenerate primers to some of the most conserved regions of the ETS domain, we have provided the first empirical evidence for the conservation of this gene family throughout the Metazoa

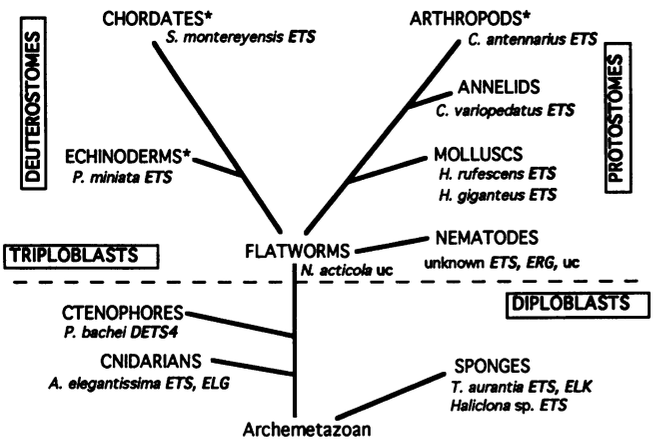

Figure 3. A hypothetical metazoan tree indicating possible evolutionary relationships (adapted from 37) of taxa surveyed for ETS sequences. Major animal groups are boxed and phyla are in upper case. An asterisk indicates phyla from which ETS domains have been characterized previously. Specific names from this study are listed under the appropriate phylum name, and are followed by the names of *ets* family groups detected in that species (uc, unclassified ETS sequence).

(summarized in Figure 3). It is commonly thought that a 'planula-like' progenitor gave rise to the diploblastic phyla, Cnidaria and Ctenophora, and to an acoelomate proto-Platyhelminthes (flatworm) which in turn gave rise to all triploblastic taxa (see 37 for an extensive discussion). The identification of *ets* proto-oncogenes in sponges, cnidarians and ctenophores confirms the existence of this gene family prior to the origin of the triploblast/diploblast dichotomy. Furthermore, comparison of partial ETS sequences obtained from these taxa suggests that a number of gene duplications had generated a multigene family prior to the origin of higher metazoans (Figures 2 & 3). Two distinct ETS sequences were detected within the genomes of the sponge *T.aurantia* and the anemone *A.elegantissima* (Figure 1). We propose that further sequences may be identified in lower metazoans by a more exhaustive search.

We have attempted to classify the partial ETS sequences obtained in this study according to the nine groups proposed by Laudet *et al.* (9) (Figure 2). With incomplete ETS domain sequences, our classification can only be considered tentative. From the lower metazoan species we have investigated, representatives of five divergent groups within the *ets* family (*ETS*, *ERG*, *ELG*, *ELK* and *DETS4*) and two unclassifiable ETS sequences (Flatworm and Nematode 3) were obtained (Figures 2 & 3). These results suggest that representatives of at least these five groups exist in both protostomes and deuterostomes, and we would predict that a *Drosophila ELK*-like *ets* gene and a human *DETS4*-like *ets* gene, neither of which have been detected to date, also are likely to exist. Based on our finding of two different *ets* genes in sponges, we also consider it likely that *ets* gene duplication events occurred prior to the origin of the simplest extant metazoans (Figure 3).

This study corroborates previous evidence that the *ets* gene family is one of the most conserved metazoan proto-oncogene families (Figure 3), implying that these genes may be essential in development in a wide range of taxa. *Ets* genes comprise a family of genes that code for transcription factors that are expressed in a developmentally restricted manner (e.g. 15, 17, 25); in this respect, they are functionally analogous to the homeodomain proteins (see 17 for discussion). Our identification of ETS domain sequences within the genomes of lower metazoans (Figure 3) extends the detection of this gene family to phyla in which HOM-C/*Hox* homeoboxes also were identified recently, namely flatworms (38), cnidarians (39, 40) and sponges (Degnan *et al.*, in prep.). Interestingly, we were unable to detect either these homeoboxes (Degnan *et al.*, in prep.) or ETS sequences (this study) in a suite of non-metazoan DNAs.

The role of *ets* proto-oncogenes and their products in metazoan development, for the most part, is not yet clear. However, it is possible that this multigene family, in addition to other metazoan gene families, was essential for establishment of the cellular environment in which multicellularity could arise. With the origin of the metazoan organism arose the need for a unique suite of genes not required by protozoans, to allow cells to coordinate the correct spatial and temporal patterns of gene expression and cytoskeletal arrangement needed for construction of a viable multicellular organism. Our identification of *ets* genes in the primitive metazoan phylum Porifera is an important addition to the recent identification of several apparently metazoan-specific genes, including HOM-C/*Hox* homeoboxes (Degnan *et al.*, in prep.), *src*-related proto-oncogenes (41) and structural genes encoding short chain collagens (42) in this phylum. The existence of several metazoan-specific genes in sponges provides strong empirical support for the hypothesis that the phylum Porifera arose from the same progenitor as diploblastic and higher metazoan taxa (37, 43). Furthermore, these observations together provide evidence for the existence of a unique suite of genes that arose concomitantly with, and hence may have allowed for, the development of metazoan animals. Our demonstration of the striking conservation of *ets* genes throughout the Metazoa provides access to experimental investigations of the structure and function of these genes in simple model organisms, and may provide a window into the molecular environment in which the multicellular organism arose and flourished.

ACKNOWLEDGMENTS

The authors thank Neal Hooker, Jim McCullagh and Shane Anderson for collection, cultivation and identification of many of the marine invertebrates, and Drs John Carbon, James Case, Jim Cooper, Ed DeLong, Eduardo Orias and Miriam Polne-Fuller for providing organisms or nucleic acids for this study. This research was supported by grants R01-CA53105 and R01-RR06640 from the National Institutes of Health.

REFERENCES

1. Davidson,E.H. (1991) *Development* **113**, 1–26.
2. Bishop,J.M. (1991) *Cell* **64**, 235–248.
3. Cross,M. and Dexter,T.M. (1991) *Cell* **64**, 271–280.
4. Moreau,J., LeGuellec,R., Leibovici,M., Couturier,A., Philippe,M. and Mechali,M. (1989) *Oncogene* **4**, 443–449.
5. Walker,C.W., Boom,J.D.G. and Marsh,A.G. (1992) *Oncogene* **7**, 2007–2012.
6. Steward,R. (1987) *Science* **238**, 692–694.
7. Perkins,K.K., Dailey,G.M. and Tjian,R. (1988) *EMBO J.* **7**, 4265–4273.
8. Lautenberger,J.A., Burdett,L.A., Gunnell,M.A., Qi,S., Watson,D.K., O'Brien,S.J. and Papas,T.S. (1992) *Oncogene* **7**, 1713–1719.
9. Laudet,V., Niel,C., Duterque-Coquillaud,M., Leprince,D. and Stehelin,D. (1993) *Biochem. Biophys. Res. Comm.* **190**, 8–14.
10. Biesalski,H.K., Doepner,G., Tzimas,G., Gamulin,V., Schroder,H.C., Batel,R., Nau,H. and Muller,W.E.G. (1992) *Oncogene* **7**, 1765–1774.

11. Paz-Ares,J., Ghosal,D., Wienand,U., Peterson,P.A. and Saedler,H. (1987) *EMBO J.* **6**, 3553–3558.
12. Leprince,D., Gegonne,A., Coll,J., deTaisne,C., Schneeberger,A., Lagrou,C. and Stehelin,D. (1983) *Nature* **306**, 395–397.
13. Nunn,M.F., Seeburg,P.H., Moscovici,C. and Duesberg,P.H. (1983) *Nature* **306**, 391–395.
14. Chen,Z., Kan,N.C., Pribyl,L., Lautenberger,J.A., Moudrianakis,E. and Papas,T.S. (1988) *Dev. Biol.* **125**, 432–440.
15. Burtis,K.C., Thummel,C.S., Jones,C.W., Karim,F.D. and Hogness,D.S. (1990) *Cell* **61**, 85–99.
16. Pribyl,L.J., Watson,D.K., Schulz,R.A. and Papas,T.S. (1991) *Oncogene* **6**, 1175–1183.
17. Chen,T., Bunting,M., Karim,F.D. and Thummel,C.S. (1992) *Dev. Biol.* **151**, 176–191.
18. Chen,Z., Burdett,L.A., Seth,A.K., Lautenberger,J.A. and Papas,T.S. (1990) *Science* **250**, 1416–1421.
19. Macleod,K., Leprince,D. and Stehelin,D. (1992) *Trends Biochem.* **17**, 251–256.
20. Majerus,M.A., Bibollet-Ruche,F., Telliez,J.B., Wasylyk,B. and Bailleul,B. (1992) *Nucleic Acids Res.* **20**, 2699–2703.
21. Xin,J.H., Cowie,A., Lachance,P. and Hassell,J.A. (1992) *Genes Dev.* **6**, 481–496.
22. Karim,F.D., Urness,L.D., Thummel,C.S, Klemsz,M.J., McKercher,S.R., Celada,A., VanBeveren,C., Maki,R.A., Gunther,C.V., Nye,J.A. and Graves,B.J. (1990) *Genes Dev.* **4**, 1451–1453.
23. Nye,J.A., Peterson,J.M., Gunther,C.V., Jonsen,M.D. and Graves,B.J. (1992) *Genes Dev.* **6**, 975–990.
24. Pribyl,L.J., Watson,D.K., McWilliams,M.J., Ascione,R. and Papas,T.S. (1988) *Dev. Biol.* **127**, 45–53.
25. Klambt,C. (1993) *Development* **117**, 163–176.
26. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning: A Laboratory Manual. Second Edition. Cold Spring Harbor Press, Cold Spring Harbor, New York.
27. Kawasaki,E.S. (1990) In Innis,M.A., Gelfand,D.H., Sninsky,J.J. and White,T.J. (eds.), PCR Protocols; A Guide to Methods and Applications. Academic Press, New York, pp. 21–27.
28. Humphreys,T. (1970) *Nature* **228**, 685–686.
29. Inoue,H., Nojima,H. and Okayama,H. (1990) *Gene* **96**, 23–28.
30. Murray,V. (1989) *Nucleic Acids Res.* **11**, 8889.
31. Ray,D., Culine,S., Tavitian,A., and Moreau-Gachelin,F. (1990) *Oncogene* **5**, 663–668.
32. Ray,D., Bosselut,R., Ghysdael,J. Mattei,M.G., Tavitian,A., and Moreau-Gachelin,F. (1992) *Mol. Cell. Biol.* **12**, 4297–4304.
33. Wang,C., Petryniak,B., Ho,I.C., Thompson,C.B. and Leiden,J.M. (1992) *J. Exp. Med.* **175**, 1391–1399.
34. Degnan,B.M. and Morse,D.E. (1993) *Molec. Mar. Biol. Biotech.* **2**, 1–9.
35. Wolff,C.M., Stiegler,P., Baltzinger,M., Meyer,D., Ghysdael,J., Stehelin,D., Befort,N. and Remy,P. (1990) *Nucleic Acids Res.* **18**, 4603–4604.
36. Stiegler,P., Wolff,C.M., Baltzinger,M., Hirtzlin,J., Senan,F., Meyer,D., Ghysdael,J., Stehelin,D., Befort,N. and Remy,P. (1990) *Nucleic Acids Res.* **18**, 5298.
37. Willmer,P. (1990) Invertebrate Relationships: Patterns in Animal Evolution. Cambridge University Press, Cambridge.
38. Webster,P.J. and Mansour,T.E. (1992) *Mech. Dev.* **38**, 25–32.
39. Murtha,M.T., Leckman,J.F. and Ruddle,F.H. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10711–10715.
40. Schummer,M., Scheurlen,I., Schaller,C. and Galliot,B. (1992) *EMBO J.* **11**, 1815–1823.
41. Ottilie,S., Raulf,F., Barnekow,A., Hannig,G. and Schartl,M. (1992) *Oncogene* **7**, 1623–1630.
42. Exposito,J.Y., LeGuellec,D., Lu,Q. and Garrone,R. (1991) *J. Biol. Chem.* **266**, 21923–21928.
43. Berquist,P.R. (1978) Sponges. University of California Press, Berkeley, CA.
44. Watson,D.K., McWilliams,M.J.,Lapis,P., Lautenberger,J.A., Schweinfest,C.W. and Papas,T.S. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7862–7866.
45. Watson,D.K., McWilliams-Smith,M.J., Kozak,C., Reeves,R., Gearhart,J., Nash,W., Modi,W. and Duesberg,P.H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1792–1794.
46. Rao,V.N., Papas,T.S. and Reddy,E.S.P. (1987) *Science* **237**, 635–639.
47. Ben-David,Y., Giddens,E.B., Letwin,K. and Bernstein,A. (1991) *Genes Dev.* **5**, 908–918.
48. Rao,V.N., Huebner,K., Isobe,M., Ar-Rushdi,A., Croce,C.M. and Reddy E.S. (1989) *Science* **244**, 66–70.
49. Thompson,C.B., Wang,C.Y., Ho,I.C., Bohjanen,P.R., Petryniak,B., June,C.H., Miesfeldt,S., Zhang,L., Nabel,G.J., Karpinsky,B. and Leiden,J.M. (1992) *Mol. Cell. Biol.* **12**, 1043–1053.
50. Dalton,S. and Treisman,R. (1992) *Cell* **68**, 597–612.
51. LaMarco,K., Thompson,C.C., Byers,B.P., Walton,E.M. and McKnight,S.L. (1991) *Science* **253**, 789–792.
52. Tei,H., Nihonmatsu,I., Yokokura,T., Ueda,R., Sano,Y., Okuda,T., Sato,K., Hirata,K., Fujita,S.C. and Yamamoto,D. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6856–6860.
53. Duterque-Coquillaud,M., Leprince,D., Flourens,A., Henry,C., Ghysdael,J., Debuire,B. and Stehelin,D. (1988) *Oncogene Res.* **2**, 335–344.
54. Boulukos,K.E., Pognonec,P., Begue,A. Galibert,F., Gesquiere,J.C., Stehelin,D. and Ghysdael,J. (1988) *EMBO J.* **7**, 697–705.
55. Lai, Z.C. and Rubin,G.M. (1992) *Cell* **70**, 609–620.