Sequences from Sea Urchin TU Transposons Are Conserved among Multiple Eucaryotic Species, Including Humans

DAN LIEBERMANN,^{1,2} BARBARA HOFFMAN-LIEBERMANN,² ANTHONY B. TROUTT,¹ LARRY KEDES,¹ AND STANLEY N. COHEN^{2*}

Departments of Genetics and Medicine, Stanford University, Stanford, California 94305,² and Veterans Administration Medical Center, Palo Alto, California 94304¹

Received 1 August 1985/Accepted 23 October 1985

Sequences homologous to various structural domains of the *Strongylocentrotus purpuratus* TU family of transposons are present in sea urchin species closely related to *S. purpuratus* and were found in close proximity to each other in linkage patterns that differed for different species. Sequence homologs of the inverted repeat outer domain (IVR-OD) segment were, in addition, present in a sea urchin related only distantly to *S. purpuratus* and in all other eucaryotic organisms surveyed. In humans, a polymorphic hybridization pattern was seen for genomic DNA obtained from different individuals. Sequence comparisons revealed that repeated sequence motifs similar to those making up the 15-base-pair direct repeat unit of the IVR-OD domain of the TU elements exist in the IVRs of transposons identified in *Drosophila melanogaster* and maize and in the transcription control regions of certain eucaryotic viral and cellular genes. The remarkable evolutionary conservation of IVR-OD homologs may reflect a biological role for these sequences in DNA transposition, the regulation of gene expression, or both.

Transposable genetic elements have been shown to be widespread in the biological world (see reference 49 for a review; 3, 13, 14, 45, 55) where they can affect gene expression (8, 28, 29, 33) and may generate genetic diversity (2, 5, 6, 10, 15, 16, 52). Recently, we reported the discovery of a heterogeneous family of modularly structured transposons, the TU elements, in the genome of the sea urchin *Strongylocentrotus purpuratus* (31, 24). TU1, the initially isolated TU element, consists of an internal (M) segment bracketed by long terminal inverted repeat (IVR) segments; these are divided into structurally distinct inner and outer domains—the latter characterized by a series of tandem repeat units 15 base pairs in length. Other members of the TU element family, which were isolated from an *S. purpuratus* genomic DNA library on the basis of their homology with the outer domain (OD) of the TU1 IVR, include M segments that are not present in TU1. The M segments of TU elements contain sequences that also are found, unassociated with the IVRs, at many other genomic locations. The inclusion of different segments of genomic DNA within different TU elements and the dispersion of TU element sequence domains to multiple locations of the *S. purpuratus* genome have suggested that these elements may play a role in the rearrangement of genomic DNA segments.

The structural heterogeneity of the TU family, the reiterated sequence motif of its IVR-OD segment, and the unique modularity of TU element structure have prompted us to investigate the occurrence of TU-related sequences in the DNAs of other sea urchin species, as well as in other eucaryotes. In this communication, we report that the IVR-OD sequences that define the TU element family in *S. purpuratus* a) have homologs even in a sea urchin species related only distantly to *S. purpuratus* (i.e., *Lytechinus pictus*) and can be found in more closely related sea urchins in close physical proximity to segments homologous to other TU element domains and b) show hybridization with DNA sequences in the chromosomes of an extraordinary range of

other eucaryotic species. Moreover, we show that the reiterated sequence motif observed in the IVR-OD segment of the TU elements is similar to sequence motifs found in the IVRs of previously studied transposable elements of *Drosophila melanogaster* and maize (*Zea mays*) as well as to sequences located in the transcriptional control regions of eucaryotic viral and cellular genes.

MATERIALS AND METHODS

General methods. Restriction enzymes were obtained from New England BioLabs, Inc., Beverly, Mass., Bethesda Research Laboratories, Inc., Gaithersburg, Md., and used as directed by the manufacturers. Fragments used as hybridization probes were radioactively labeled by nick translation (42). Agarose gels were run in TBE buffer (0.089 M Tris-borate [pH 8.3], 20 mM EDTA). Fragments of bacteriophage λ HindIII and pBR322 *Hinf*I or *Hae*II-digested pBR322 DNA were used as molecular length standards.

Preparation of λ phage libraries, recombinant plasmids, and genomic DNA. λ Phage libraries, each prepared from single sea urchin individuals of the four different species, were obtained as a generous gift from R. Britten and E. Davidson; the construction of the libraries has been described previously. These included a λ phage partial *Hae*III library of *S. purpuratus* (1), a λ phage partial *Eco*RI library of *S. drobachiensis* (43), a λ phage partial *Eco*RI library of *S. franciscanus* (25), and a λ phage partial *Eco*RI library of *L. pictus* (E. Davidson, personal communication).

The constructed plasmids pTU1 and pTU2 were both described previously by Hoffman-Liebermann et al. (24). λ Phage DNA was obtained from Bethesda Research Laboratories, and *Escherichia coli* DNA was from Worthington Diagnostics, Freehold, N.J. DNAs from chicken embryo, mouse L cells, monkey CV-1 cells, and human HeLa cells were prepared according to procedures described by Maniatis et al. (32). Sea urchin sperm DNAs from *S. purpuratus* and *L. pictus* were prepared as described previously by Kedes et al. (26). Sea urchin sperm DNAs from *S. drobachiensis* and *S. franciscanus* were obtained as a gen-

* Corresponding author.

FIG. 2. Southern blot analysis of genomic DNA obtained from the four different sea urchin species and hybridized to IVR-OD (A)-, IVR-ID (B)-, M (C)-, or IS (D)-specific probes. In panels A, B, and C, the DNA was digested with *Eco*RI endonuclease, whereas in panel D, *Kpn*I digests were carried out, since *Eco*RI cuts within the IS segment (24, 31). Hybridization was performed as described in Materials and Methods under conditions that require ~80% homology (11). With the IVR-OD probe only, hybridization was also performed under conditions that require ~90% homology (50% formamide, 4× SSC [50°C]), and similar blots were obtained. All DNA digests were electrophoresed in 0.65% agarose gels and treated as described in Materials and Methods. Pu, *S. purpuratus*; Dr, *S. drobachiensis*; Fr, *S. franciscanus*; Pi, *L. pictus*. Numbers on the left indicate size in kilobases.

erous gift from R. Britten and J. Roberts. DNAs from yeast (*Saccharomyces cerevisiae*), maize, *D. melanogaster*, *Aplysia vacaria*, and *Xenopus laevis* were obtained as generous gifts from J. Chenevert, T. Huynh, P. Beachy, R. Britten, and T. Mohun. DNAs from the leukocytes of different humans were a generous gift from S. Embury.

Southern blots. For analytical purposes, digests of cloned TU-element DNA were electrophoresed on 0.8 to 1.5% agarose minigels (0.05 to 0.1 µg per well). Sea urchin genomic DNA digests were electrophoresed on 0.65 to 1.5% agarose gels (1.5 µg per well) as indicated. For other animal species, amounts of DNA equal to the genome equivalents contained in 1.5 µg of sea urchin DNA (8.6×10^8 base pairs per haploid genome) were used in all digestions, and λ DNA was included as an internal marker in samples subjected to complete digestion. After electrophoresis, the gels were blotted onto nitrocellulose by the procedure of Southern (51). The blots were preincubated in hybridization buffer (4× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate]; 50 mM sodium phosphate [pH 7]; 0.05% each of bovine serum albumin [BSA], Ficoll [Pharmacia Fine Chemicals, Piscataway, N.J.], and polyvinylpyrrolidone) for 3 to 6 h at 65°C and then incubated at the same temperature in hybridization buffer containing the appropriate nick-translated probe at 0.5×10^6 cpm/ml. Blots were washed in 0.1× SSC–0.1% sodium dodecyl sulfate–50 mM sodium phosphate (pH 7) at 50°C for 3 h and exposed to prefogged X-ray film (XAR-5; Eastman Kodak Co., Rochester, N.Y.) at –80°C for ~12 h.

Linkage analysis experiments with TU element domain-

specific probes and phage genomic libraries. Nick-translated, domain-specific probes from TU elements (see Fig. 1) were hybridized with multiple plaque lifts from culture dishes containing various numbers of phage plaques (1,000 to 6,000) from the various genomic phage libraries. The number of plaques per dish was set according to the number of copies of each domain-specific probe within each of the genomes tested. In some instances, to confirm linkage results, hybridizing plaques were purified and rehybridized with the appropriate probe.

Plaque lifts were prepared and analyzed as described previously (7, 24). DNA isolated from purified phage (32) was used in the restriction enzyme experiments described below.

RESULTS

Occurrence of TU element sequence domains in other sea urchin species. To determine whether homologs of TU element segments occur in other sea urchin species, we used DNA fragments specific to the separate structural domains of previously identified TU elements of *S. purpuratus* (31, 24) (Fig. 1A) as interspecies hybridization probes against DNA blots of restriction endonuclease-digested genomic DNA. Included among the species tested was *L. pictus*, which is estimated to have diverged from *S. purpuratus* almost 200 million years ago (16) (Fig. 1B). Extensive hybridization with the IVR-OD probe was seen for genomic DNA from all four of the sea urchin species tested (Fig. 2A). Whereas no homology with *L. pictus* DNA was observed for the IVR inner domain (ID) segment of TU1, the M segments

FIG. 3. (A) Southern blot analysis of restriction enzyme digests of genomic DNAs obtained from different sea urchin species. Genomic DNAs from *S. purpuratus* (Pu), *S. drobachiensis* (Dr), *S. franciscanus* (Fr), and *L. pictus* (Pi) were digested with *FnuclI* and *HpaII* endonucleases and electrophoresed in 1.5% agarose gels. Hybridization conditions were as described in Materials and Methods for these experiments and for the experiments reported in all figures below, except where indicated. Black dots indicate the positions of the bands discussed in the text. (B) Southern blot analysis of bacteriophage λ DNA inserts derived from different sea urchin species. DNAs obtained from six plaque-purified IVR-OD hybridizing clones from each of the genomic libraries of *S. drobachiensis*, *S. franciscanus*, and *L. pictus* were digested with restriction enzymes that are known not to cut within the IVR-OD segment (24, 31). Of the six clones tested in each case, four, five, and five, respectively, yielded two IVR-OD hybridizing bands with at least one of the restriction enzymes used. An example of one such digestion for two different clones from each library is shown here. For *S. purpuratus*, the samples shown are those derived from the pTU1 and pTU2 plasmids (24, 31). Digestions with the following enzymes are shown: Pu1, *XhoI*; Pu2, *HindII*; Dr1, *Sall*; Dr2, *AvaI*; Fr1, *AvaI*; Fr2, *Sall*; Pi1, *AvaI*; Pi2, *BamHI*. In both A and B, hybridization was with the IVR-OD probe. Numbers at the left and center indicate size in kilobases.

of TU1, TU5, or TU6, or the probe containing the insertion sequence (IS) segments of TU4, all of these DNA regions hybridized to sequences present in genomic DNAs of *S. drobachiensis* and *S. franciscanus*. The *Mrep* segment of TU1 hybridized with the DNAs of all four sea urchin species tested, although hybridization with *L. pictus* DNA was extremely weak under the experimental conditions used (i.e., moderate stringency; ~80% homology [11]) (Fig. 2C). The IS segment of TU3, which is relatively scarce even in the genomic DNA of *S. purpuratus*, showed significant hybridization only to the DNA of *S. drobachiensis*, the sea urchin species most closely related to *S. purpuratus*.

A high proportion of the TU1 IVR-OD homologs found in *S. drobachiensis* and *S. franciscanus* appear to be similar to each other in overall structure as well as sequence (Fig. 3). Southern blots of genomic DNA digested with *FnuclI* and *HpaII*, both of which cut within the IVR-OD sequence (Fig.

1A), yielded bands of similar size (0.21 and 0.36 kilobases [kb]) for *S. purpuratus*, *S. drobachiensis*, and *S. franciscanus* (Fig. 3A). An additional hybridizing band representing a 0.64-kb *FnuclI*-*HpaII* fragment present in the IVR-OD of the *S. purpuratus* element TU1 (Fig. 1) was seen in the appropriately digested DNAs of *S. drobachiensis* but was absent from *S. franciscanus* DNA. The dissimilar pattern of hybridization observed for *S. purpuratus* DNA versus *L. pictus* DNA is in keeping with the ancient phylogenetic divergence of these organisms. However, this assay does not measure the degree of variability of the IVR-OD homologs in the different sea urchin species, since even a single base change could be responsible for the appearance of variable-sized hybridizing bands.

DNA from bacteriophage λ clones containing IVR-OD hybridizing sequences derived from genomic DNA libraries of *S. drobachiensis*, *S. franciscanus*, or *L. pictus* was cut with one of several restriction endonucleases that are known not to cleave within the IVR-OD sequence (31) and subjected to Southern blot analysis with the IVR-OD probe. As found for λ clones containing TU elements derived from *S. purpuratus*, cloned DNA from the other sea urchin species yielded two separate hybridizing bands (Fig. 3B), indicating that two discrete regions of homology with the IVR-OD probe are present in each of these genomic DNA inserts less than 15 kb in length.

Linkage of IVR-OD homologs to homologs of other TU element structural domains. The presence of TU element domain homologs in various sea urchin species does not address the question of whether these regions of homology are physically linked. To infer the frequency of linkage of the IVR-OD homologs to M-segment homologs of TU elements, we separately hybridized the IVR-OD probe and three different M-segment probes from *S. purpuratus* to multiple plaque lifts from DNA libraries derived from different sea urchin genomes. The results of this linkage analysis (Table 1) show that about 50% of the *S. purpuratus*- or *S. drobachiensis*-derived plaques that hybridized to the IVR-OD probe also hybridized to a TU1 M-segment probe that lacks the *Mrep* region (Fig. 1). In contrast, only 5% of the IVR-OD hybridizing plaques derived from *S. franciscanus* and none of the plaques derived from *L. pictus* DNA, hybridized to the TU1 M probe. Three and 1% of the IVR-OD hybridizing plaques from *S. purpuratus* and *S. drobachiensis*, respectively, hybridized with the M-region probe derived from TU5, whereas 50% of the plaques containing *S. franciscanus* DNA hybridized to this probe. The TU6-M segment probe showed homology to only a small percentage of the IVR-OD hybridizing plaques containing DNA from the three sea urchin species of the genus

TABLE 1. Linkage analysis^a of IVR-OD and M sequences in the different sea urchin species

Species	% (No./total) of IVR-OD hybridizing plaques which hybridize to:		
	TU1M	TU5M	TU6M
<i>S. purpuratus</i>	50 (85/180)	3 (5/180)	3 (5/180)
<i>S. drobachiensis</i>	60 (280/470)	1 (4/470)	1 (5/470)
<i>S. franciscanus</i>	5 (16/322)	50 (156/322)	1 (3/322)
<i>L. pictus</i>	0 (0/150)	0 (0/150)	0 (0/150)

^a Linkage analysis was carried out as described in Materials and Methods. Double plaque lifts were hybridized to an IVR-OD probe and the indicated M probe. A total of 5×10^4 plaques were examined (approximately one-half haploid genome).

Linkage Group	Possible Structure	Number of plaques			
		PU	Dr	Fr	Pi
1. IVR _{OD} —IVR _{ID} —M _{rep} —M		132	600	—	—
2. IVR _{OD} —IVR _{ID} —M		—	—	—	—
3. IVR _{OD} —IVR _{ID} —M _{rep} —		9	16	8	—
4. IVR _{OD} —IVR _{ID} —		24	19	56	—
5. IVR _{OD} —M _{rep} —M		—	—	46	—
6. IVR _{OD} —M		—	—	—	—
7. IVR _{OD} —M _{rep} —		50	25	45	—
8. IVR _{OD} —		140	280	460	300
9. IVR _{ID} —M _{rep} —M		6	8	—	—
10. IVR _{ID} —M		—	—	—	—
11. IVR _{ID} —M _{rep} —		—	—	—	—
12. IVR _{ID} —		21	16	12	—
13. M _{rep} —M		—	—	22	—
14. M _{rep} —		6500	9500	9000	100
15. —M		—	—	—	—

FIG. 4. Linkage analysis for IVR-OD, IVR-ID, TU1-M, and TU1M_{rep} sequences in the different sea urchin species. Linkage analysis was carried out as described in Materials and Methods. Quadruplicate plaque lifts from a genome equivalent, i.e., 10^5 plaques (*S. purpuratus* [24]), or one-half genome equivalent (other species) of the respective libraries were hybridized to an IVR-OD, IVR-ID, TU1-M, or TU1-M_{rep} probe. Indicated are the number of plaques per haploid genome in each linkage group. Schematic diagrams of possible element structures are shown. —, TU1M; ~, any M other than TU1M. Also indicated are cloned elements in which the deduced structure is as depicted (24; Fig. 1A). Further support for the putative structures shown in the figure was obtained from restriction enzyme analysis of cloned DNA, including also the clones shown in Fig. 3C. A dash under number of plaques indicates that no evidence was obtained for a linkage group having the structure proposed.

Strongylocentrotus. As was the case for the M segment of TU1, the M segments of TU5 and TU6 showed no homology to any of the IVR-OD hybridizing plaques derived from *L. pictus* genomic DNA. Thus, the frequency of linkage of IVR-OD homologs to a particular M-segment homolog appears to vary substantially in different sea urchin species.

Additional information about the frequency of physical association of TU element domain homologs in different sea urchin species was obtained from a linkage analysis experiment in which IVR-OD, IVR-ID, TU1-M, and TU1-M-M_{rep} domain-specific probes were hybridized against quadruplicate plaque lifts derived from the different sea urchin genomic DNA libraries. The results of this experiment (Fig. 4) showed that the linkage patterns are similar in *S. purpuratus* and *S. drobachiensis*. Moreover, the linkage pattern characteristic of TU1-like elements, abundant in *S. purpuratus* and *S. drobachiensis* as noted above, was absent from *S. franciscanus*. Instead, a new type of linkage group, in which sequences homologous to TU1-M-M_{rep} segments are linked to IVR homologs that lack the IVR-ID region, was seen (Fig. 4, linkage group 5). *S. franciscanus* also contained TU1-M-M_{rep} homologs that are not linked to sequences homologous to either IVR-OD or IVR-ID (Fig. 4, linkage group 13). In the sea urchin species most distantly related to *S. purpuratus* (i.e., *L. pictus*), sequences that hybridize to the IVR-ID or TU1-M segment probes were not detected (Fig. 2 and 4); however, sequences homologous to the M_{rep} segment of TU1 do exist in *L. pictus*, where they showed no

linkage to homologs of the IVR-OD probe (Fig. 4, linkage group 14). Finally, our linkage analysis indicated the existence in all three sea urchin species of the genus *Strongylocentrotus* of segments that are homologous to the IVR-ID region but are not linked to homologs of the IVR-OD segment (Fig. 4, linkage groups 9 and 12).

Sequences that are contiguous in the genome will, at some frequency, be separated in the library clones, because the end of the clone can fall within a particular sequence. The contribution of such events to linkage groups that appear at a low frequency must, therefore, be considered. It seems, however, most unlikely that such events have had a significant contribution to these linkages, given the structural features of TU elements and the types of linkage observed; the structural heterogeneity of these elements, as revealed by detailed structural analysis (24; Fig. 1A); and the small size of the structural domains—less than 1 kb compared with a 15-kb average size of the genomic inserts within the library clones. It is possible, therefore, that TU elements having the various structures suggested by our linkage analysis experiments (Fig. 4) do indeed exist, including another family of TU-related transposons having IVR-ID segments as termini.

Despite the existence in *S. purpuratus*, *S. drobachiensis*, and *S. franciscanus* of sequences homologous to the M region of TU1, the TU1-M homologs are not identical in all of these three sea urchin species. Genomic blots of DNA from *S. purpuratus* and *S. drobachiensis* both showed the 0.7-kb *Xho*I-*Hind*III band seen with TU1 DNA and an additional hybridizing band 0.48 kb in size (Fig. 5). The *S. franciscanus* genomic DNA blot, in contrast, yielded only a

FIG. 5. Southern blots of *Xho*I-*Hind*III restriction enzyme digests of TU1 and genomic DNA from the different sea urchin species electrophoresed in parallel. DNA digests were electrophoresed in a 1.5% agarose gel. The TU1M probe used was the *Xho*-*Hind*III M fragment indicated in Fig. 1A. Pu, *S. purpuratus*; Dr, *S. drobachiensis*; Fr, *S. franciscanus*; Pi, *L. pictus*. Numbers on the left indicate size in kilobases.

FIG. 6. (A) Southern blot analysis of genomic DNAs obtained from various animal species. DNA were digested with *Eco*RI endonuclease and electrophoresed in a 0.65% agarose gel. Hybridization was as described in Materials and Methods but was carried out at 60 instead of 65°C. E.c., *E. coli*; Ye., *S. cerevisiae*; Ma., maize; Dr., *D. melanogaster*; Ap., *A. vacaria*; S.U., sea urchin; Xe., *X. laevis*; Ch., chick; Mo., mouse; Mon., monkey; Hu., human. (B) Southern blot of restriction enzyme digests of genomic DNA from different humans. DNA from leukocytes of different humans was digested with *Eco*RI and electrophoresed in 0.65% agarose gels. Hybridization was as described in Materials and Methods at 65°C. In both A and B, blots were exposed for ~16 h.

single strongly hybridizing band that was approximately 1.5 kb in size.

Sequences homologous to the IVR-OD segment of TU1 exist in all eucaryotic organisms surveyed. The discovery of sequences homologous to the IVR-OD segment in a sea urchin species that is estimated to have diverged from *S. purpuratus* almost 200 million years ago led us to examine whether homologs of this DNA segment exist also in other eucaryotic organisms. We hybridized the IVR-OD probe against DNA blots of *Eco*RI-digested genomic DNA obtained from a variety of eucaryotic organisms, as well as against DNA isolated from *Escherichia coli* or bacteriophage λ . Sequences homologous to the IVR-OD probe were present in all eucaryotic organisms surveyed but were lacking in the procaryote-derived DNAs tested (Fig. 6A). No hybridization with the IVR-ID or TU1-M probes was observed with any of these species (data not shown). Further analysis showed a polymorphic pattern of hybridization of the IVR-OD probe with DNA obtained from leukocytes of six different humans (Fig. 6B).

DISCUSSION

The results reported here indicate that sequences homologous to various structural domains of the TU elements of *S. purpuratus* are present in other sea urchin species. Sequences homologous to the IVR-OD segment common to the

TU elements are present also in *L. pictus*, a sea urchin related only distantly to *S. purpuratus*, as well as in representatives of other phyla of multicellular eucaryotes and in yeasts. The striking evolutionary conservation observed for sequences homologous to the TU element IVR-OD segment was not seen for other structural domains of the TU elements or for other repetitive sequences obtained from sea urchins (unpublished results), including a member of the csp2108 family of sequences, known to occur in a variety of sea urchin species (47).

Analysis of bacteriophage λ clones containing IVR-OD hybridizing sequences derived from genomic DNA libraries from different sea urchin species indicates that two separate segments homologous to the IVR-OD regions are present within cloned genomic DNA segments less than 15 kb in length. Moreover, linkage analysis indicated that IVR-OD homologs are found in close physical proximity to homologs of other TU element domains in sea urchin species closely related to *S. purpuratus*. Taken together, these findings suggest that the IVR-OD homologs present in the DNA of other sea urchins are organized into TU-element-like structures (i.e., two IVR-OD segments bracketing a middle region).

Analysis of linkage of the various TU element domains indicates that different structural associations of these domains predominate in different members of the *Strong-*

FIG. 7. Repetitive sequence motifs common to transposon IVRs and transcription regulators. Arrows in the schematic diagram indicate the 5'→3' orientation of sequence motifs. For transposons, number of repetitive motif includes both IVRs. In the case of the FB, simian virus 40, BKV, and BKV cellular homologs, one or both of the common sequence motifs are contained within larger repeating units (31, 72, 68, and 21 base pairs in length, respectively). Elements (reference[s]): FB (40), DS (14), simian virus 40 (4, 19, 21–23, 27, 38, 54), simian virus 40 cellular homologue (9, 18, 34, 41, 46), BKV (48), BKV cellular homologue (44), herpes simplex virus thymidine kinase (35–37), β-globin (12). D.R.U., direct repeating unit. ^(a) The DS segment shown is part of a double DS that originates from an inverted insertion of one DS into another one (14).

ylacentrotus genus. The predominance of particular linkages in different but closely related sea urchin species supports the notion that the TU elements were assembled from modular domains during evolution rather than being derived from a single ancestral structure. In addition, the M segments of TU elements appear themselves to be members of families of dispersed repetitive sequences (24) that exist in different sea urchin species at different frequencies (Fig. 4; unpublished data). Potentially, the TU elements may be able to disperse and amplify other sequences in the genome, thus contributing to genetic diversity among sea urchin species (5, 6, 10).

The overall similarity in structure between the TU elements and the heterogeneous family of foldback (FB) transposons of *D. melanogaster* (39, 53) has been noted previously (24). These two groups of elements both appear to be members of a structurally distinct class of eucaryotic transposons characterized by a) IVRs that contain short direct repeat segments and b) heterogeneity of the M region located between the IVR termini (30). Although it is not known whether single FB transposons display the degree of structural modularity we have observed for the TU elements, it has been shown that the *Drosophila* TE elements, which consist of two FB elements flanking an M segment, can immobilize many different genomic sequences (20).

Comparison of the IVR sequences of these two families of transposable elements (Fig. 7) indicates that similarity between *S. purpuratus* TU and *Drosophila* FB elements exists

also at the nucleotide level (40). The IVR-OD segments of the TU elements and the inverted repeat termini of the FB elements contain direct repeat motifs that resemble each other and also resemble motifs present in the IVRs of a recently studied example of a DS (dissociation) element of maize, part of a double DS (14; Fig. 7). In each case, multiple copies of particular sequence motifs are tandemly repeated within the IVR segments of the transposon (30 and 50 repeats in the IVR regions of the maize transposon, 47 repeats in FB4 of *D. melanogaster*, 56 repeats in TU1 of *S. purpuratus*); moreover, two consensus sequences, one of which is highly pyrimidine rich in one strand and purine rich in the other, can be derived from the repeated sequences of the three species.

Further analysis indicated that the consensus sequences derived from the repetitive sequence motifs of the TU1, FB4, and DS elements are similar to sequences found in multiple-repeat units within the transcriptional regulatory regions of certain cellular and viral genes (Fig. 7). Included among the transcriptional regulatory sequences of homology with the consensus sequences of the transposon IVRs are: the simian virus 40 and human BK papovavirus (BKV) enhancer cores (4, 21, 27, 54); transcriptional regulatory sequences of certain genes carried by the mammalian viruses simian virus 40 (19, 22, 23, 38), BKV (48), and herpes simplex virus (35–37); a transcriptional regulatory sequence found upstream from the β-globin gene in several different mammalian species (12); as well as other tandemly repeated

cellular sequences (9, 18, 34, 41, 44, 46), whose role in the normal regulation of gene expression is still unclear. In each case, a motif consisting of multiple repeats of transcriptional regulatory sequences that include either one or two types of pyrimidine-purine-rich sequence elements can be seen.

The curious prevalence in genetically diverse eucaryotic species of sequences homologous to the IVR-OD sequence of the TU elements suggests an important biological role for these sequences—possibly related to processes such as transposition, the regulation of transcription, or both. The extent to which the IVR-OD homology observed in diverse eucaryotes results from hybridization of the probe to transcriptional regulatory regions such as those noted above or to other types of IVR-OD homologs present in these eucaryotes is not known. Further studies involving detailed analysis of the TU element homologs found in different sea urchin species, as well as characterization of the IVR-OD homologies that exist in other animal species, may define more precisely the nature of these widely prevalent sequences and elucidate their role in the eucaryotic genome.

ACKNOWLEDGMENTS

We acknowledge the contribution of David Gilbert to this work; he first demonstrated the hybridization of the IVR-OD to *Drosophila* DNA.

This work was supported by grant MV-44 from the American Cancer Society, Public Health Service grants GM 27241 (to S.N.C.) and GM 17995 from the National Institutes of Health, and a grant from the Veterans Administration to L.H.K. D.L. was a postdoctoral fellow of the California Division of the American Cancer Society when certain of these studies were carried out. A.B.T. was supported by Public Health Service grant 5 T32 CA 09302 awarded by the National Cancer Institute.

LITERATURE CITED

- Anderson, D. M., R. H. Scheller, J. W. Posakony, L. B. McAllister, S. G. Trabert, C. Beall, R. J. Britten, and E. H. Davidson. 1981. Repetitive sequences of the sea urchin genome: distribution of members of specific repetitive families. *J. Mol. Biol.* 145:5–28.
- Baltimore, D. 1985. Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome. *Cell* 40:481–482.
- Barker, R. F., D. V. Thompson, D. R. Talbot, J. Swanson, and J. L. Bennetzen. 1984. Nucleotide sequence of the maize transposable element Mu1. *Nucleic Acids Res.* 12:5955–5967.
- Benoist, C., and P. Chambon. 1981. In vivo sequence requirements of SV40 early promoter region. *Nature (London)* 290:304–310.
- Britten, R. 1981. DNA sequence organization and repeat sequences, p. 9–23. In D. M. Bennett et al. (ed.), *Chromosomes today*, vol. 7. Allen & Unwin Inc., Winchester, Mass.
- Britten, R. J. 1982. Genomic alterations in evolution, p. 41–64. In J. T. Bonner (ed.), *Evolution and development*. Springer-Verlag KG, Berlin.
- Benton, W. D., and R. W. Davis. 1977. Screening lambda gt recombinant clones by hybridization to single plaques in situ. *Science* 196:180–182.
- Cairns, J. 1981. The origin of human cancers. *Nature (London)* 289:353–357.
- Conrad, S. E., and M. R. Botchan. 1982. Isolation and characterization of human DNA fragments with nucleotide sequence homologies with the simian virus 40 regulatory region. *Mol. Cell. Biol.* 2:949–965.
- Davidson, E. H., and J. W. Posakony. 1982. Repetitive sequence transcripts in development. *Nature (London)* 297:633–635.
- Davis, R. W., D. Botstein, and J. R. Roth. 1980. *Advanced bacterial genetics: a manual for genetic engineering*, p. 222–223. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Dierks, P., A. van Ooyen, M. D. Cochran, C. Dobkin, J. Reiser, and C. Weissman. 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit β -globin gene in mouse 3T6 cells. *Cell* 32:695–706.
- Doring, H. P., and P. Starlinger. 1984. Barbara McClintock's controlling elements: now at the DNA level. *Cell* 39:253–259.
- Doring, H. P., E. Tillmann, and P. Starlinger. 1983. DNA sequence of the maize transposable element dissociation. *Nature (London)* 307:127–130.
- Douglas, H. E., and J. W. Valentine. 1984. "Hopeful monsters," transposons, and Metazoan radiation. *Proc. Natl. Acad. Sci. USA* 81:5482–5483.
- Dover, G. 1982. Molecular drive: a cohesive mode of species evolution. *Nature (London)* 299:111–116.
- Durham, J. W. 1966. Classification, p. 270–295. In R. C. Moore (ed.), *Treatise on invertebrate paleontology (U) Echinodermata 3 (1)*. The Geological Society of America and the University of Kansas Press, New York.
- Dynan, W. S., J. D. Saffer, W. S. Lee, and R. Tjian. 1985. Transcription factor Sp1 recognizes promoter sequences from the monkey genome that are similar to the simian virus 40 promoter. *Proc. Natl. Acad. Sci. USA* 82:4915–4919.
- Everett, R. D., D. Baty, and P. Chambon. 1983. The repeated GC-rich motifs upstream from the TATA box are important elements of the SV40 early promoter. *Nucleic Acids Res.* 11:2447–2464.
- Finnegan, D. J. 1985. Transposable elements in eukaryotes. *Int. Rev. Cytol.* 93:281–326.
- Fromm, M., and P. Berg. 1982. Deletion mapping of DNA regions required for SV40 early region promoter function in vivo. *J. Mol. Appl. Genet.* 1:457–481.
- Hartzell, S. W., B. J. Byrne, and K. N. Subramanian. 1984. Mapping of the late promoter of simian virus 40. *Proc. Natl. Acad. Sci. USA* 81:23–27.
- Hartzell, S. W., J. Yamaguchi, and K. N. Subramanian. 1983. SV40 deletion mutants lacking the 21-bp repeated sequences are viable, but have noncomplementable deficiencies. *Nucleic Acids Res.* 11:1601–1616.
- Hoffman-Liebermann, B., D. Liebermann, L. H. Kedes, and S. N. Cohen. 1985. TU elements: a heterogeneous family of modularly structured eucaryotic transposons. *Mol. Cell. Biol.* 5:991–1001.
- Johnson, S. A., E. H. Davidson, and R. J. Britten. 1984. Insertion of a short repetitive sequence (D881) in a sea urchin gene: a typical interspersed repeat. *J. Mol. Evol.* 20:195–201.
- Kedes, L. H., R. H. Cohn, J. C. Lowry, A. C. Y. Chang, and S. N. Cohen. 1975. The organization of sea urchin histone genes. *Cell* 6:359–369.
- Khoury, G., and P. Gruss. 1983. Enhancer elements. *Cell* 33:313–314.
- Kleckner, N. 1981. Transposable elements in prokaryotes. *Annu. Rev. Genet.* 15:341–404.
- Klein, G. 1981. The role of gene dosage and genetic transpositions in carcinogenesis. *Nature (London)* 294:313–318.
- Levis, R., M. Collins, and G. M. Rubin. 1982. FB elements are the common basis for the instability of the W^{DZL} and W^c *Drosophila* mutations. *Cell* 30:551–565.
- Liebermann, D., B. Hoffman-Liebermann, J. Weinthal, G. Childs, R. Maxson, A. Mauron, S. N. Cohen, and L. Kedes. 1983. An unusual transposon with long terminal inverted repeats in the sea urchin *Strongylocentrotus purpuratus*. *Nature (London)* 306:342–347.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*, p. 280–281, 371–372. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- McClintock, B. 1956. Controlling elements and the gene. *Cold Spring Harbor Symp. Quant. Biol.* 21:197–216.
- McCutchan, T. F., and M. F. Singer. 1981. DNA sequences similar to those around the simian virus 40 origin of replication are present in the monkey genome. *Proc. Natl. Acad. Sci. USA* 78:95–99.
- McKnight, S. L. 1982. Functional relationships between transcriptional control signals of the thymidine kinase gene of

- herpes simplex virus. *Cell* 31:355-365.
36. McKnight, S. L., and R. Kingsbury. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. *Science* 217:316-324.
37. McKnight, S. L., R. C. Kingsbury, A. Spence, and M. Smith. 1984. The distal transcription signals of the herpesvirus tk gene share a common hexanucleotide control sequence. *Cell* 37:253-262.
38. Mishoe, H., J. N. Brady, M. Radonovich, and N. P. Salzman. 1984. Simian virus 40 guanine-cytosine-rich sequences function as independent transcriptional control elements in vitro. *Mol. Cell. Biol.* 4:2911-2920.
39. Potter, S., M. Truett, M. Phillips, and A. Maher. 1980. Eukaryotic transposable genetic elements with inverted terminal repeats. *Cell* 20:639-647.
40. Potter, S. S. 1982. DNA sequence of a foldback transposable element in *Drosophila*. *Nature (London)* 297:201-204.
41. Queen, C., S. T. Lord, T. F. McCutchan, and M. F. Singer. 1981. Three segments from the monkey genome that hybridize to simian virus 40 have common structural elements. *Mol. Cell. Biol.* 1:1061-1068.
42. Rigby, P. W., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* 113:237-251.
43. Roberts, J. W., J. W. Grula, J. W. Posakony, R. Hudspeth, E. H. Davidson, and R. J. Britten. 1983. Comparison of sea urchin and human mtDNA: evolutionary rearrangement. *Proc. Natl. Acad. Sci. USA* 80:4614-4618.
44. Rosenthal, N., M. Kress, P. Gruss, and G. Khoury. 1983. BK virus enhancer element and a human cellular homolog. *Science* 222:749-755.
45. Rosenzweig, B., L. W. Liao, and D. Hirsh. 1983. Sequence of the *C. elegans* transposable element Tc1. *Nucleic Acids Res.* 11:4201-4209.
46. Saffer, J. D., and M. F. Singer. 1984. Transcription from SV40-like monkey and DNA sequences. *Nucleic Acids Res.* 12:4769-4788.
47. Scheller, R. H., D. M. Anderson, J. W. Posakony, L. B. McAllister, R. J. Britten, and E. H. Davidson. 1981. Repetitive sequences of the sea urchin genome II: subfamily structure and evolutionary conservation. *J. Mol. Biol.* 149:15-39.
48. Seif, I., G. Khoury, and R. Dhar. 1979. The genome of human papovavirus BKV. *Cell* 18:963-977.
49. Shapiro, J. A. (ed.). 1983. Mobile genetic elements. Academic Press, Inc., New York.
50. Smith, A. B. 1981. Implications of lantern morphology for the phylogeny of post-Paleozoic echinoids. *Paleontology* 24:779-801.
51. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
52. Syvanen, M. 1984. The evolutionary implications of mobile genetic elements. *Annu. Rev. Genet.* 18:271-293.
53. Truett, M. A., R. S. Jones, and S. S. Potter. 1981. Unusual structure of the FB family of transposable elements in *Drosophila*. *Cell* 24:753-762.
54. Weiher, H., M. Konig, and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancer. *Science* 219:626-631.
55. Zuker, C., J. Capello, H. F. Lodish, P. George, and S. Chung. 1984. *Dictyostelium* transposable element DIRS-1 has 350-base-pair inverted terminal repeats that contain a heat shock promoter. *Proc. Natl. Acad. Sci. USA* 81:2660-2664.