EVOLUTION OF A SPECIFIC CLONED REPEAT DNA SEQUENCE IN SEA STARS

by

Deborah Earlene Cool
B.Sc., Simon Fraser University, 1980

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in the Department
of
Biological Sciences

Deborah Earlene Cool 1982
Simon Fraser University
December 1982

Approval

Name: Deborah Earlene Cool

Degree: Master of Science

Title of Thesis: Evolution of a specific cloned repeat DNA sequence in sea stars

Examining Committee:

Chairman: Dr. C. L. Kemp

Dr. Michael J. Smith, Senior Supervisor

Dr. D. L. Baillie

Dr. A. T. Beckenbach

Dr. B. M. Honda, Assistant Professor,
Biological Sciences, SFU

Date approved 1982-11-19

PARTIAL COPYRIGHT LICENSE

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

Title of Thesis/Project/Extended Essay

Evolution of a specific cloned repeat DNA sequence in sea stars

Author: _____

(signature)

Deborah Earlene Cool

(name)

Dec. 14, 1982

(date)

ABSTRACT

Isolated sperm DNA from the starfish, Pisaster ochraceus contains a repetitive sequence representing approximately 0.5% of the haploid genome. This fraction of the genome was isolated and individual members of the repeat DNA class were cloned into the BamHI site of the pBR322 plasmid vector. This recombinant DNA was subsequently transformed into the bacteria host E. coli strain χ1776. Cloned repeat sequences, 3.9 kb in length, were compared using restriction enzyme analyses. One cloned representative member of the repeat DNA family, pPol431, was selected to determine the precise reiteration frequency in the genomes of the parent species, P. ochraceus and four other starfish species, P. brevispinus, Evasterias troschelii, Pycnopodia helianthoides, and Dermasterias imbricata. The amount of sequence homology within the repeat class in these five sea stars was assessed. Furthermore, the arrangement of repeat DNA family members was classified as clustered or dispersed in the sea star genomes.

The cloned repeat sequences contain three internal EcoRI restriction endonuclease sites which produce fragments 1.1, 1.3, 0.8 and 0.7 kb long. Both HincII and AvaI restriction enzymes cleave the repeat sequence. The locations of the enzyme cleavage sites within the repeat sequence have been mapped.

Filter hybridization of genomic DNA totally digested with the restriction enzyme PstI indicates that the 3.9 kb cloned fragment is part of a larger repeating element which is 5.3 kb long. The organization of the repeat family in three starfish genomes is maintained in a clustered arrangment. The repeat family is probably not represented in either oocyte or ribosomal RNA transcripts. Methylated cytosine residues in the 5'-CG-3' of the HpaII restriction site arrangement of the repeat element are only found in Pisaster ochraceus genomic DNA.

Heat denatured ^{32}P-labeled insert from pPol431 contains a maximum of 10% of its nucleotides organized into short inverted repeat segments. After denaturation these sequences immediately reanneal and form short double stranded regions called "snap-back" repeats.

There are approximately 500 copies of the cloned repeat in P. ochraceus parent species, 300 in Pisaster brevispinus, 150 in Evasterias troschelii, 180 in Pycnopodia helianthoides and 80 copies in Dermasterias imbricata. Genomic DNA from each sea star species was reassociated with ^{32}P-labeled repeat DNA from pPol431 and thermally denatured to determine the denaturation temperature (T_m) of hybrid DNA duplex. The depression in T_m of hybrid duplex relative to that of native duplex DNA (91.5^{o}C) indicates the amount of sequence divergence in the

repeat families from each species genome. The results show that the sequences of the repeat family are more conserved than unique DNA sequences between the sea star species. However, sequence homology and repeat sequence copy number per genome decrease with increasing phylogenetic distance from P. ochraceus.

ACKNOWLEDGEMENTS

The author wishes to express her appreciation to Dr. M.J. Smith for his encouragement, guidance and inspiration throughout this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY OF SPECIAL ABBREVIATIONS

Cot	concentration (M) x time (s)
cpm	counts per minute
Ci	Curie (3.7×10^{10} dps)
dps	decays per second
g	gram
kb	kilobase, equal to 1000 nucleotide base pairs
M	molar
mCi	milliCurie
mg	milligram
ug	microgram
ml	millilitres
ul	microlitres
RMS	root mean square
s	second
T_m	thermal denaturation temperature

INTRODUCTION

DNA reassociation kinetics provided the first quantitative evidence (Britten and Kohne, 1968) that there existed DNA sequences which are repeated in the genomes of higher organisms. Later it was established (Britten and Smith, 1970) that many of the repetitive sequences in calf DNA were interspersed among single copy sequences. Genomic DNA sequence organization in many animals has since been categorized in terms of possessing a short-period or "Xenopus" interspersion pattern (Davidson et al., 1973; Graham et al., 1974; Angerer et al., 1975; Goldberg et al., 1975; Chamberlain et al., 1975; Deininger and Schmidt, 1979) or, a long-period "Drosophila" interspersion pattern (Manning et al., 1975; Crain et al., 1976). The short interspersion pattern consists of repetitive DNA sequences on the order of 300 bp long interrupting single copy DNA at intervals of approximately 1000 base pairs. The long-period interspersion pattern contains repeats averaging five kilobases in length and the single copy stretches are even longer. Eden and Hendrick (1978) have demonstrated that the domestic chicken, Gallus gallus, genomic sequence organization falls into an intermediate category. Approximately half of the genome contains single copy DNA at 4.5 kb in length which is interspersed with repeats 2.0 kb long. The remainder of the genome has very long single copy regions, 17.5 kb in length,

before interruption by repeat DNA, 2.0 kb long. This arrangement is reminiscent of the Drosophila pattern of genomic sequence organization. It is clear however that eukaryotic genomes contain both long and short repetitive sequences creating more complex genomes in which single copy and repeat DNA sequences may not be exclusively organized in the "Xenopus" or "Drosophila" pattern (Britten et al., 1976; Galau et al., 1976; Cochet et al., 1979; Smith et al., 1980; Moore et al., 1981; Moyzis et al., 1981).

Repetitive DNA can be subdivided into "families" which are sets of sequences sufficiently homologous to form stable base-paired structures when the DNA is renatured under standard conditions (0.18 M Na^+, pH 6.8 and 60°C) (Britten et al., 1974). The number of members of repeat families can range from several million down to a few per genome. Single stranded DNA may also fold-back upon itself in a first order kinetic mechanism. These "snap-back" sequences are detected in the genomes of many organisms (Britten, 1981).

The functional roles and evolutionary properties of most repeat DNA is not known. Some repeat DNA such as the tandemly arranged, short repeats called satellite DNA, found at the centromeres and telomeres may be structurally important during cell division (Rubin, 1977). Other long and dispersed repeats may be transposable elements which can be transcribed: for

example; the Drosophila copia, 412 and 297 repeat families (Potter et al., 1979); or, the CS2108 or CS2034 families in sea urchin (Scheller et al., 1981). Lastly repeat DNA can consist of coding sequences for ribosomal or histone genes, although these constitute a minor fraction of repeat DNA. A model was proposed by Britten and Davidson (1969, 1971) to explain the role of short interspersed repeat DNA. Since the short repeats are found adjacent to single copy DNA it was suggested by these authors that perhaps they were involved in some transcriptional regulatory role.

It has been proposed (Doolittle and Sapienza, 1980; Orgel and Crick, 1980) that some DNA sequences serve no function in the organism and would have no affect on the phenotype thus are not subject to direct classical phenotypic selection. The most direct selection pressure of a DNA sequence is to first survive within the cells; sequences which do not contribute to the phenotype have adopted the ability of self-preservation and are maintained by "non-phenotypic selection". Therefore, these elements need no phenotypic rationale to explain their origin or maintenance within the cell. These sequences may have taken advantage of replication and recombination mechanisms in order to promote the amplification and dispersion of the sequence within the genome (Dover, 1980; Dover and Doolittle, 1980). The sequences referred to may be transposable elements in both

eukaryotes and prokaryotes as well as many kinds of repeat DNA sequence families. The amplification and dispersion of these DNA sequences occur either randomly (referred to as sequence independent or "ignorant") or are preferred (called sequence dependent or "selfish") (Dover and Doolittle, 1980). Many observations, on the other hand, suggest that in fact repeat sequences in eukaryotes may be under strong selective pressures, although there can be no experimental evidence yet to support their direct effect on the phenotype. For example, there is conservation of repeat sequence and frequency of some repeat DNA families found in interspecific comparisons (Britten, 1981); there is great differences in concentration of repeat sequences found in transcripts in various sea urchin tissues (Scheller et al., 1978); there is a particular pattern of repeat DNA sequence expression which is conserved in the sea urchin egg RNA (Moore et al., 1980); and there are repetitve sequences which are developmentally regulated in Dictyostelium discoideum (Zuker and Lodish, 1981).

The discovery of complex patterns of genomic sequence organization (Britten, 1982) has precipitated the need to know whether these arrangements are conserved between species suggesting an importance to function and even perhaps to species diversity (Davidson, 1982). Furthermore, how the patterns are originated, maintained or changed during evolution have yet to

be defined (Doolittle and Sapienza, 1980; Orgel et al., 1980; Dover et al., 1982). It had been shown by kinetic and thermal denaturation studies of DNA from many organisms (Britten and Kohne, 1968) that repetitive sequence family members were not identical. Furthermore in similar studies it was observed in sea urchin and other animal genomes that repeat DNA reassociated incompletely with related species repetitive DNA fractions. In fact, both amount and sequence homogeneity diminished with increasing phylogenetic distance. However, it was impossible to assess whether this phylogenetic relationship between organisms was true for all kinds of repeat DNA families since individual families from the parent species could not be isolated and characterized quantitatively in order to compare its properties in related species. Studying the evolution of repeat DNA was imprecise. Now however, with the aid of recombinant DNA technology studies of individual repeat families are facilitated by the ability to isolate individual members from eukaryotic genomes. Recently, information has been accumulating about specific eukaryotic families; for example, in the sea urchin (Klein et al., 1978; Moore et al., 1978; Anderson et al., 1981; Posakony et al., 1981; Scheller et al., 1981); in birds (Musti et al., 1980; Eden et al., 1981) in Drosophila (Rubin, 1977; Potter et al., 1979; Wensink et al., 1979) and in plants (Bedbrook et al., 1980a,b; Flavell et al., 1981).

Analysis of results obtained from these studies of the various organisms has begun to reveal the processes involved in repeat DNA evolution. First of all, repeat DNA families amplify in the genome by some unknown mechanism. This may involve unequal crossing over which produces an accummulation of successive duplications or the mechanism of sequence amplification may be a result of saltatory replication (Britten et al., 1968; Davidson, 1982). Amplification occurs at rates which appear both species specific and dependent on the repeat family itself since there is a vast difference between the sea urchin rates of repeat amplification (Moore et al., 1978; Britten, 1982) and many plants (Preisler and Thompson, 1981a,b). Repetitive sequence families "diffuse" around the genome by excising and inserting in a manner analogous to prokaryotic transposons or perhaps by circularized episomes (Stanfield and Lengyel, 1979). The episomes consist of only repetitive elements that can base pair with homologous sequences within the genome and insert by a single crossover event. There is still no direct evidence to support that these mechanisms or any others describe the movement of all kinds of repeat family members within the genome. Furthermore, the means of selection as to which sequences are amplified and/or dispersed is unknown. As a result of these chromosomal rearrangements of repeat sequences, family members can be found in the genome in either individual long or short repeats, as clusters in a uniform array, or as

tandemly arranged repetitive units containing scrambled subelements (Wensink et al., 1979; Musti et al., 1981; Scheller et al., 1981).

There is one important aspect of repeat DNA evolution which has not been explained. There is a decreasing number of repeat copies of a given family within related species relative to their increasing phylogenetic distance from the parent species. Also, a family of repeat DNA sequences found in two related species has more sequence homogeneity amongst its members within a species than amongst repeat DNA members between species (Dover et al., 1982). The pattern of within-species homogeneity and between-species heterogeneity is known as concerted evolution (Zimmer et al., 1980; Dover et al., 1982). The mechanism for this kind of evolution of repeat DNA is unlike others such as drift and selection. Fixation of variants in a repeat DNA family probably cannot be entirely responsible for concerted evolution since there must be fixation of the same mutations throughout the entire repeat DNA family. This seems unlikely if the family is large and dispersed. Although if the family elements are tandemly arranged then it may be feasible for fixation of unequal chromatid exchange events to occur (Smith, 1974; Tartof, 1974) resulting in expansion or contraction of a repeat DNA family cluster.

The process of selection has been disputed as to whether or not it is possible to phenotypically select for an apparently non-functional portion of the genome. Therefore a more appropriate mechanism to describe the evolution of repeat DNA families must be found and account for the fact that when speciation occurs there is a change in the dominant repeat DNA families which have been selectively amplified in the genome. Perhaps if we can answer this specific phenomenon we may be able to understand the role and evolutionary significance of repeat DNA in eukaryotes.

The great diversity of starfish species (Phylum Echinodermata, Class Asteroida) found on the coast of British Columbia provide a unique opportunity to study the evolution of repeat DNA in deuterostomes that have small genomes and a short-period interspersion pattern of sequence organization. Single copy rate of divergence (Smith et al., 1982) and some paleontological data have been used to determine the phylogenetic distance between five starfish species; Pisaster ochraceus, Pisaster brevispinus, Evasterias troschelii, Pycnopodia helianthoides, and Dermasterias imbricata. The Genera Pisaster, Evasterias and Pycnopodia are representatives of the Order Forcipulatida whereas Dermasterias is in the Order Spinulosida. The two orders were distinct 425 myr ago. The time since divergence between Pisaster and Dermasterias is 500 myr, 40 myr between

Pisaster and Evasterias or Pycnopodia, and only 5-10 myr between the two Pisaster sister species (Smith et al., 1982).

The P. ochraceus haploid genome size is 0.65 pg and it contains all frequency classes of repeat sequences including the fold-back sequences. Repeat sequences make up about 35% of the genome in this species. There is some uncertainty in the amount of repetitive DNA in the Dermasterias genome which has a genome size of 0.54 pg. Only about 24% of the genome contains repeat DNA and perhaps another 9% is very slow repeat sequences which are almost kinetically indistinquishable from single copy DNA. Reassociated Pisaster or Dermasterias DNA to a value of Cot 10 consists only of repetitive DNA sequences. Of this repeat DNA 4-5% is found in long repetitive segments (2600 bp). The DNA in these long repetitive families shows a high level of intragenomic sequence conservation (Smith and Boal, 1978; Smith et al., 1980). Furthermore, there is evidence of long regions of single copy DNA, on the order of 6500 bp representing 15% of single copy sequences in Dermasterias. This biological system naturally lends itself to studying genomic sequence organization and the evolution of repetitive DNA.

This study was undertaken to investigate the nature and evolution of repeat DNA in sea stars. A repeat DNA family from Pisaster ochraceus was isolated, cloned and characterized in the

parent species and its evolutionary properties determined over at least 800 myr of starfish divergence time. Long genomic Pisaster DNA was extracted from starfish sperm and a single family of repeats was selected using restriction enzyme digestion of the DNA. The repeat members were isolated from agarose gels and cloned into pBR322 and transformed into E. coli χ1776. One representative member, pPol431 was selected and used to characterize the family of repeats in the parent species P. ochraceus and in the five related species.

MATERIALS AND METHODS

Genomic DNA isolation from sea star sperm.

Genomic DNA was isolated from sea star sperm using the following DNA extraction procedures. In the first method developed by Marmur (1961), sperm was homogenized and resuspended in 0.1 M NaCl, 0.05 M Na_2EDTA, 0.05 M Tris, pH 7.4. The mixture was adjusted to 2% SDS and heated for 10 minutes at 60°C. Pronase B (Sigma) was added to a final concentration of 200 ug/ml and incubated overnight at 37°C. The protein was removed with the addition of 1/5 volume of 5 M $NaClO_4$ and an equal volume of phenol:SEVAG (1:1), shaken for 10 minutes and centrifuged at 5000 rpm for 10 minutes. SEVAG consists of chloroform:isoamyl alcohol at a ratio of 24:1. The aqueous phase was removed and

extracted two times with equal volumes of SEVAG. To precipitate the DNA, 1/10 volume of 3 M Na Acetate and two volumes of 95% ethanol were added. The DNA was wound onto a glass stirring rod and resuspended in the buffer solution described above. RNA was degraded with RNAse A (bovine pancreas) (Sigma) at 20 ug/ml for 2 hours at 37°C and the extraction and DNA precipitation procedures were repeated. DNA isolated by this procedure was used in all thermal denaturation and kinetic studies described below.

The second DNA isolation protocol was designed to minimize genomic DNA shearing during the extraction process (Blin et al., 1976). The homogenized sperm was dispersed in 0.2 M Na_2EDTA, 0.2 M Tris, pH 8.0 and warmed to 50°C for 1 hour. An equal volume of 1.5% SDS, 200 ug/ml Proteinase K (EM Biochemicals), both in the same buffer, was added and incubated overnight at 50°C. A solution of redistilled phenol (500 g) and 70 ml of m-cresol was adjusted to pH 8.0 and then to 0.1% 8-hydroxyquinoline. This mixture was added to an equal volume of DNA solution. An equal volume of SEVAG was further added and gently shaken for 30 minutes. The two phases were separated by centrifugation at 5000 rpm for 15 minutes and the aqueous phase poured off. The 8-hydroxyquinoline and m-cresol were removed by dialysis, for at least two days, into 0.02 M Na_2EDTA, 0.02 M Tris, pH 8.0. The DNA was treated with RNAse A, then Pronase B

and extracted with SEVAG as described in the Marmur method above. Long, intact DNA was required for restriction enzyme digestion of sea star genomic DNA and its subsequent gel electrophoresis.

Restriction enzyme digestion of genomic and plasmid DNA and gel electrophoresis.

Restriction endonucleases have been used extensively in comparative studies of DNA sequence organization. The isolation procedures and enzymatic properties of these enzymes which are found in bacteria have been described in detail elsewhere (Roberts, 1980). Their remarkable feature is that these endonucleases cleave double stranded DNA molecules by recognizing a specific sequence of four to six nucleotides and hydrolyzing the phosphodiester backbone at a specific nucleotide in that sequence producing fragments of discrete lengths. These fragments of digested DNA can be sized quantitatively using gel electrophoresis methods. A Poisson distribution of fragment lengths usually appear with enzymatic digestion of long genomic DNA which is observed in an agarose gel as a smear upon staining with ethidium bromide. However, when there are regions in long genomic DNA which are repetitive and contain a restriction enzyme recognition site in the repeating unit, then a discrete band of fragments will be seen over the background smear.

The volume of restriction enzyme reactions used here ranged from 20 ul to 50 ul in the buffer systems recommended by the suppliers. Genomic DNA was incubated with a minimum two-fold enzyme unit to DNA mass ratio for at least 5 hours to overnight at 37°C. All other enzyme reactions with plasmid DNA, bacterial DNA, or lambda DNA were incubated for only 2 hours at 2 units of enzyme to 1 ug DNA. One Unit of restriction enzyme is defined by Bethesda Research Labs as that amount of enzyme required to completely digest 1.0 ug of Lambda DNA (or equivalent) in one hour under the appropriate conditions in a volume of 50 ul. The reactions were stopped with 0.1% SDS and 0.02 mM Na_2EDTA and heated at 70°C for 10 minutes. The digests were made 1% glycerol and 0.025% bromophenol blue tracking dye.

The agarose gel was prepared from electrophoresis grade low Mr agarose (Biorad) in running buffer (40 mM Tris Base, 2 mM Na_2EDTA, 20 mM Acetic Acid, pH 8.1). The gel was submerged in the appropriate running buffer and the DNA samples were loaded into the wells with a Hamilton syringe. The samples were electrophoresed generally at 40V for 6 hours, and then the gel was stained for approximately 30 minutes in 1 ug/ml ethidium bromide solution. A photo of the stained gel was taken under u.v. transillumination at 254 nm. The sizes of the bands observed were determined by reference to markers of known

molecular weights which had been coelectrophoresed in the gel. Agarose concentrations used in this study were varied from 0.7 to 1.5% to measure the ranges of fragment lengths of digested DNA. Gels for the analysis of very small fragments were prepared from acrylamide.

5% acrylamide gels were prepared with 1.5 ml acrylamide (30% w/v stock of bisacrylamide: acrylamide, 1:29), 0.9 ml 10X TBE running buffer (see below), 90 ul fresh ammonium persulfate (100 mg/ml stock), and 6.51 ml double distilled H_2O. This solution was mixed and 15 ul TEMED added to a 0.05% final concentration. A 10 ml syringe was filled immediately with the acrylamide and injected into a water tight gel forming apparatus. The gels are electrophoresed at 100V for 1.5 hours. No more than 0.3 ug of DNA was loaded into each well and fragments ranging in sizes from 1000 to 50 base pairs could be sized accurately with this gel system.

To extract specific restriction enzyme digest fragments from agarose gels, the gel containing the DNA was placed in a dialysis bag along with a minimum volume of 0.5X TBE (1X TBE is 89 mM Tris Base, 89 mM Boric Acid, 2.5 mM Na_2EDTA, pH 8.3) running buffer so that the agarose was not touching the walls of the dialysis bag. The dialysis bag was placed in an electrophoresis apparatus along with 0.5X TBE and

electrophoresed at 50 mAmps overnight. Before removing the bag, the current was reversed for 2 minutes to remove DNA which may have adhered to the dialysis bag. The solution containing the DNA was removed, adjusted to 0.3 M Na Acetate and 2 volumes of cold 95% ethanol were added. After freezing the samples at -20^{o}C they were centrifuged at 40,000 rpm for 30 minutes in the SW50.1 rotor. The average yield was 60-80%.

Radioisotope labeling of DNA fragments.

DNA fragments were labeled by a modified nick translation method (Rigby et al., 1977). To label the DNA, a maximum of 30 uCi of alpha-^{32}P-dCTP (Amersham; 2500 Ci/mmole) was lyophilized in a 1.5 ml eppendorf tube to which various concentrated stock solutions required for labeling were added. The final 25 ul reaction contained 0.1 to 0.3 ug DNA, 2.0 uM dCTP, 15 uM TTP, 15 uM dGTP, 15 uM dATP, 500 mM Tris HCl, pH 7.5, 50 mM $MgCl_2$, 40 ug/ml BSA, 10 mM dithiothreitol, 0.1 ug DNAse and 5 units of DNA Polymerase (BRL). The reaction was incubated at 16^{o}C for 2 hours and stopped with 5 ul each of 2 M Tris HCl, pH 7.0 and 0.5 M Na_2EDTA and 40 ul of double distilled H_20 at 4^{o}C. The reaction mixture was extracted with an equal volume of saturated phenol and SEVAG. Unincorporated isotope was separated from labeled DNA by G-75 Sephadex column chromatography. The labeled DNA is excluded from the Sephadex. Tracer DNAs with specific

activities of $1x10^8$ to $1x10^9$ cpm ug^{-1} were routinely obtained.

Ligation of isolated Pisaster ochraceus 3.9 kb fragments to pBR322 and transformation into E.coli χ1776.

The ligation reaction contained 0.2 ug of pBR322 (Bolivar et al., 1977) which had been digested with BamHI and 0.1 ug of 3.9 kb fragments from P. ochraceus, 66 mM Tris HCl (pH 7.6), 6.6 mM $MgCl_2$, 10 mM dithiothreiotol, 0.4 mM ATP, 1 unit of T4 Ligase in a total volume of 100 ul. The reaction was left overnight at 16°C. A unit of ligase is that amount of T4 DNA ligase (BRL) that catalyzes the conversion (via exchange reaction) of 1 nmol ^{32}PPi into (alpha/beta ^{32}P)-ATP at 37°C for 20 minutes (Weiss et al., 1968).

E. coli χ1776 (Curtis III et al., 1977) was transformed with the ligation reaction. Host and recombinant DNA facilities for "A" level of containment were within the NSERC guidelines. 100 ml of Luria broth (1% w/v bacto-tryptone, 0.5% w/v yeast extract, 0.5% w/v NaCl, pH 7.5) supplemented with diaminopimelic acid (DAPA) at 100 ug/ml and thymidine at 10 ug/ml were innoculated with 10 ml of a fresh overnight culture of E. coli. The cultures were grown to an absorbance of 0.3 at 600 nm. The cells were centrifuged at 3000 rpm for 10 minutes at room

temperature in a JA20 Beckman rotor and the pellet was resuspended in 10 ml of fresh ice cold transfection buffer (70mM $MgCl_2$, 30mM NaAcetate, 30 mM $CaCl_2$) (Mandel and Higa, 1970; Enea et al., 1975) and left on ice for 20 minutes. After incubation on ice the cells were centrifuged at 3000 rpm for 10 minutes at 2°C and resuspended in 3.0 ml of transfection buffer. 200 ul aliquots of the cells were distributed into Eppendorf tubes and a maximum of 20 ng of ligated pBR322 and insert DNA was added to the cells in each tube. The mixtures were incubated on ice for 10-15 minutes, heated at 42°C for 2 minutes and then immediately put back on ice for 1 hour. 2.0 ml of Luria broth containing DAPA (100 ug/ml), thymidine (10 ug/ml) and ampicillin (25 ug/ml) were added to the pooled aliquots, centrifuged at 3000 rpm for 10 minutes at room temperature and resuspended in the same fresh media. The newly transformed cells were grown overnight in Luria broth and then plated from which positive colonies were selected.

Screening for Positive Transformants.

The method for drug-selecting recombinant E. coli bacteria entails first the isolation of unique colonies which are ampicillin resistant and tetracycline sensitive. The newly transformed culture was streaked onto ampicillin agar plates containing diaminopimelic acid and thymidine and incubated at

37°C for 48 hours. Individual colonies were picked with sterile toothpicks into microtitre wells containing 200 ul of Luria broth, ampicillin (25 ug/ml), DAPA (100 ug/ml) and thymidine (10 ug/ml). These cultures were replica plated on agar plates containing the essential nutrients and either ampicillin or tetracycline.

Cells which grew only on the ampicillin plate and not on the tetracycline plates (10 ug/ml) were further screened with P. ochraceus repeat DNA probe by filter hybridization methods. The positive cultures were spotted from the microtitre wells onto nitrocellulose filter paper discs, which had been placed on fresh agar plates containing the same nutrients and drugs as in the wells. The colonies were grown for at least 48 hours at 37°C until the clone colony size was approximately 2 mm in diameter. The bacterial clones attached to the nitrocellulose filters were lysed and denatured in situ according to the method of Thayer (1979). The filters were air dried and baked in a vacuum oven for 2 hours at 80°C.

A P. ochraceus repeat DNA probe was prepared from the BamHI 3.9 kb DNA fragments isolated from agarose gels. The 3.9 kb fragments were ^{32}P-labeled by nick translation and used to screen by filter hybridization (as described below) the transformed colonies whose plasmid DNAs had been bound to

nitrocellulose paper.

Amplification of Positive Clones.

Bacteria which contain recombinant plasmids were numbered according to the well that the colonies occupied in the microtitre plate. The initials preceeding the number indicate that the cloning vector is a plasmid, "p", and the species name Pisaster ochraceus, "Po". The colonies of four plasmids which had hybridized to ^{32}P-labeled DNA in the in situ colony filter hybridization experiment above were selected and amplified. They were designated pPo328, pPo1289, pPo1431 and pPo1675. 50 ml of fresh Luria broth containing the approriate nutrients and ampicillin, were inoculated with 100 ul of the culture from the microtitre well. The culture was shaken in a New Brunswick shaker incubator overnight at 37°C and 2000 rpm. This fresh overnight culture was aliquoted into six 500 ml erlenmeyer flasks containing fresh Super broth (2.4% yeast extract, 1.2% tryptone, 5% (v/v) glycerol, 2.2% K_2PO_4, 0.48% KH_2PO_4), DAPA (100 ug/ml), thymidine (10 ug/ml) and ampicillin (25 ug/ml) and shaken until growth reached mid-log phase or to an absorbance of 0.3 at 600 nm. Chloramphenicol (Clewell, 1972) was added to a final concentration of 12.5 ug/ml. The cultures were shaken overnight at 37°C. The cells were collected by centrifugation at 6000 rpm for 15 minutes at 4°C. The pellet

was resuspended in a maximum of 3.0 ml of 25% sucrose, 50 mM Tris HCl, pH 8.0. At 4°C the cell suspension was adjusted to 0.1 M NaCl and 20 mM Na_2EDTA and incubated with 1.5 mg/ml lysosyme (Sigma) for 15 minutes. These spheroblasts were lysed by a 5-fold dilution with 1 mM Tris HCl, pH 8.0, 1 mM Na_2EDTA, and 0.2% Sarkosyl. The cells were incubated on ice for 30 minutes with periodic gentle agitation. Cell debris was pelleted along with most of the genomic DNA by sedimentation in a Beckman JA-20 rotor at 20000 rpm for 1 hour at 2°C. The supernatant containing the plasmids was decanted leaving a jelly-like pellet. Protein in the lysate was extracted with an equal volume of SEVAG, gently mixed and centrifuged to separate the phases. The procedure was repeated and the final supernatant was made to 15 ml with the dilution buffer (1 mM Tris HCl, pH 8.0 and 1 mM Na_2EDTA). 23 ml of CsCl saturated with dilution buffer was added to a final CsCl density of 1.593 g/ml. Ethidium bromide was added to a final concentration of 250 ug/ml. The samples were transferred to a Beckman heat sealable tube. The plasmid containing solutions were sedimented at 45000 rpm for 24 hours at 20°C in a Beckman VTi50 rotor. Ethidium bromide stained plasmid DNA was observed wih a 375 nm uv lamp as a discrete band midway in the gradient. The band of closed circular plasmid DNA was removed from the gradient with a syringe and 20 gauge needle. The ethidium bromide was extracted with isopropyl alcohol saturated with the CsCl. The DNA was

dialyzed into low salt buffer (10 mM Tris HCl, pH 8.0) and precipitated with 2 volumes of ethanol at -20°C. The plasmids were collected by centrifugation at 10000 rpm for 1 hour at 0°C in a JA-20 rotor. The pellet was lyophilized and resuspended in a small volume of 10 mM Tris HCl, pH 8.0 and 1 mM Na_2EDTA.

Nitrocellulose filter hybridization and autoradiography in clone screening.

The ^{32}P-labeled 3.9 kb BamHI DNA fragment from pPol431 was used to screen the plasmid DNA bound to the nitrocellulose filters. The filter discs were placed in a heat sealable plastic bag (Kapak) and prehybridized for 4 hours at 68°C with 4X SET (4 X 0.15 M NaCl, 0.03 M Tris HCl, pH 8.0, 1 mM Na_2EDTA), 5X Denhardt's solution (1% Bovine Serum Albumin, 1% Polyvinylpyrrolidone, 1% Ficoll) (Denhardt, 1966), 0.1% Na pyrophosphate, 2.5 mM Na phosphate buffer, pH 6.8, 0.1% SDS, 50 ug/ml sheared and denatured calf thymus DNA. After prehybridization approximately 5 x 10^5 cpm/filter of denatured probe was added and the DNA was hybridizaed to the filters with shaking at 68°C for at least 18 hours. The filters were washed at 68°C with 4X SET, 0.1% SDS, 2.5 mM phosphate buffer, 0.1% NaPPi. The wash solution was replaced every hour for 3 hours and finally the filters were washed with 1X SET, 0.1% SDS,

2.5 mM Na phosphate buffer and 0.1% NaPPi for 1 hour. After air drying the filters, they were exposed to Kodak BBI, preflashed film with the aid of Dupont enhancer screens at -80°C. The film was exposed for at least 2 hours and developed with Kodak D19 developer and Kodak fixer. Black spots on the X-ray film correspond to those colonies which contained recombinant, P. ochraceus DNA inserted, plasmids.

Southern blot of genomic DNA and hybridization with isolated insert from pPol431.

The restriction enzyme digested genomic or cloned DNA were transferred from agarose electrophoresis gels to nitrocellulose paper by the Southern transfer method (Southern, 1975). Either long genomic DNA or specific cloned DNA sequences were digested with various restriction enzymes in the appropriate reaction buffer at 37°C. The DNAs are electrophoresed for at least 6 hours at 40V in a 0.8% or 1.0% agarose gel and photographed under uv transillumination with a ruler aligned along side the gel. If the DNA was long it was partially depurinated in the electrophoretic gel which was soaked twice for 10 minutes in a fresh solution of 0.25 M HCl. The DNA was denatured for 30 minutes in two changes of 0.5 M NaOH, 1.5 M NaCl and neutralized in 0.5 M Tris HCl, pH 7.5, and 1.5 M NaCl. A large horizontal electrophoresis apparatus was the most convenient stand to use

for the transfer blot with the gel supported in the middle and the two buffer tanks on either side containing the transfer solution, 10X SET (1X SET is 0.15 M NaCl, 0.03 M Tris HCl, pH 8.0, 1 mM Na_2EDTA). Two 3MM Whatman paper wicks, wetted in 10X SET and cut wider than the gel and long enough to extend from one buffer well to the other, were placed over the middle gel support. The neutralized gel was placed on the Whatman paper wicks and overlaid with a sheet of BA85 0.45 um nitrocellulose paper (Schleicher and Schuell), which had been wetted in double distilled H_2O followed by 10X SET. Two layers of 3MM Whatman paper, 3 inches of facial tissue and two inches of paper towels were layered on top of the nitrocellulose filter in that order. Each layer was cut so that there was no overlap of layers. A weight was placed on top of the paper towels and the system was left overnight or approximately 12 hours. Each layer was removed carefully, the electrophoretic gel well positions marked and numbered with a pencil. The nitrocellulose paper was rubbed for 10 minutes in 4X SET to remove adhering agarose. The filter was air dried and baked in a vacuum oven at 80°C for 2 hours.

The ^{32}P-labeled probe was hybridized to the Southern transferred DNA on the nitrocellulose filter in exactly the same manner as described above for the colony <u>in situ</u> hybridization screening. Generally, 2×10^6 cpm of ^{32}P-labeled probe was

added to the nitrocellulose filter hybridization solution. The filters were first prehybridized for four hours in the 4X SET solution containing 50 ug/ml sheared, denatured E. coli DNA at 68°C. The probe was heat denatured and added to the prehybridization mixture. The filters were hybridized for 18 to 24 hours at 68°C. After hybridization the filters were washed as described above. The hybridized DNA containing filters were air dried and exposed to preflashed X-ray film, with enhancer screens at -80°C. Exposure time was from 2 hours to 7 days.

Dot Blot hybridization of egg RNA from P. ochraceus with ^{32}P-labeled pPol431 insert.

5, 10 or 20 ug of oocyte RNA from P. ochraceus, 3 ug pPol431 and 3 ug lambda Dm40-1 DNA were placed in solutions containing 40% glyoxal, 50% dimethyl sulfoxide, 0.01 M Na phosphate buffer, pH 7.0 in a final volume of 10 ul (Thomas, 1980). The samples were incubated for 1 hour at 50°C, placed on ice, and spotted onto nitrocellulose filter paper discs that had been wetted in double distilled H_2O and soaked 1 hour in 20X SET. The spotted filters were baked for 2 hours at 80°C in a vacuum oven. The filters were prehybridized at 42°C for 8-20 hours in 50% formamide v/v, 5X SET, 50 mM Na phosphate buffer, pH 6.5, 0.02% Denhardt solution and 250 ug/ml sheared, denatured salmon sperm or E. coli DNA. The hybridization buffer contained 4 parts

prehybridization buffer and 1 part 50% (w/v) dextran sulfate to which was added 1 x 10^6 cpm of denatured ^{32}P-labeled pPol431 insert DNA. The filters were hybridized with probe for 20 hours at 42°C. The blots were washed for 5 minutes four times with 2X SET, 0.1% SDS at room temperature and washed twice at 15 minutes in 0.1X SET and 0.1% SDS at 50°C. The dot blots were exposed to preflashed BBl Kodak film with intensifying screens at -80°C for 2 hours to 4 days.

Reassociation and hybridization of DNA in solution.

Genomic DNA from each of five sea star species was sheared in the Virtis 60 homogenizer at 50,000 rpm in 33 ml of 66% glycerol, 0.02 M Na Acetate in a dry-ice and ethanol bath (Britten et al., 1974). The sheared DNA was made 0.3 M in Na Acetate and precipitated with two volumes of 95% ethanol at -20°C. The DNA was centrifuged at 10000 rpm for 30 minutes at 0°C. The pellet was resuspended in 0.3 M Na Acetate, passed through an equilibrated Chelex 100 column to remove divalent cations and reprecipitated with 2 volumes of ethanol. The final DNA pellet was resuspended in 0.12 M Na phosphate buffer (equimolar Na mono-and di-basic phosphate), pH 6.8. The final concentration of phosphate buffer was determined by refractometry. The DNA concentration was determined by absorbance at 260 nm DNA concentration. The single strand DNA

fragment length was determined on a 1.5% alkaline agarose gel prepared in 30 mM NaOH, 2 mM Na_2EDTA which was also the running buffer (Sharp et al., 1973; McDonell, 1977). The gel was neutralized in 1 M Tris HCl, pH 7.5, stained with 2 ug/ml ethidium bromide and photographed. The average length of sheared DNA was determined by reference to coelectrophoresed standards of known length.

DNA reassociations and hydroxyapatite (HAP) chromatography were done by standard methods as described by Britten et al. (1974). The genomic driver DNAs prepared as outlined above and ^{32}P-labeled insert from pPol431 were mixed in a ratio of 1 to 2×10^5 ug of driver to 1 ug of tracer DNA. The amount of tracer was determined from the specific activity, usually 10^7 cpm per ug. Aliquots of the ^{32}P-labeled DNA and excess genomic DNA, generally containing about 2000 cpm, were denatured in a boiling water bath for 10 minutes and immediately placed in a temperature controlled water bath at 60°C or 50°C to reassociate. At various times samples were removed and immediately frozen in liquid nitrogen. These reaction mixtures were stored at -20°C for subsequent hydroxyapatite chromatography to determine the amount of duplex formation.

Hydroxyapatite chromatography was used to determine the amount of reassociation which had occurred between driver and tracer

DNA in individual samples. Hydroxyapatite at 60°C in 0.12 M Na phosphate will selectively bind double stranded DNA fragments and let single strand DNA fragments elute. The amount of reassociated ^{32}P-labeled DNA which binds to HAP at 60°C and 0.12 M Na phosphate was measured by scintillation counting. The product of the time of reassociation in secs and the concentration (moles/litre) of driver DNA in the reassociation mixture (Cot) was plotted versus fraction unreassociated tracer DNA. The reassociation data were analyzed using a computer-programmed least-squares fitting procedure developed by R.J. Britten and revised by Pearson et al. (1977). With this program, the number, size, and second-order reaction rate of up to 10 sequence frequency components in the DNA population can be determined. The operator has the option of holding constant particular parameters such as component size, rate, or total extent of reassociation.

For hydroxyapatite thermal chromatography, reassociated DNA was loaded onto hydroxyapatite columns at 50°C in 0.12 M Na phosphate buffer. The temperature of the column was raised in 3° or 5° increments to 100°C. At each increment single stranded DNA was eluted with 3 x 1 ml washes of 0.12 M Na phosphate, 0.06% SDS. The amount of ^{32}P-labeled DNA at each temperature interval was determined by scintillation counting. The denaturation temperature, T_m, is defined as the

temperature at which half the duplex DNA has been eluted from the hydroxyapatite column.

The thermal denaturation temperature of molecular duplex hybrids between sheared genomic DNA and labeled insert DNA fragments from pPol431 can be used to ascertain the degree of base sequence mismatch of such duplexes both intra- and interspecifically. The concentration of salt and temperature during reassociation affects the stability of the duplex. The standard reassociation conditions of 0.18 M Na^+, 60°C and pH 6.8 were used as well as a lower temperature of 50°C in order to obtain reassociated fragments which were not stable at 60°C and thus would represent more divergent members of the family. Furthermore, reassociation in 0.3 M NaCl, 0.01 M Pipes, pH 6.7 and 64°C which is equivalent to 0.12 M Na phosphate at 60°C was also done for samples to be digested with S1 nuclease. This enzyme digests only single stranded tails at an appropriate salt concentration and ratio of DNA to enzyme is used (Smith et al., 1975). An assay of the appropriate enzyme to substrate ratio was done prior to the reassociation experiment. After DNA reassociation the samples were adjusted to the correct salt conditions for single stranded DNA digestion by S1 nuclease: 0.15 M NaCl, 0.005 M Pipes, pH 6.7, 0.025 M Na Acetate, 0.0004 M $ZnSO_4$, 0.025 M beta-mercaptoethanol. The duplex DNA was incubated at 37°C for 1 hour with 6 units of S1 nuclease

(Boehringer Mannheim) per ug of DNA. Sheared carrier E. coli DNA was added to each reaction. One unit of S1 nuclease has been defined (Boehringer Mannheim) as the enzyme activity which catalyzes the formation of 1 ug acid soluble deoxynucleotides after 30 minutes incubation with denatured DNA at 37°C under assay conditions.

After S1 nuclease digestion an aliquot of the reactions was electrophoresed on 2% agarose gels to determine the duplex fragment length. Undigested but reassociated P. ochraceus driver DNA was coelectrophoresed with these samples. The electrophoretic gels were dried and autoradiographed as previously described.

To determine the fraction of ^{32}P-labeled pPol431 insert DNA which was fold-back sequence an aliquot of the tracer was boiled and immediately quenched in liquid nitrogen. The amount of fold-back DNA was assayed by hydroxyapatite chromatography. The T_m of fold-back was determined by thermal elution from hydroxyapatite.

For other thermal denaturation experiments the fold back DNA sequence was stripped from the probe in the following manner. About 10^6 cpm of ^{32}P-labeled pPol431 insert was boiled and quenched in liquid nitrogen. The sample was thawed and passed

over a hydroxyapatite column at 60°C. The single stranded eluate was used as probe for reassociation with driver DNAs. An aliquot of this DNA was sized on a 1.5% alkaline agarose gel. The fold-back, stripped insert DNA was reassociated with starfish driver DNAs from all five species in 0.12 M Na phosphate at 60° and 50°C. The hybrid duplexes were thermally denatured on hydroxapatite columns.

Isolation of recombinant phage DNA from a Charon 4 lambda Dm40-1 clone.

A lambda genomic clone (Dm 40-1) containing Drosophila 18S and 28S ribosomal DNA sequences in Charon 4 phage cloning vector (Blattner, 1977) was a gift from N. Davidson. Isolated phage were lysed in 0.1% SDS and 50 mM Na_2EDTA at 50°C for 10 minutes. The DNA was extracted with phenol:SEVAG (1:1,V:V), saturated with 10 mM Tris HCl buffer, pH 8.0. The extraction was repeated and the phenol removed with repeated ether extractions. The residual ether was evaporated at 50°C for 5 minutes. As a final purification step, the DNA was precipitated in 0.5 M NH_4 Acetate, 20 mM $MgCl_2$, 0.2 mM Na_2EDTA and with 1 volume of isopropanol at -20°C. The DNA was centrifuged at 10000 rpm for 1 hour at -10°C and the pellet washed with ice cold 80% ethanol. The sample was centrifuged again at 20000 rpm for 10 minutes and the supernatant discarded. The 80% wash and

subsequent centrifugation was repeated two more times. Finally the pellet was lyophilized and resuspended in 50 ul of 10 mM Tris HCl, pH 7.5 and 0.1 mM Na_2EDTA.

The phage DNA was digested with HindIII restriction enzyme in a 30 ul reaction mixture at 2 units of enzyme to 1 ug of phage DNA. The digest was electrophoresed at 45V for 6 hours in a 1% gel and then Southern blotted onto nitrocellulose filter paper. The filters were hybridized with 1 x 10^6 cpm of ^{32}P-labeled insert as previously described and then exposed for 2 to 48 hours to preflashed X-ray film using enhancer screens at -80°C.

Selection of P. ochraceus Charon 4 phage clones with ^{32}P-labeled insert.

Genomic DNA from P. ochraceus had been previously cloned into Charon 4 lambda phage (Blattner, 1977). The phage titre of the library had been determined to be 1 x 10^{10} pfu/ml. 50 ml of fresh E. coli K802 bacterial culture was centrifuged at 10000 rpm for 15 minutes and resuspended in 0.8 ml of SM (0.1 M NaCl, 0.01 M Tris HCl, pH 7.5, 0.01 M $MgSO_4$, 0.02% gelatin). 100 ul of the resuspended bacteria was infected with 10 ul of diluted phage at 10^6 pfu/ml. The mixture was heated at 37°C for 15 minutes, added to 7.5 ml of 0.7% agarose in NZCYM (1% NZamine,

0.5% NaCl, 0.5% yeast extract, 0.1% casamino acids, 10 mM $MgSO_4$) (Maniatis et al., 1978), mixed and poured onto 15 cm agar plates containing 1.2% agar in NZCYM. The plates were incubated at 37°C for 14 hours.

To screen the genomic library for positive clones containing repeat DNA, in situ hybridization with labeled pPol431 isolated insert was used (Benton and Davis, 1977). Nitrocellulose filter paper discs were layered onto cooled plates containing the plaques and the position of the plaques oriented by stabbing the filter and agar with a needle containing india ink. The filters were peeled off and laid onto a solution of 1.5 M NaCl, 0.1 M NaOH for 1 minute, then 1.5 M NaCl, 0.5 M Tris HCl, pH 7.5, for 1 minute, rubbed clean of adherent agarose for 1 minute, and soaked in 4X SET for approximately 5 minutes. The filters were air dried and baked at 80°C for 2 hours in a vacuum oven.

^{32}P-labeled insert from pPol431 was hybridized to 6 filters, each containing 2500 plaques, using the same procedures as described for Southern blot hybridizations. 30 ml of hybridization solution and 6×10^6 cpm were added to the filters and hybridized overnight at 65°C. The discs were autoradiographed and positive spots on the film were aligned to the correct plaques.

Determination of genomic methylation sites in P. ochraceus, P. brevispinus and Dermasterias repeat families.

Long genomic DNA from the starfish species, P. ochraceus, P. brevispinus and Dermasterias were each digested with the restriction enzymes MspI and HpaII using the reaction conditions specified by the supplier at 4 units of enzyme to 1 ug of DNA. A 30 ul reaction mixture contained 5 ug of DNA and 20 units of enzyme and was incubated at 37°C overnight. The digests were electrophoresed on 2% agarose gels and then Southern blotted as previously described. ^{32}P-labeled insert was hybridized to the blots and the filters were autoradiographed. Both MspI and HpaII recognize the same restriction site (5'-C/CGG-3') but HpaII cannot cleave the DNA if the cytosine residue is methylated in the 5'-CG-3' location. However, MspI will cleave the DNA even with the methyl group.

RESULTS

Cloning of the P. ochraceus 3.9 kb BamHI repeat fragment.

Long genomic Pisaster ochraceus DNA was digested with the restriction endonuclease BamHI and the fragments electrophoresed on horizontal agarose gels. The ethidium bromide stained gels show that the genome is cleaved into many fragments of various

lengths creating a smear in the gel. This suggests that most enzyme sites are randomly located about the genome and upon digestion a Poisson distribution of fragment sizes is generated. In the genomic DNA digests there are however BamHI sites that occur with some regular periodicity producing bands of fragments of uniform length which can be seen upon gel electrophoresis (Figure 1). A distinctive band in an electrophorogram of a total BamHI digest of genomic DNA presumably indicates a repetitive DNA fragment. There is a distinctive 3.9 kb band found in a P. ochraceus genomic digest by BamHI.

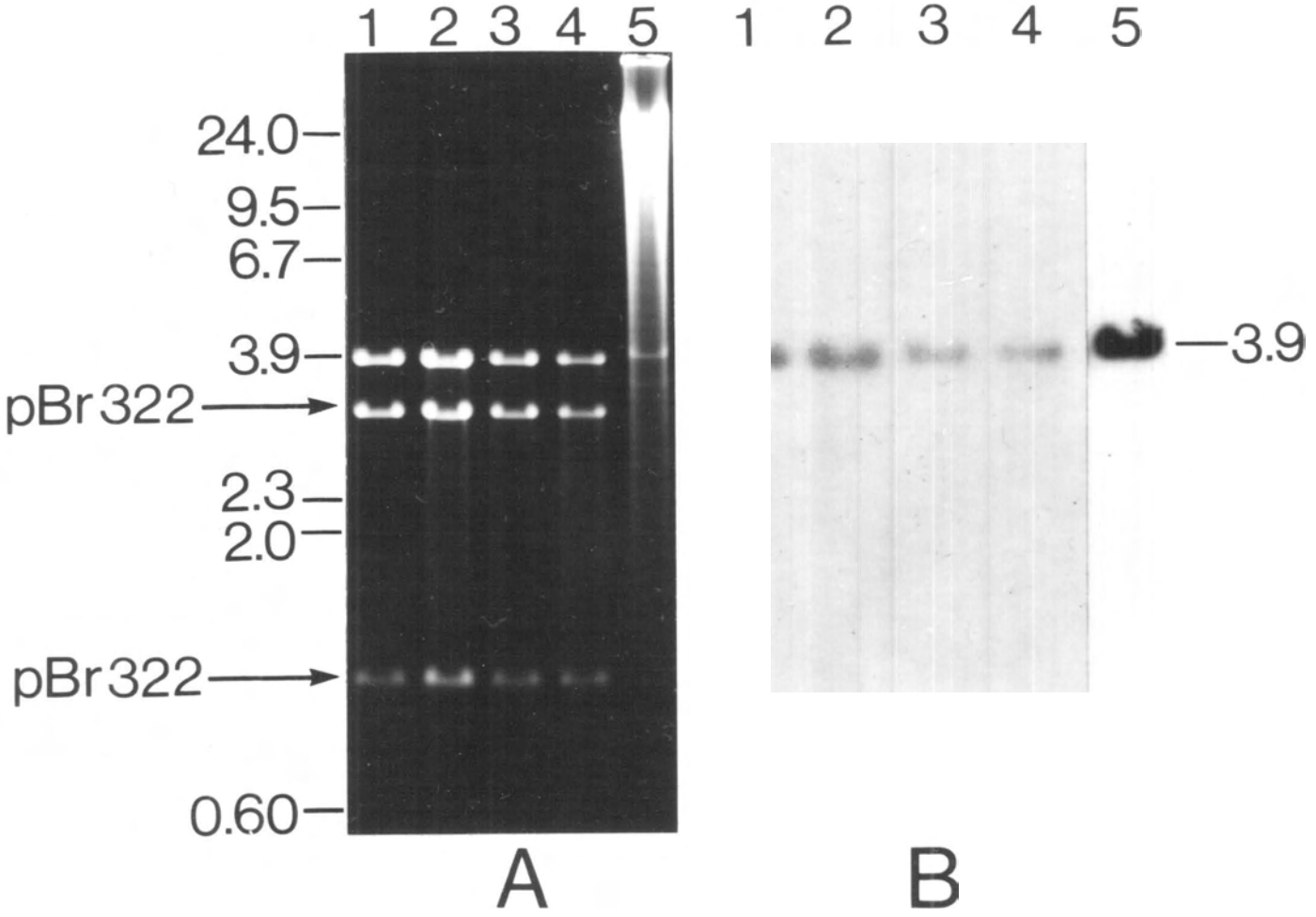
Selection of four positive clones.

The 3.9 kb P. ochraceus DNA band was isolated from preparative electrophoretic gels and ligated into BamHI cut pBR322. E. coli χ1776 was transformed with this ligation mixture. The colonies which could grow on ampicillin but not tetracycline were screened using the Thayer in situ hybridization technique (Thayer, 1979). These colonies were grown on nitrocellulose filter paper, lysed and then the DNA was denatured in situ. The single stranded DNA binds irreversibly to the nitrocellulose. The colonies were ordered in microtitre wells so that positive colonies could be retrieved. ^{32}P-labeled 3.9 kb genomic band excised from agarose gels was hybridized to the filters overnight at 5 x 10^5 cpm per filter and the filters were

Figure 1. Selected repetitive DNA clones each contain a member homologous to a 3.9 kb fragment observed in Pisaster ochraceus genomic DNA.

(A) Ethidium bromide stained agarose gel electrophoresis of genomic and cloned DNA. Lanes 1,2,3 and 4 contain PstI and BamHI double digests of selected recombinants, pPo328, pPol289, pPol431 and pPol675 respectively. Note that each contains a 3.9 kbp insert fragment. Lane 5 is a total digest of P. ochraceus genomic DNA with BamHI restriction endonuclease.

(B) The restriction enzyme digested DNA was Southern transferred from the agarose gel in (A) to nitrocellulose filter paper and hybridized with ^{32}P-labeled isolated insert from pPol431. The autoradiogram of hybridized probe to DNA blots shows reaction with genomic DNA (Lane 5) (24 hour exposure) as well as with Pst I and Bam HI double digests of pPo328, pPol289, pPol431, and pPol675 (lanes 1,2,3, and 4 respectively) (2 hour exposure). The autoradiogram demonstrates that the pPol431 insert hybridizes to the 3.9 kbp genomic band as well as clones of pPo328, pPol289, pPol431 and pPol675.



35B

exposed for 4 hours. There were 9 recombinant clones obtained from 700 positives screened which is a 1.3% transformation efficiency. The positive cultures were transferred from the appropriate microtitre well to a large overnight culture. Four strong positives were selected and designated pPo328, pPol289, pPol431 and pPol675. These plasmids were amplified and their DNA isolated. To demonstrate that each contained a homologous member of the repetitive DNA family an insert probe was prepared from pPol431. 20 ug of pPol431 plasmid DNA was digested with 40 units of PstI and 40 units of BamHI restriction enzymes in PstI reaction buffer and a volume of 100 ul at 37^{o}C for 2 hours. The PstI plus BamHI double digest ensures that the insert fragments would be distinguishable from pBR322 vector fragments. The sample was loaded onto a 1.0% agarose gel, electrophoresed at 40V until the insert was clearly separated in the gel from pBR322 DNA. PstI does not cleave the insert but it does cleave pBR322 at 3.2 kb from the BamHI site so that pBR322 DNA can be easily separated from the insert. The insert DNA was cut from the agarose gel and electroeluted as described above. This isolated insert DNA was ^{32}P-labeled by nick translation.

The DNA from each plasmid was digested with PstI and BamHI and electrophoresed in the same gel with P. ochraceus total genomic DNA digested with BamHI. The electrophoretic gel was Southern transferred to nitrocellulose which was hybridized with

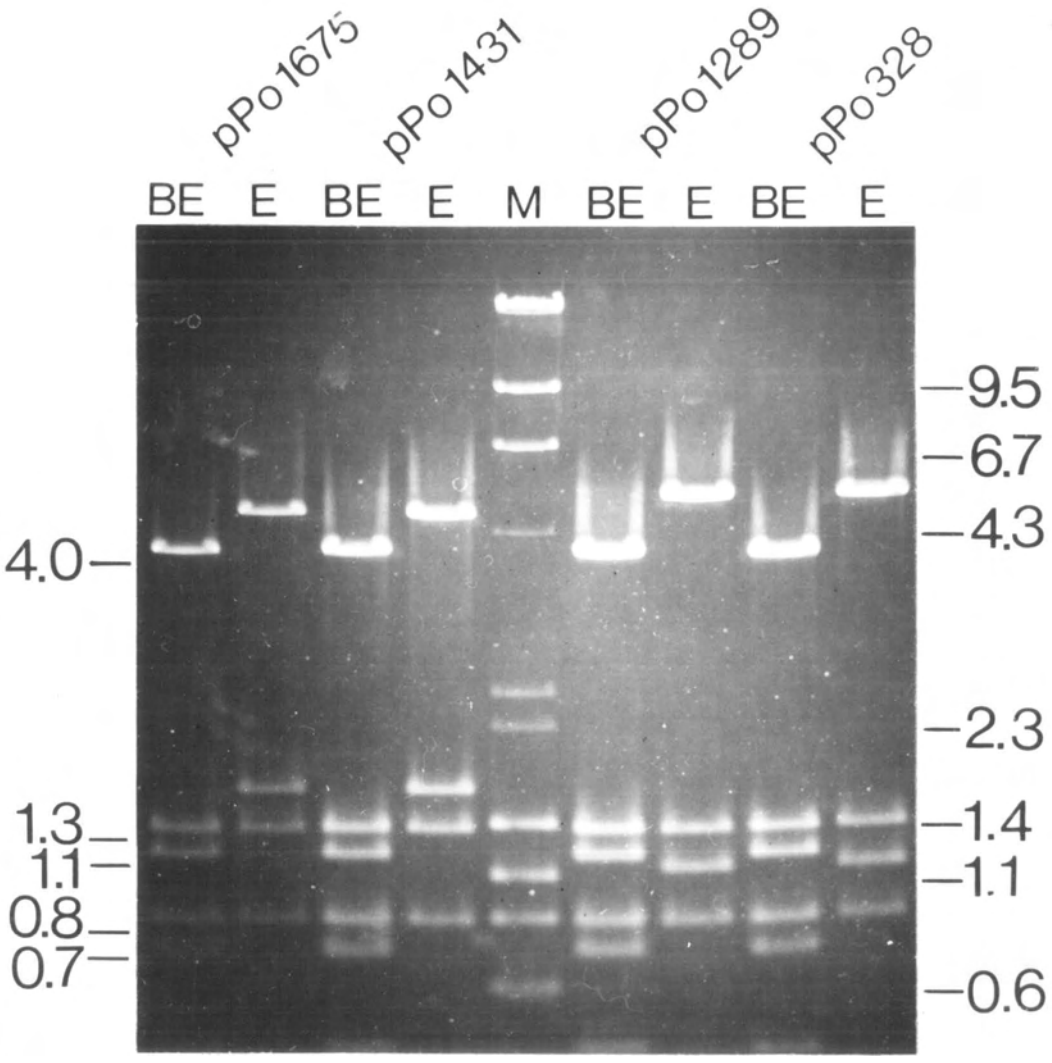
^{32}P-labeled pPol431 insert. It can be seen in Figure 1 that both the 3.9 kb band in the genomic digest and the four cloned inserts hybridized strongly to the labeled repeat from pPol431.

Restriction enzyme mapping of the Pisaster repetitive DNA clones.

Restriction enzymes cleave double stranded DNA by recognizing a specific nucleotide sequence in the DNA and then hydrolyzing the phosphodiester backbone. Different enzymes recognize different specific sequences. An efficient method of analyzing the relationships of different DNA fragments such as the clones that have been isolated here, is to digest them with various enzymes, electrophorese the digestion products, and compare sizes of the fragments generated. Figure 2 shows the results of four clones digested by EcoRI and by both EcoRI and BamHI. It can be seen that the double digest generates bands that are identical in all four clones yet the EcoRI digestion alone produces bands which are different in pPo328 and pPol289 from those in pPol431 and pPol675. However, there are two common bands in all four EcoRI digests. The two common EcoRI bands within all plasmid inserts are 1.3 and 0.8 kb long. The pPo328 and pPol289 EcoRI end fragments of the insert plus the pBR322 DNA are 5.0 and 1.1 kb in length. pPol431 and pPol675 EcoRI digestion produces 4.7 and 1.5 kb bands that also include pBR322 DNA and the EcoRI end

Figure 2. Repeat DNA insert orientation in the four recombinant plasmids.

Four chimeric plasmids containing 3.9 kbp inserts were digested with Eco RI (E) or with Bam HI and Eco RI (BE) restriction endonucleases. The fragments generated were separated by electrophoresis a 1.0% agarose gel at 50V for 5h and stained with ethidium bromide. (Lane 1) pPol675 (BE); (Lane 2) pPol675 (E); (Lane 3) pPol431 (BE) ; (Lane 4) pPol431 (E); (Lane 5) Markers, lambda DNA digested with Hind III and ϕX174 DNA digested with Hae III; (Lane 6) pPol289 (BE); (Lane 7) pPol289 (E); (Lane 8) pPo328 (BE); (Lane 9) pPo328 (E). The same sized fragments are generated upon Eco RI and Bam HI digestion of all the plasmids.


pPo1675
pPo1431
pPo1289
pPo328
BE E BE E M BE E BE E
9.5
6.7
4.3
4.0
2.3
1.3
1.1
1.4
1.1
0.8
0.7
0.6

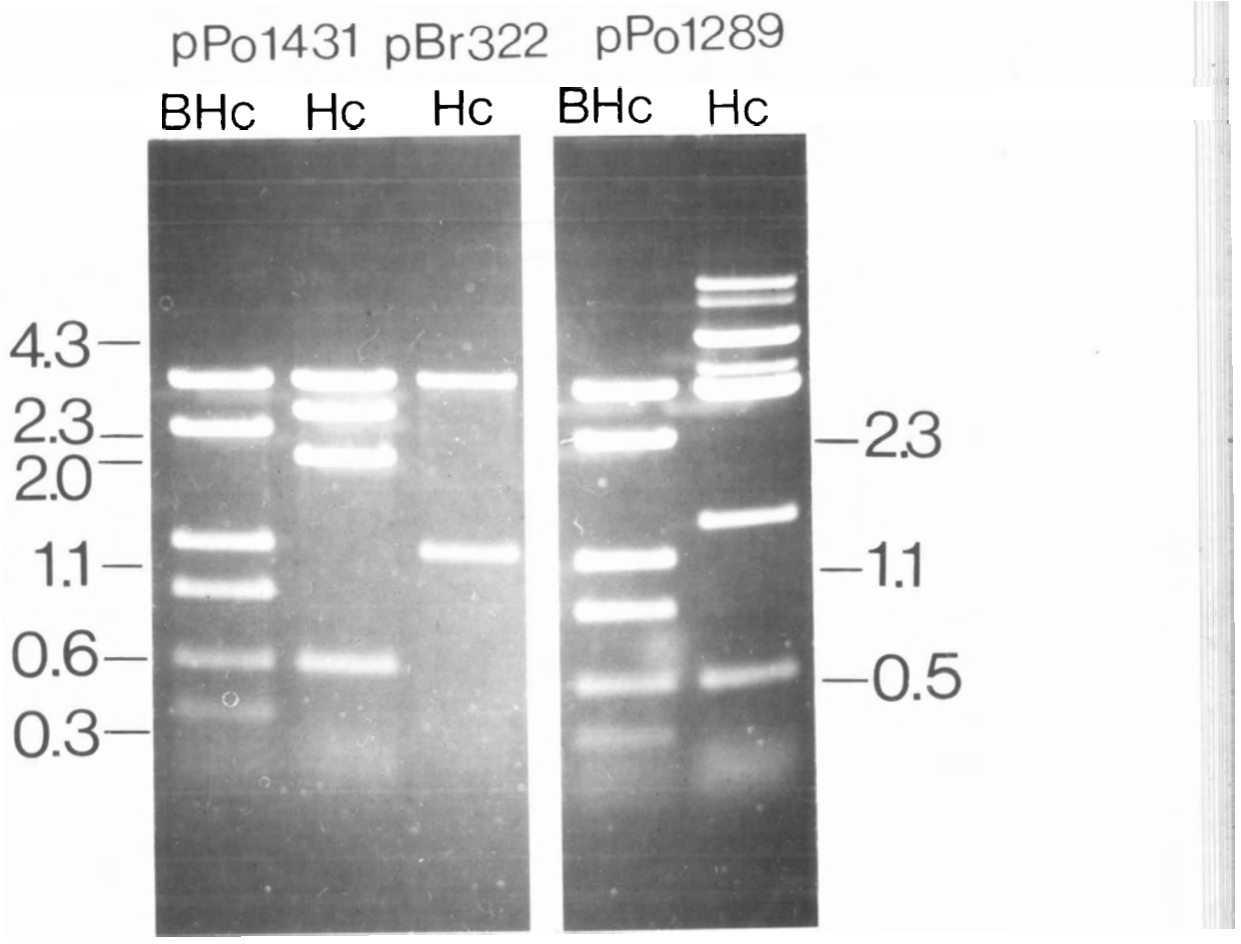
fragments of the insert. The lengths of the EcoRI/BamHI bands in all four clones are 1.1, 1.3, 0.8 and 0.7 kb. This suggests that all four clones contain copies of the same P. ochraceus segment but this segment is inserted in reverse orientation in the case of pPol431 and pPol675 from that found in pPo328 and pPol289.

To clarify the question of insert orientation in the plasmids, pPol289 and pPol431 were digested with HincII which cuts at 5'-GTPyPuAC-3' and with both HincII and BamHI (Fig. 3). Double digestion of both plasmids produced 2250, 1120 and 530 bp fragments whereas upon HincII digestion of pPol431 produced 530, 1950 and 2530 bp fragments and pPol289 produced 530, 1400 and 3100 bp fragments. A pBR322 fragment, 3250 bp long, was also observed in the HincII digests. The 2250 and 1120 bp segments of the inserts are attached to the vector pBR322 so that their order relative to the HincII sites in pBR322 can be assessed from the sizes of the fragments produced upon HincII digestion (Figure 3).

The insert fragments in the four clones are not exactly identical. When these chimeric molecules are digested with HaeIII which recognizes 5'-GG/CC-3' sequences and the fragments electrophoresed on a 5% polyacrylamide gel, the pPo328, pPol289 and pPol431 display distinct band patterns from one another

Figure 3. Determination of the HincII restriction enzyme sites in pPol431 and pPol289 and insert orientation.

Both pPol431 and pPol289 were digested with HincII or with HincII and BamHI in order to determine the HincII restriction enzyme sites in the recombinant DNA and to verify that the insert in pPol289 is in the reverse orientation to that of pPol431. After digestion, the fragments were separated on a 0.8% agarose gel and stained with ethidium bromide. HincII sites are located at 1120 bp and 1650 bp in pPol431 and are at 2250 and 2780 bp in pPol289 from the 5' end that is closest to the EcoRI site in pBR322. Note that digestion of pPol289 with HincII is incomplete.


pPo1431
pBr322
pPo1289
BHc
Hc
Hc
BHc
Hc
4.3
2.3
2.0
1.1
0.6
0.3
2.3
1.1
0.5

whereas pPol675 was identical to pPol431 (Figure 5). There are 8 different bands between pPol289 and pPol431 (two are probably due to reverse orientation of the insert), 4 different bands between pPo328 and pPol41 (2 are due to reverse orientation) and 5 different bands between pPo328 and pPol289 which have the same orientation. Thus there is some sequence diversity in the repeat family represented by these four clones and shown by HaeIII restriction endonuclease digestion.

Restriction Enzyme Mapping of pPol431 and pPol289.

Figures 4 and 6 show the order of the restiction enzyme sites found in pPol431 and pPol289 inserts. The two plasmids were digested with BamHI, EcoRI, HincII or AvaI as previously described and the fragments were electrophoresed on 0.8% to 1.0% agarose gels. Using internal gel markers the sizes of the fragments were determined. AvaI and HincII could be ordered directly since there were only two internal sites in the insert and the ends could be determined from the known restriction sites in pBR322. However, EcoRI sites had to be ordered using partial digestion of plasmid DNA. 3 ug of pPol431 and pPol289 were digested with 10 units of EcoRI in a total reaction volume of 50 ul and the reaction mixtures were incubated at 37^{o}C. 10 ul aliquots were removed at 4, 8, 10, 15 and 30 minutes of incubation and added to the stop mix and heated at 65^{o}C for 5

Figure 4. Structure of recombinant plasmids, pPol431 and pPol289.

Recombinant plasmids pPol431 and pPol289 represent the opposite orientation of the P. ochraceus repeat DNA insert. The 3.9 kb P. ochraceus BamHI restriction fragments were inserted into the pBR322 BamHI cleavage site situated in the tetracycline resistant gene of the plasmid pBR322.

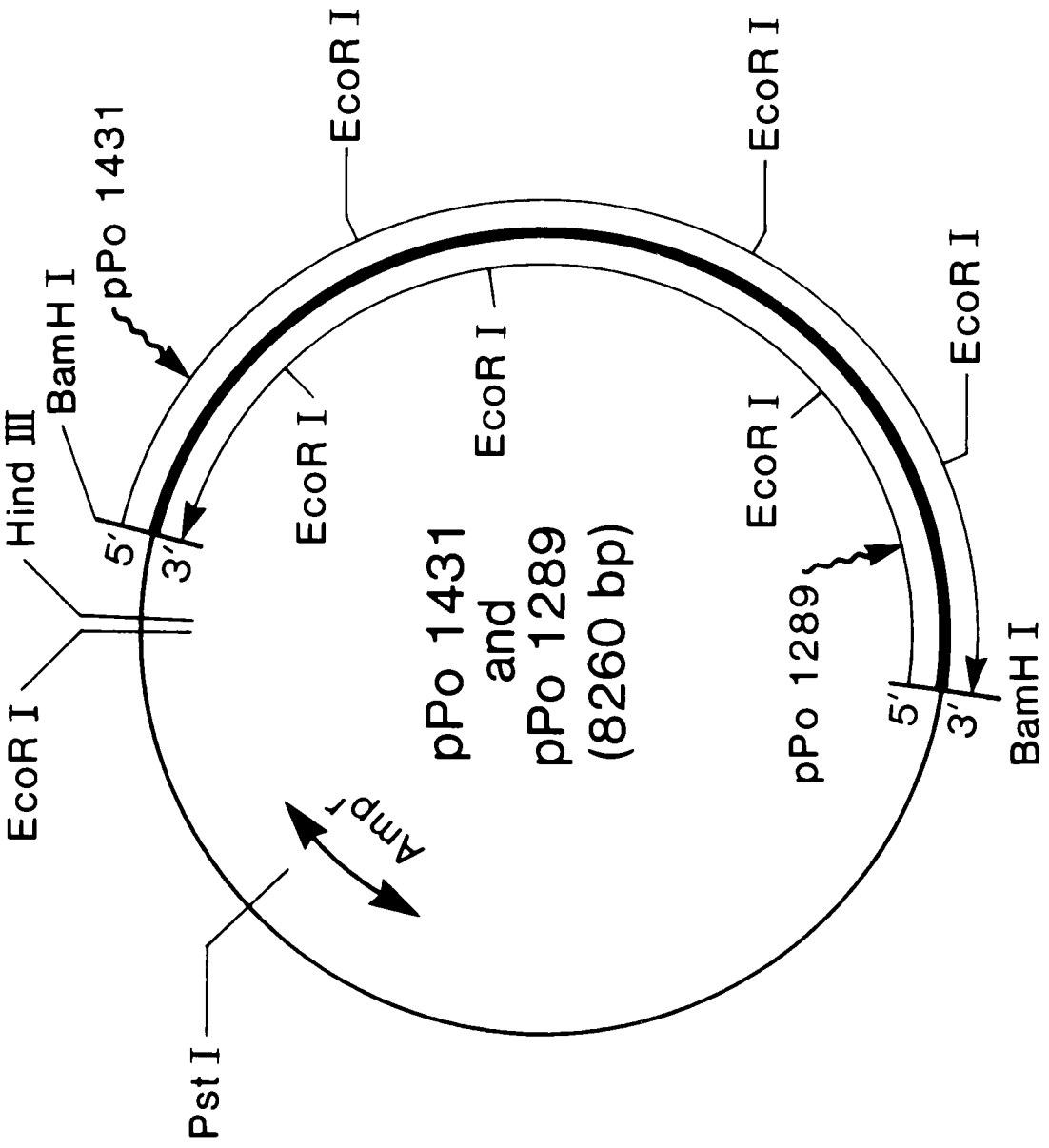

pPo 1431
and
pPo 1289
(8260 bp)
EcoR I
EcoR I
EcoR I
EcoR I
EcoR I
EcoR I
Hind III
BamH I
pPo 1431
5′
3′
pPo 1289
5′
3′
BamH I
Pst I
Amp r

Figure 5. HaeIII restriction enzyme digestion of plasmids pPo328, pPol289, pPol431 and pPol675.

Four recombinant plasmids were digested with HaeIII restriction endonuclease and electrophoresed on a 5% polyacrylamide gel for 1.5 hours at 100 V. The bands were observed by staining with 1 ug/ml ethidium bromide. (Lane 1) Marker, ϕX174 DNA digestd with Hae III; (Lane 2) pPo328; (Lane 3) pPol289; (Lane 4) pPol431; (Lane 5) pPol675; (Lane 6) pBR322 digested with HaeIII; (Lane 7) Marker, ϕX174 DNA digested with AluI. Sequence divergence of the insert DNA is demonstrated by the variation in the restriction patterns. Band distribution exceeds that expected from distinct plasmid/insert junctions generated by reverse orientation of inserts (see Fig. 4).

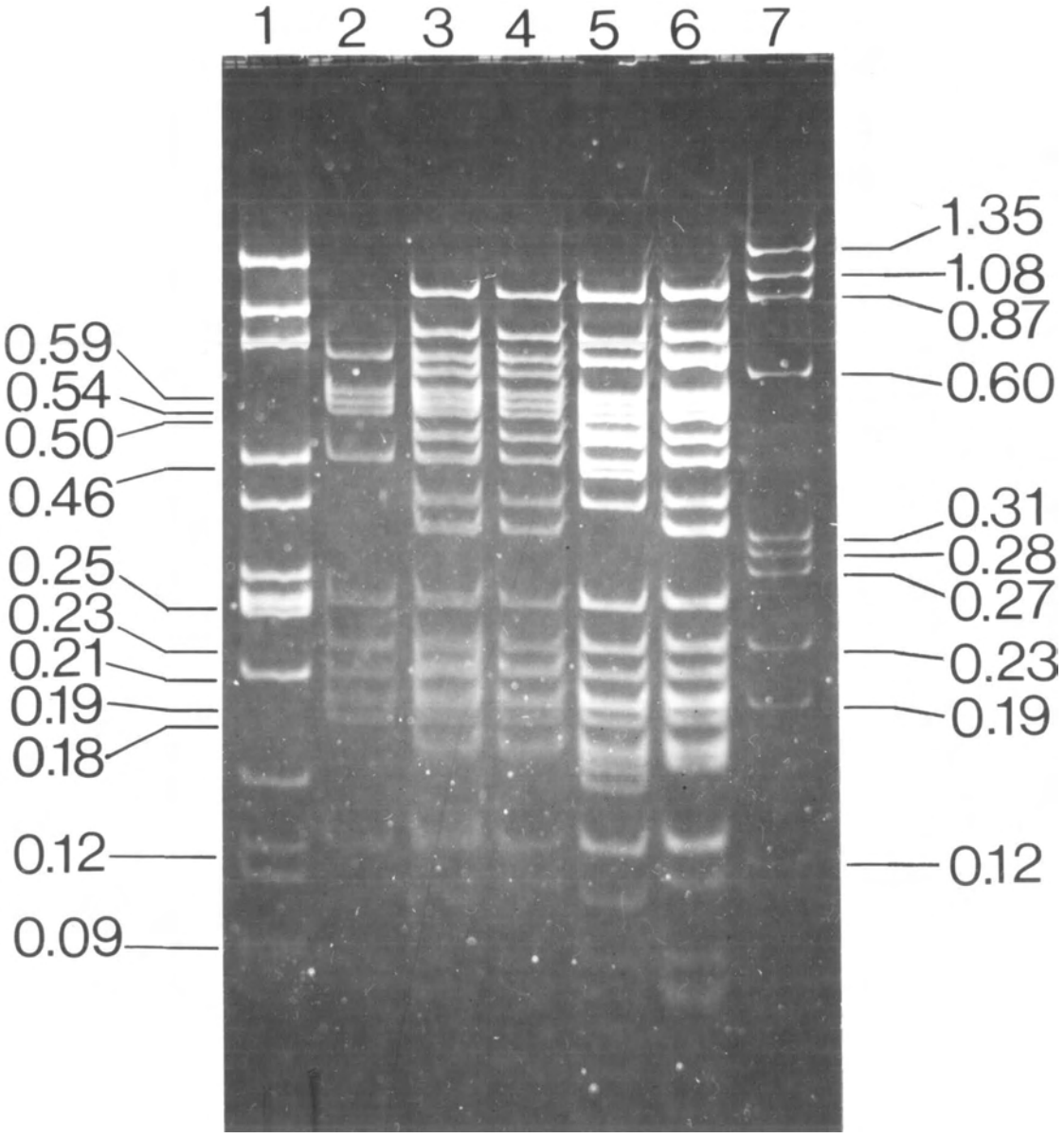
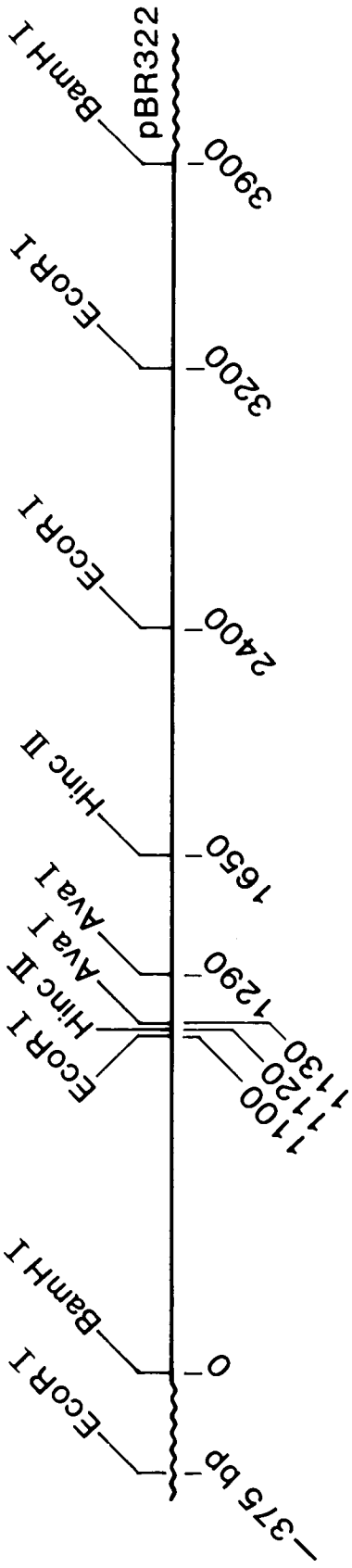

1 2 3 4 5 6 7
0.59
0.54
0.50
0.46
0.25
0.23
0.21
0.19
0.18
0.12
0.09
1.35
1.08
0.87
0.60
0.31
0.28
0.27
0.23
0.19
0.12

Figure 6. Detailed restriction enzyme map of the insert from pPol431.

The map was determined by digesting the recombinant plasmid with various restriction enzymes and the sizes and order of the fragments generated were estimated by agarose gel electrophoresis. The gels were calibrated with DNA fragments of known molecular weight. The positions of the restriction enzymes sites are indicated as the number of nucleotide base pairs from the BamHI insertion site of the plasmid pBR322. The EcoRI site of pBR322 is 375 bp upstream from the 5' end of the insert represented by pPol431.


EcoR I
BamH I
0
375 bp
pBR322
BamH I
3900
EcoR I
3200
EcoR I
2400
Hinc II
1650
Ava I
Ava I
1290
Hinc II
EcoR I
1100
1120
1130

minutes. The partials were sized on 0.6% agarose gels. The sizes of fragments expected for a given arrangment of the two EcoRI internal fragments of the insert were predicted and then compared to those sizes observed in the gel. The restriction map of pPol431 is shown in Figure 6. The order of the EcoRI fragments internal to the BamHI site is 1.1, 1.3, 0.8 and 0.7 kb.

Sequence organization of the repeat pPol431 family in Pisaster ochraceus and Dermasterias genomic DNA.

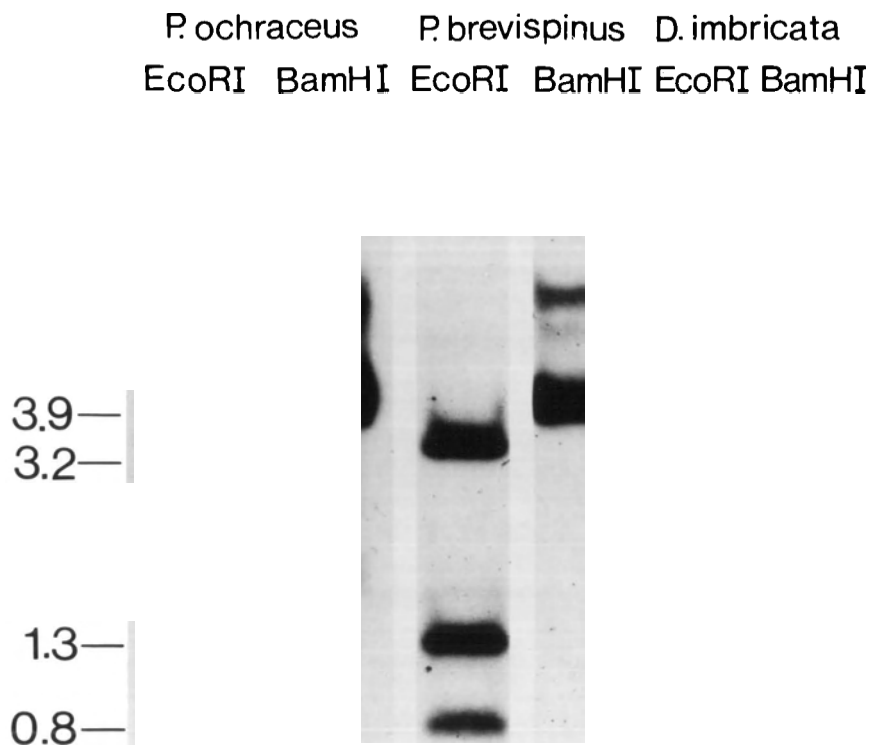
Genomic DNA digested by restriction endonucleases and electrophoresed in agarose gels were Southern transferred to nitrocellulose filter paper and hybridized with ^{32}P-labeled insert from the repeat clone, pPol431. The filters were washed and then exposed for various lengths of time to X-ray film. If the repeat family is clustered in the genome then discrete bands, of high intensity, will appear in the autoradiogram because the restriction enzyme site would occur in a periodic arrangement. If the repeat elements were totally dispersed then for each repeat there would probably be a distinct restriction enzyme fragment containing the repeat and it would appear as a unique autoradiogram band. Since there are on the order of 500 copies of the repeat family in the P. ochraceus genome (see below) then there would be 500 different bands reacting with the

probe producing a smear in the autoradiogram after a long exposure.

P. ochraceus, P. brevispinus and Dermasterias genomic DNA was digested with EcoRI or BamHI. The fragments were electrophoresed in 0.8% agarose gels, Southern transferred and hybridized with 1 x 10^6 cpm per species of ^{32}P-labeled insert from the repeat pPol431 clone. Figure 7 shows that after 48 hours exposure there is an intense band in the BamHI digests at 3.9 kb in P. ochraceus and P. brevispinus and a weaker band at 5.9 kb in Dermasterias. The appearance of one band suggests that most members of the repeat family are clustered. This experiment however does not address the question of the number of such clusters which appear in the genome. The lengths of the EcoRI bands are 3.2, 1.3 and 0.8 kb in the two Pisaster blots which accounts for the two internal EcoRI fragments in the pPol431 repeat (Fig. 6). If the repeat is clustered in the genomic DNA then the 3.2 kb band will contain the 1.1 and 0.7 ends of the 3.9 kb insert plus another 1.4 kb connecting two 3.9 kb repeating elements. This hypothesis was verified (see below). The autoradiograph of Dermasterias digested by EcoRI shows a band, 5.9 kb long, suggesting that there is clustering of a repeat element containing sequences homologous to pPol431 insert but which is longer than that represented by the 3.9 kb insert of P. ochraceus.

Figure 7. Conservation of repeat sequences homologous to pPol431 insert DNA in Pisaster brevispinus and Dermasterias imbricata.

Approximately 8 ug each of genomic DNA from P. ochraceus, P. brevispinus or Dermasterias imbricata was digested with EcoRI and BamHI restriction endonucleases and electrophoresed for 6 hours at 40V in a 1.0% agarose gel. The DNA was transferred to nitrocellulose filter paper and hybridized with 1 x 10^6 cpm of ^{32}P-labeled insert from pPol431 per species. An intense 3.9 kb band in the Pisaster BamHI digests corresponds to the electrophoretic band observed in the ethidium bromide stained gel which was originally cloned. The internal EcoRI fragments of the repeat element, 1.3 and 0.8 kb, hybridize strongly to the probe. One other strong band is observed in both Pisaster EcoRI digests at 3.2 kb. Autoradiogram bands in Dermasterias hybridized blots are much less intense and are longer (5.9 kb) than those found in either Pisaster digest. The repeats are clustered in the genomes since only a few discrete bands are observed in the autoradiograms of the three species.

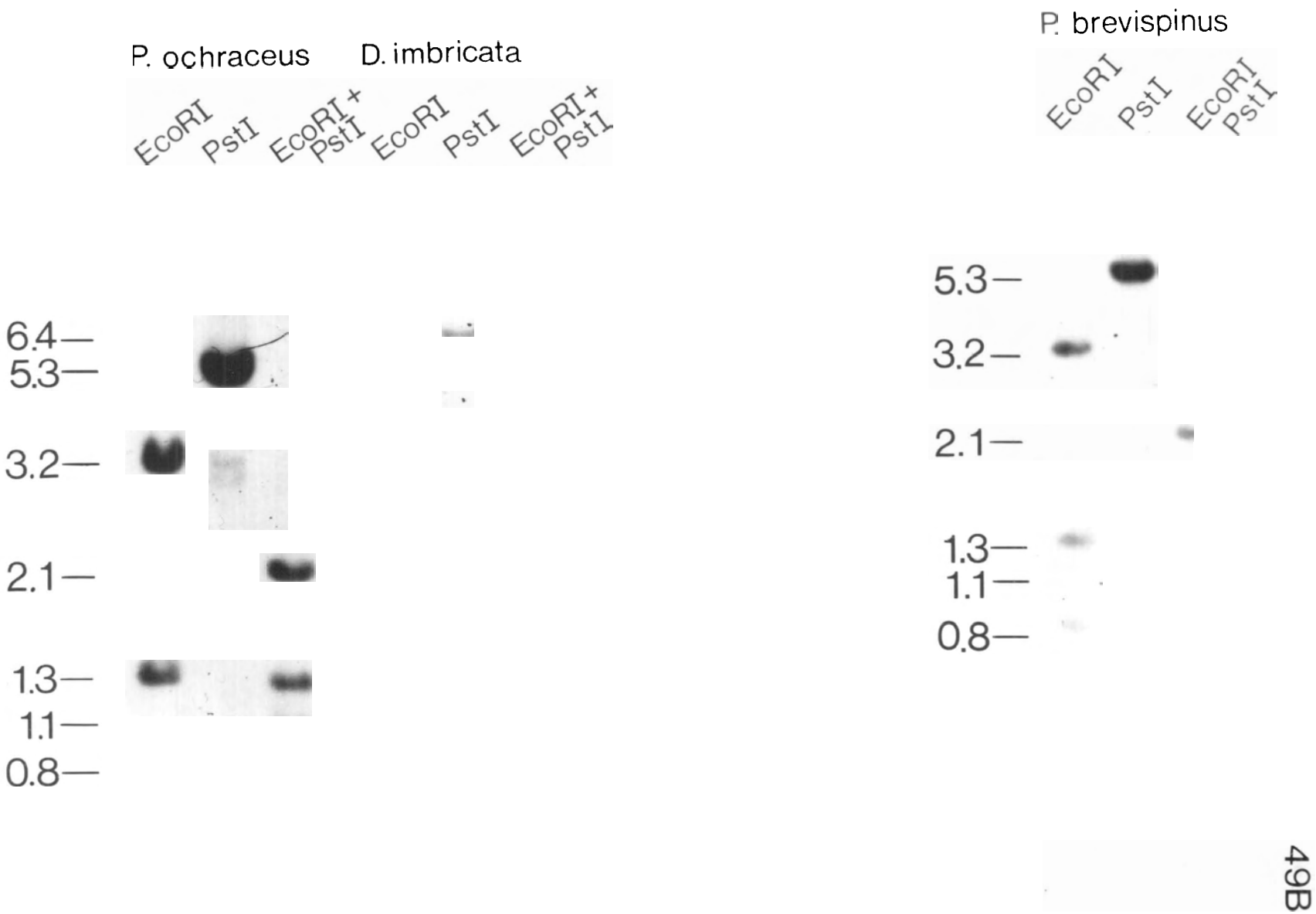

P. ochraceus
P. brevispinus
D. imbricata
EcoRI
BamHI
EcoRI
BamHI
EcoRI
BamHI
3.9
3.2
1.3
0.8

The three genomic DNAs were digested with EcoRI, PstI and with both enzymes. The electrophoresed products were Southern blotted and hybridized with ^{32}P-labeled insert from pPol431 to determine if the 3.2 EcoRI band shown in Figure 7 not only contained the terminal 1.1 and 0.7 kb but also the connecting DNA sequence in the Pisaster species. Figure 8 shows that PstI cuts the repeat DNA sequence in the genomic DNA between two 3.9 kb BamHI fragments. This generates an intense 5.3 kb band observed after only 4 hours exposure in both Pisaster species. When genomic DNA from these two species is double digested with both EcoRI and PstI, Southern transferred and hybridized with labeled insert the predicted bands appear at 2.1, 1.3, 1.1, and 0.8 kb. Dermasterias hybridized genomic blots showed an intense band at 5.9 kb and less intense band at 4.9 kb in the PstI digest. Also, another intense 5.9 kb band was observed in the EcoRI digest. Bands of fragment lengths 5.5 and 4.9 kb appeared in the double digest of Dermasterias showing that the PstI site is 0.4 kb from one of the EcoRI sites in that genome. These bands in Dermasterias are much less intense than those found in the Pisaster species but are still observed with short exposure time indicating that the bands probably represent clusters in the genome.

The minimum length of the repeat appears to be 5.3 kb in the two Pisaster genomes but this may not be the maximum length since

Figure 8. Southern blots of genomic digests of three different starfish species demonstrates that the repeat family appears clustered in all three genomes.

PstI, EcoRI or PstI and EcoRI restriction enzymes digested 5 ug of genomic DNA from P. ochraceus, P. brevispinus and Dermasterias. The fragments were electrophoresed for 16h at 15V in 0.7% agarose gels; P. brevispinus DNA was electrophoresed separately on a 1.0% gel at 15V for 16h. The DNA was Southern transferred to nitrocellulose paper and hybridized with 2 x 10^6 cpm per lane. PstI cleaves between the 3.9 kb cloned repeats of the genome generating a 5.3 kb band in the autoradiograms of P. ochraceus and P. brevispinus. Double digests generated the fragments expected if the repeat family was completely clustered. The EcoRI blots are identical to those observed in Fig. 7. The Dermasterias autoradiograms show a 5.9 kb PstI band which is reduced in size by 0.4 kb upon double digestion with EcoRI and PstI.


P. ochraceus
D. imbricata
P. brevispinus
EcoRI
PstI
EcoRI+ PstI
6.4
5.3
3.2
2.1
1.3
1.1
0.8

49B

there could be two PstI sites between the 3.9 kb BamHI elements. To exclude that possibility, Pisaster ochraceus genomic DNA was partially digested with PstI and BamHI, the DNA Southern transferred and hybridized with ^{32}P-labeled insert from pPol431. If multiples of 5.3 kb occurred then the PstI site was almost adjacent to the BamHI site at the 1100 bp EcoRI fragment end of the pPol431 insert. If bands 4.3 and 4.8 kb or their multiples appeared, then the PstI site was located 0.4 kb from the 700 bp EcoRI end of the inserted repeat sequence. A prominent 4.8 kb band was observed in the autoradiogram of the Southern blots of partially digested DNA hybridized with labeled repeat DNA (data not shown). This band can only be generated if the PstI site is 0.4 kb from the BamHI site which is adjacent to the 700 bp EcoRI fragment in pPol431.

The position of the PstI site between BamHI elements was mapped and found to be 0.4 kb from a BamHI site in the 5.3 kb repeat. The PstI and EcoRI double digest autoradiogram bands in Pisaster DNA are 2.1, 1.3, 1.1 and 0.8 kb long. The 1.3 and 0.8 kb represent the EcoRI internal fragments but the 2.1 and the 1.1 kb contain the 1.1 and 0.7 kb ends of the pPol431 insert. The observed fragments showed that the PstI site between the two 3.9 kb repeats was 0.4 kb from the 0.7 kb EcoRI fragment of the first element and 1.0 kb from the 1.1 kb EcoRI fragment of the second element. The genomic organization of the 5.3 kb repeat

family is shown in Figure 9.

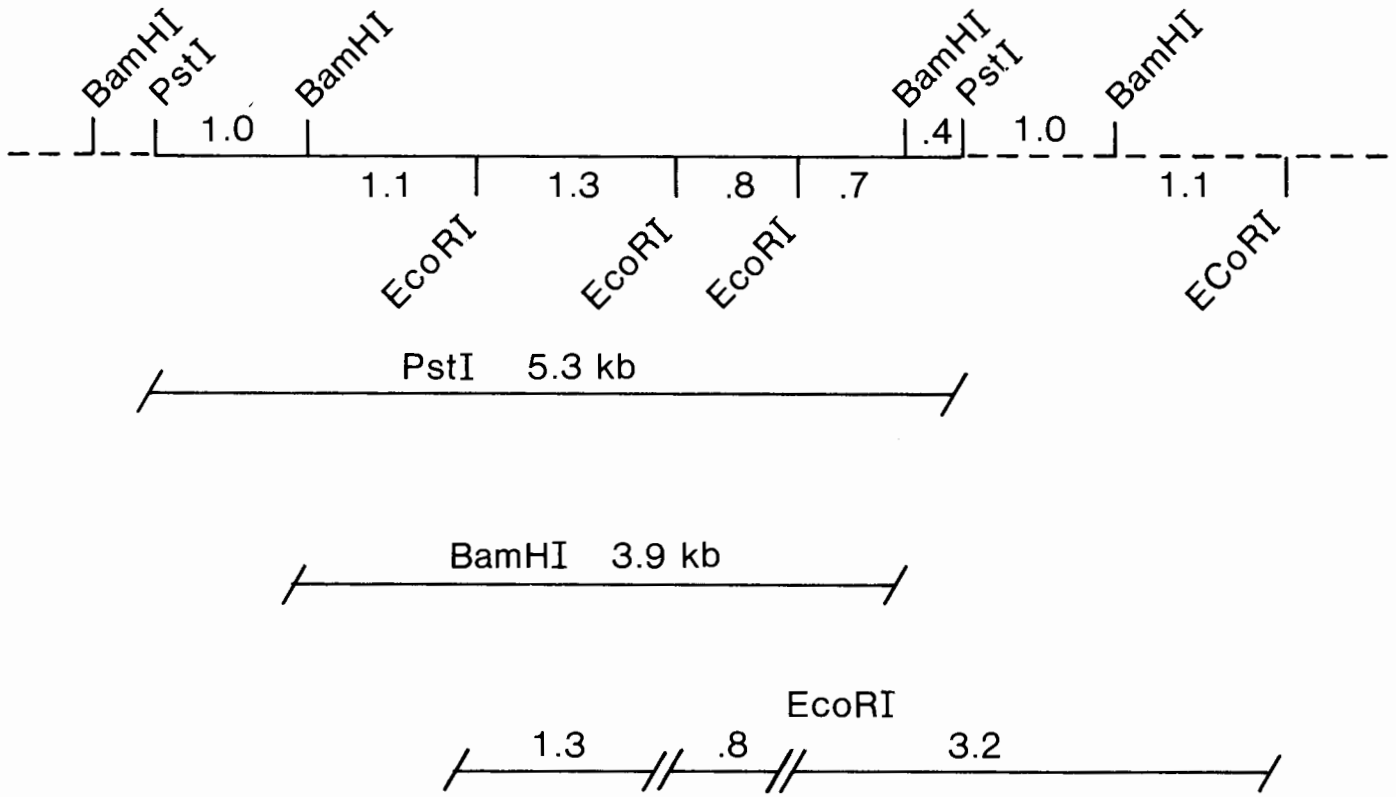
Long exposure of Southern blots indicate that members of the repeat family are clustered with perhaps smaller groups or even single elements present elsewhere in the genome. To further verify that the family of repeats is indeed clustered the P. ochracreus genomic Charon 4 phage library was screened with the labeled insert from pPol431 and the number of positive plaques were counted. If the average insert size in the Charon 4 phage is 15 kb each repeat DNA containing phage would have a capacity for approximately 3 of the 5.3 kb PstI elements. The kinetic analyses indicate that there are about 500 copies of the repeat in the genome (see below). Consequently there should be about 500/3 phage which contain repeat DNA per genome. In this survey 1.6×10^4 phage were screened representing 40% of the genome. Of these, 98 hybridized with repeat probe where 66 were expected to contain repeat sequence if the family was totally clustered.

Methylation of P. ochraceus genomic repeat sequence.

To determine if the genome repeat family was methylated in the three different starfish, genomic DNA was digested with a pair of restriction endonuclease isoschizomers, one of which is inhibited by site specific methylation. MspI and HpaII are restriction enzymes which recognize the same nucleotide sequence

Figure 9. Organization of the repeat family in the Pisaster genome.

The repeat fragment lengths expected from digestion of genomic DNA with PstI, BamHI or EcoRI are indicated below the figure illustrating the organization of individual elements in a genomic repeat DNA cluster. The intervening PstI site was precisely mapped by partial digestion of genomic DNA with BamHI and PstI followed by Southern transfer and filter hybridization with ^{32}P-labeled insert from pPol431. PstI, EcoRI and BamHI restriction sites are indicated in the clustered repeat family as it appears in the genome. The distance between restriction sites are indicated in kilobase pairs.


BamHI
PstI
BamHI
BamHI
PstI
BamHI
1.0
.4
1.0
1.1
1.3
.8
.7
1.1
EcoRI
EcoRI
EcoRI
ECoRI
PstI 5.3 kb
BamHI 3.9 kb
EcoRI
1.3
.8
3.2

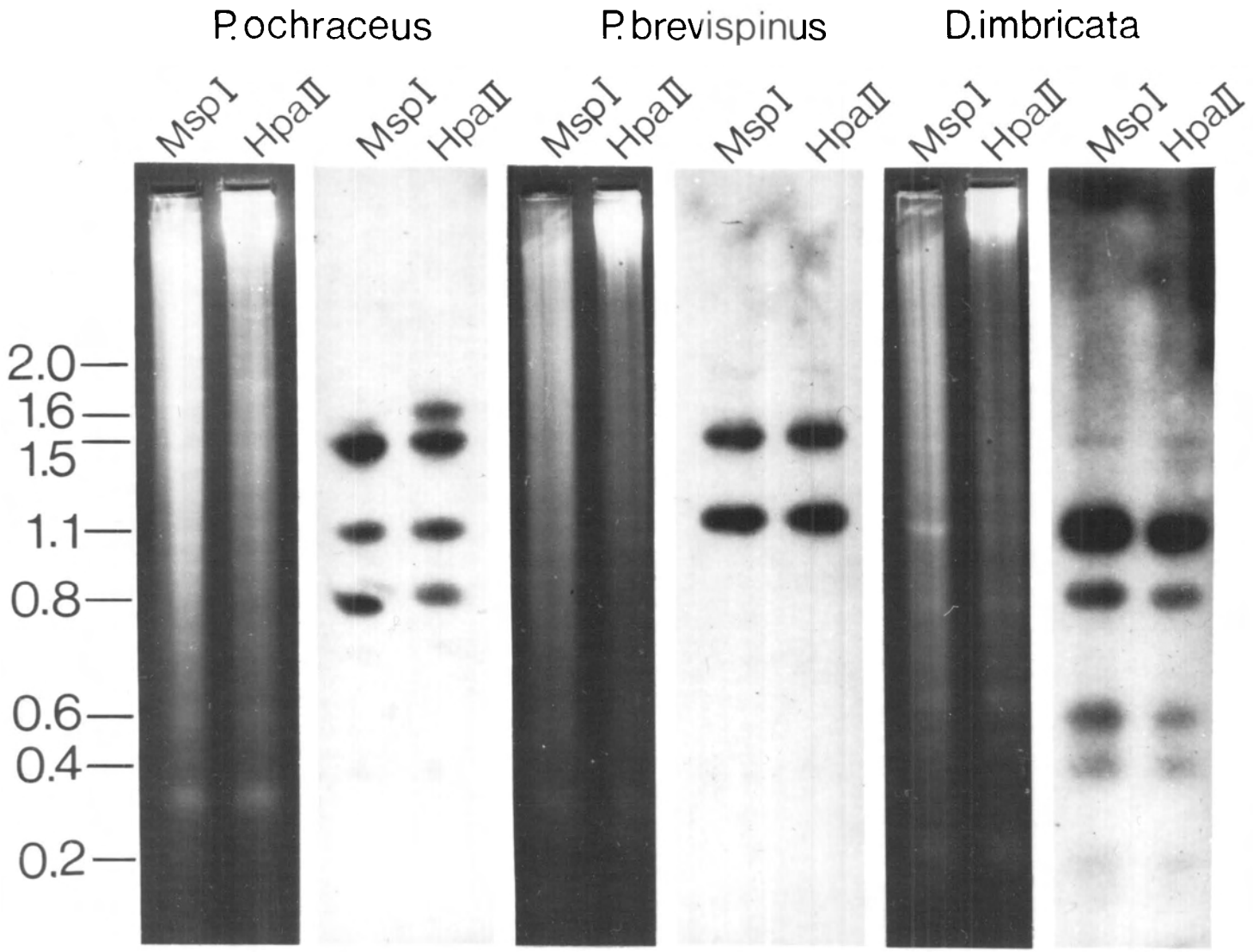
5'-C/CGG-3'. However, if the second cytosine residue is methylated as is commonly found in DNA, then only MspI will cleave at this site. DNA digested with both these enzymes will show different electrophoretic patterns if some of the cytosine residues are methylated. If the repeat family is methylated then hybridized Southern blots of these two digests will also be different. The ethidium bromide stained gel in Figure 10 shows that there is a large amount of high molecular weight DNA in the HpaII genomic digests in all three species. The MspI digestions are much more extensive. This demonstrates that there is a high degree of methylation in sperm DNA from all three starfish species. The autoradiograms show that there is only one band at 1.6 kb in the P. ochraceus HpaII digest which is not present in the MspI digest (Fig. 10). Since a new band does not appear in the MspI autoradiogram, then the HpaII 1.6 kb fragment must have methylated sequences at specific positions, such that MspI digestion creates fragments equal in length to the lower molecular weight bands. In fact, the 1.6 kb methylated fragment may well contain two 0.8 kb fragments.

Hybridization kinetic estimation of genomic frequency of the repeat family represented by pPol431.

Reassociation of the labeled repeat insert with total genomic DNA at standard conditions (0.18 M Na^+, 60^oC, pH 6.8), will

Figure 10. Methylation of repeat DNA elements in the genomes of P. ochraceus, P. brevispinus and Dermasterias imbricata.

P. ochraceus, P. brevispinus and Dermasterias sperm DNAs were digested with restriction endonucleases MspI or HpaII and electrophoresed in a 2% agarose gel at 15V for 16 hours. HpaII will not cleave if the cytosine residue in the CG region of the restriction sites, 5'-CCGG-3', are methylated. The isoschizomer MspI is methyl insensitive. Genomic DNA is methylated since the ethidium bromide stained gel clearly shows that in all three species HpaII does not digest the genome as completely as MspI. Hybridizatiation of ^{32}P-labeled insert from pPol431 with Southern transfers and autoradiography of the blots showed that only one band at 1.6 kb in P. ochraceus HpaII digest is different between the HpaII and MspI genomic blots.


P. ochraceus
P. brevispinus
D. imbricata
MspI
HpaII
MspI
HpaII
MspI
HpaII
MspI
HpaII
MspI
HpaII
MspI
HpaII
2.0
1.6
1.5
1.1
0.8
0.6
0.4
0.2

54B

proceed at a second order rate dependent on the number of copies of sequences homologous to the repeat element in the driver DNA. That is, the ratio of the tracer labeled repeat DNA reassociation rate to the driver DNA single copy rate of reassociation indicates the repeat sequence frequency in the driver DNA genome. Since the rate of reassociation for single copy DNA in each starfish genome is known (Smith et al., 1982), it is possible to determine the frequency of the repeat family.

Genomic DNA from P. ochraceus, P. brevispinus, Evasterias, Pycnopodia or Dermasterias sperm was sheared to average fragment lengths listed in Table I. The genomic driver DNA was present in the hybridization mixture at a 1 x 10^5 mass excess to the ^{32}P-labeled pPol431 repeat insert. Reassociations were done at 60°C in 0.12 M Na phosphate buffer. The mixtures were heat denatured and reassociated at 60°C to various Cot values. The amount of hybridized tracer was assayed by hydroxyapatite chromatography. The data were analyzed for second order reassociation reactions using a least squares fit computer program for one component. These hybridization reactions are illustrated in Figure 11. The second order reassociation rates, genomic frequency of repetitive DNA homologous to the pPol431 insert, and extent of tracer hybridization are given in Tables I and II.

Figure 11. Reassociation kinetics of ^{32}P-labeled tracer from pPol431 reacted with a mass excess of sea star genomic DNAs.

A 10^5 mass excess of short driver DNAs from P. ochraceus, P. brevispinus, Evasterias troschelii, Pycnopodia helianthoides, or Dermasterias imbricata were reassociated with ^{32}P-labeled insert from pPol431. Table I lists the driver and tracer DNA fragment lengths as well as the second order reassociation rates. The DNAs in the reaction mixtures were heat denatured and then reassociated at 60^{o}C in 0.12 M Na phosphate buffer, pH 6.8, to various Cots (moles/l x sec). At each Cot value the amount of duplex formed was assayed by hydroxyapatite chromatography. The solid lines represent computer analyzed least squares best-fit lines for the reassociation of a single second order component in each reaction. Key: Open triangles, P. ochraceus driven reaction; Closed Circles, P. brevispinus; Open Circles, Pycnopodia; Open Squares, Evasterias; Closed squares, Dermasterias.

Table I footnotes.

a; Calculated as the ratio of the length corrected reassociation rate of repeat insert to rate of reassociation of single copy DNA in the excess driver DNA.

b; RMS is the root mean square deviation of the data from the least squares solution.

c; Single copy rate corrected to driver DNA length.

Table I. The reassociation kinetics of sea star genomic DNA reassociated with ^{32}P-labeled repetitive DNA insert from pPol431.

Driver DNA	Driver Length bp	Tracer Length bp	Reassociation Rates ($M^{-1}s^{-1}$) Observed	Length Corrected for s.c.[c]	Corrected to Tracer Length at Driver Length	Frequency[a]	RMS[b] %
P. ochraceus	390	800	1.170	.001596	.8169	512	3.28
P. brevispinus	275	800	0.667	.001340	.3911	292	1.81
Evasterias	790	500	0.234	.002340	.2940	120	1.05
Pycnopodia	270	500	0.285	.001130	.2094	185	1.60
Dermasterias	275	800	0.210	.001610	.1231	77	0.92

Table II footnotes.

Reassociation kinetic data plotted as Cot vs percent of ^{32}P-pPol431 in hybrid (Fig. 11) shows that the tracer does not completely react with the species driver genomic DNA. The total amount of driver DNA reacted is shown in the Table. Single copy DNA reaction extent corrected to 100% is also listed for comparison (Smith et al., 1982).

Table II. Extent of hybridization between genomic sea star DNA and ^{32}P-labeled repeat DNA from pPol431.

Driver DNA	pPol431 Observed 60°C	pPol431 Corrected to 100% Homologous Reaction	Single Copy P. ochraceus DNA 60°C
P. ochraceus	79.9%	100.0%	100.0%
P. brevispinus	80.0	100.0	98.0
Evasterias	62.6	76.4	36.9
Pycnopodia	64.0	77.4	30.7
Dermasterias	36.8	55.2	6.0

There are fewer copies per genome of sequences homologous to the pPol431 insert the more distant the phylogenetic relationship of the sea stars from P. ochraceus. The last common ancestor between Pisaster ochraceus and Pisaster brevispinus, Evasterias, Pycnopodia and Dermasterias was approximately 5-10 myr, 20-40 and 500 myr respectively (Smith et al., 1982). The genomic frequency of the repeat elements in the five star fish species are approximately 500 in P. ochraceus, 300, 125, 190 and 80 in P. brevispinus, Evasterias, Pycnopodia and Dermasterias respectively. There are more copies in Pycnopodia than in Evasterias although Pycnopodia is a little further removed from P. ochraceus than Evasterias. It can be seen from Figure 11 that the Evasterias and Pycnopodia hybridization curves are almost coincident. The accuracy of these frequency estimates is on the order of 2-4% in each case (see standard error values for rate estimates in Table I).

The total amount of ^{32}P-labeled pPol431 insert DNA which hybridizes with each species is also listed in Table II. Once again the further the phylogenetic distance, the less the total reaction; for example, 55.2% of the tracer hybridizes with Dermasterias genomic DNA. There is essentially complete reaction of the tracer with either of the Pisaster species. Near the end of the reaction there may be so few copies remaining in the driver DNA (even with a mass excess) that

tracer self-reaction begins to become the dominant reassociation reaction. It is also possible that the 3.9 kb sequence in the tracer is not entirely represented in the Dermasterias genome or that some of the genomic members are so diverged that they form only short duplex regions that do not bind to hydroxyapatite. This could also lead to strand displacements of a poorly base-matched tracer by a less divergent homologous fragment.

The repeat family contains a palindromic sequence which reassociates at low Cot.

A small fraction of the 3.9 kb repeat DNA sequence which was cloned appears to consist of inverted repeats or fold-back sequences of sufficient length to bind to hydroxyapatite. This property was observed in the kinetic experiments where approximately 10-20% of the tracer bound to the hydroxyapatite columns at very low Cot values (Fig. 11). To locate the region of the repeat sequence which contained the fold-back, the 1.3, 1.1, 0.8 and 0.7 kb EcoRI fragments of the insert from pPol431 were isolated, ^{32}P-labeled by nick translation and assayed for fold-back sequence. Samples of the labeled fragments were diluted in 0.12 M phosphate buffer, heat-denatured, and immediately quenched in liquid nitrogen. In this procedure only those sequences which are intramolecularly homologous, that is are fold-back sequences, can reassociate after heat

denaturation. The amount of fold-back was measured by hydroxyapatite chromatography at 60°C in 0.12 M phosphate buffer. The ^{32}P-labeled DNA that immediately reassociated remained bound to the column after washing at 60°C. All four fragments bound between 8 and 20% of the counts (Table III). The hydroxyapatite bound DNA was thermally eluted from the columns, cooled and passed over another hydroxyapatite column. Greater than 60% of each bound again to the hydroxyapatite. As a control experiment, pBR322 DNA was ^{32}P-labeled by nick translation and also assayed in the same manner since it does not contain fold-back sequences. Less than 3% of the counts bound to hydroxyapatite columns. Therefore, there are fold-back sequences throughout the 3.9 kb repeat DNA insert from pPol431 which can bind to hydroxapatite at a low Cot.

Sequence divergence of pPol431 repeat family within starfish DNA.

HaeIII restriction enzyme analysis indicated intraspecific sequence divergence between the clones pPo328, pPol289, pPol431 and pPol675 (Fig. 5). A reliable measure of the average amount of sequence divergence of the repeat DNA family in the genome is its thermal denaturation temperature, the T_m, of reassociated duplexes between driver and tracer. The ^{32}P-labeled insert from pPol431 was reacted to five different starfish driver DNAs

Table III footnotes.

Four EcoRI repeat DNA fragments (Fig. 6) from pPol431 were ^{32}P-labeled and assayed for the presence of fold-back sequence. The fragments were boiled, quenched and passed over a hydroxyapatite column at 60^{o}C. The amount bound was heat denatured to remove the fold-back from the column and then repassed over hydroxapatite.

Table III. Presence of the fold-back sequence in each EcoRI fragment from pPol431.

EcoRI Fragment From pPol431 Insert bp	% Bound First Passage	% Bound Second Passage
700	16.8	58.0
800	20.6	68.0
1100	11.0	68.0
1300	7.5	84.0

and then thermally denatured in small temperature increments from hydroxyapatite columns. The T_m is the temperature at which half the reassociated duplexes are eluted from hydroxyapatite columns (Fig. 12). In one experiment the reassociated duplexes formed at 60°C and native ^{32}P-labeled insert, were first S1-nucleased and then thermally eluted from a HAP column prepared at 50°C. A more precise measure of the T_m is obtained after S1-nuclease digestion of reassociated duplexes. The number of counts found in the duplexes formed from randomly sheared fragments at a given temperature increment can be overestimated due to long single-stranded tails of the tracer in the duplex. These tails were enzymatically removed by S1-nuclease. The observed and length corrected T_m values are listed in Table IV. In another experiment ^{32}P-labeled pPol431 insert and sea star genomic DNAs were reassociated at both 50 and 60°C in 0.12 M phosphate buffer, pH 6.8, and the T_ms of duplexes were assayed on hydroxapatite (Table V). The lower, 50°C temperature for reassociation will allow less stable duplexes to form than those at 60°C resulting in a lower T_m value. The difference in T_m measurements will indicate presence of repeat DNA family members whose sequences are more diverse. A third thermal denaturation study of the five starfish species used tracer that was first denatured and passed over HAP to remove the fold-back portion of the probe. The tracer was treated in this manner because the presence of

Figure 12. Hydroxyapatite thermal denaturation of ^{32}P-labeled insert DNA from pPol431.

pPol431 was digested with BamHI and electrophoresed in an agarose gel. The 3.9 kb BamHI repeat insert was isolated from the gel and ^{32}P-labeled by nick translation. The labeled double-stranded DNA was denatured and eluted in 3^{o}C increments from a hydroxyapatite column prepared in 0.12 M Na phosphate buffer. The data was plotted as the cummulative fraction of tracer eluted versus the temperature (^{o}C). T_m was determined as the temperature at which 50% of the ^{32}P-labeled DNA had eluted. The observed T_m is 89.0^{o}C and the length corrected T_m is 91.5^{o}C with the average length of tracer being 480 bp.

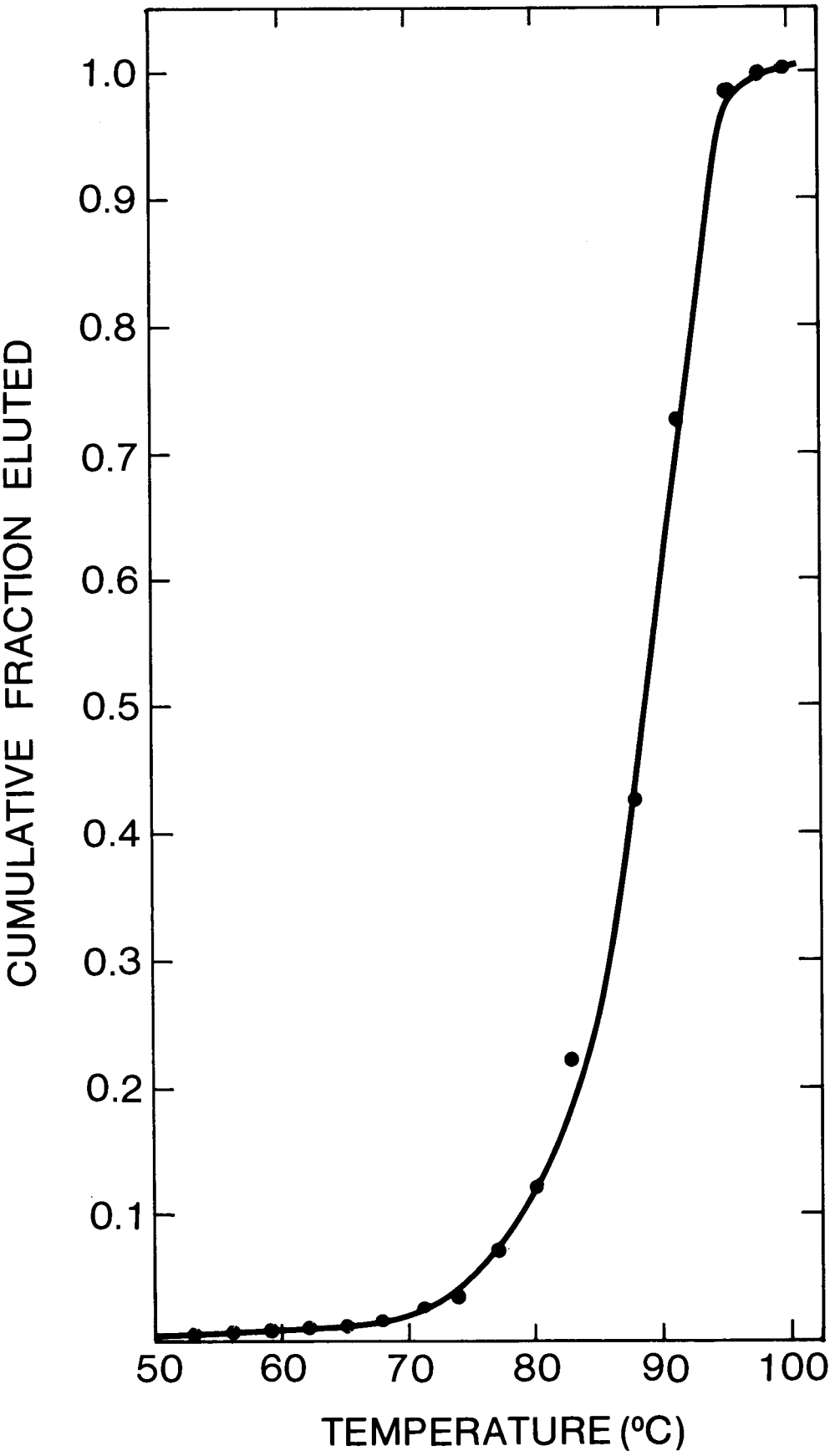

CUMULATIVE FRACTION ELUTED
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
50
60
70
80
90
100
TEMPERATURE (°C)

Table IV footnotes.

Sheared genomic driver DNA from five starfish species was reassociated in a mass excess to ^{32}P-labeled insert from pPol431 to various Cot values: P. ochraceus and P. brevispinus, Cot 5; Evasterias, Pycnopodia and Dermasterias, Cot 10. The reassociated duplexes formed at 60°C and 0.18 M Na^+, were S1-nuclease digestion at 6 units to 1 ug of single stranded DNA before thermally eluting from hydroxyapatite columns.

a; The S1-nuclease digested duplexes were electrophoresed in 1.5% agarose gels and autoradiographed. The size of the tracer in the duplex is listed. Native tracer or driver duplex was sized on a 1.5% alkaline agarose gel.

b; T_m is the temperature at which half the reassociated duplexes become single stranded during thermal elution from hydroxyapatite columns.

c; The T_m varies depending on the lengths of DNA molecules reassociating in the reaction mixture. All T_m values have been adjusted in order to make comparisons. The heteroduplex length was calculated as 55% of the length of the shortest member of the fragment pair. The depression in T_m (dT_m) due to fragment length was calculated from the relationship $dT_m = 650/0.55L$, where L is the shortest fragment length (Britten et al., 1974). The dT_m was added onto the observed T_m to obtain the length corrected values.

Table IV. Thermal denaturation temperatures of sea star DNA and ^{32}P-labeled pPol431 insert formed at 60°C after S1-nuclease digestion.

Driver DNA	Driver Length bp	Probe Length[a] in Duplex bp	Observed[b] T_m (°C)	Length Corrected[c] T_m (°C)
P. ochraceus	280	450	86.0	90.2
P. brevispinus	330	410	84.0	87.9
Evasterias	400	350	80.0	83.3
Pycnopodia	400	350	80.0	83.3
Dermasterias	330	350	74.5	78.4
Native		480	89.0	91.5

Table V footnotes.

The presence of a fold-back sequence in the probe, ^{32}P-labeled insert from pPol431, may have affected the true T_m values between reassociated duplexes listed in Table IV. Therefore, an aliquot of the labeled tracer was boiled, quenched and passed over hydroxyapatite at 60°C (See Mat. and Met.) in order to "strip" the sequence of the fold-back. The front peak was used as the tracer reassociated with sea star genomic DNA. The DNAs were hybridized at 60° or 50°C in 0.18 M Na^+.

a; Driver DNA reassociated with untreated ^{32}P-labeled pPol431 insert.

b; Driver DNA was reacted with repeat DNA tracer that had the fold-back sequence removed prior to reassociation.

c; The observed T_m values were adjusted as described in Table IV for "Length Corrected T_m". The lengths used in the corrections are listed in Table V. The length of the tracer in all experiments was 450 bp.

Table V. Thermal denaturation temperatures of ^{32}P-labeled repeat DNA tracer reacted with genomic driver DNA from five sea star species at 60° and 50°C.

Driver DNA	pPo1431 Tracer[a]				Stripped Tracer[b]				Cot[c]		Cot	
	Shortest Length in Duplex	60° OBS T_m	60° ADJ[d] T_m	50° OBS T_m	50° ADJ T_m	60° OBS T_m	60° ADJ T_m	50° OBS T_m	50° ADJ T_m	pPo1431 Tracer 60°C	pPo1431 Tracer 50°C	Stripped Tracer 60°C	Stripped Tracer 50°C
P. ochraceus	380	80.0	83.1	80.0	83.1	80.1	83.2	79.0	82.1	343	341	131	131
P. brevispinus	300	77.0	80.9	75.5	79.4	76.5	80.4	75.0	78.9	122	121	77	78
Evasterias	450	75.5	78.1	74.5	77.1	75.5	78.1	74.5	77.1	136	132	113	112
Pycnopodia	310	72.0	75.8	70.1	73.9	72.0	75.8	70.0	73.8	153	149	125	124
Dermasterias	250	69.0	73.7	67.5	72.2	68.5	73.2	67.5	72.2	209	203	158	156

fold-back sequences in the repeat DNA sequence will interfere with the true T_m values of reassociated duplexes (Table V). Figures 13 and 14 are plots of the fraction of ^{32}P-labeled pPol431 insert eluted *vs* temperature increment for the S1-nucleased duplexes reassociated at 60°C and for stripped tracer reassociated at 50o and 60°C.

Thermal chromatograms of S1 nuclease digested duplex all show a significant high temperature melting component (Figure 13). This high T_m component could be due to regions of conserved high sequence homology, regions of high GC content, to regions of snap-back or fold-back within the labeled probe itself, or to tracer self-reaction. The high melting component due to fold-back thermal denaturation seems an unlikely alternative because the high T_m would require an appreciably long fold-back sequence or an extremely high GC content within the sequence. However, the proposition was analyzed in the following manner. The fold-back sequence was removed from the tracer prior to reassociation with the driver DNAs at both 50°C and 60°C criteria. Although the high melting component is not apparent in the thermal chromatograms (Fig. 14) of *Evasterias*, *Pycnopodia* and *Dermasterias* and is reduced in the *Pisaster* species, there can still be tracer self-reaction which would raise all T_m values artificially high.

Figure 13. Thermal denaturation chromatograms of hybrid DNA duplexes of ^{32}P-labeled pPol431 insert DNA and genomic DNA from five starfish species.

Total sheared genomic starfish DNA from the five sea star species was reassociated with ^{32}P-labeled pPol431 insert DNA to the following Cot values: *P. ochraceus*, 5; *P. brevispinus*, 5; *Evasterias*, 10; *Pycnopodia*, 9.6; and *Dermasterias*, 10. At these Cot values the reassociated DNA was treated with S1 nuclease to destroy single strand DNA fragments. The endonuclease resistant duplex was loaded on hydroxapatite columns prepared in 0.12 M Na phosphate buffer, pH 6.8 and thermally eluted from 50^{o}C in 3^{o}C temperature intervals to 98^{o}C. The open bars represent the fraction of ^{32}P-labeled DNA eluted at each temperature interval. The chromatograms have been superimposed over a thermal chromatogram of native pPol431 insert DNA (black background) which is illustrated in Panel A. A) Native ^{32}P-labeled duplex from pPol431 insert; B) *P. ochraceus* driver DNA; C) *P. brevispinus*; D) *Evasterias*; E) *Pycnopodia*; and F) *Dermasterias*.

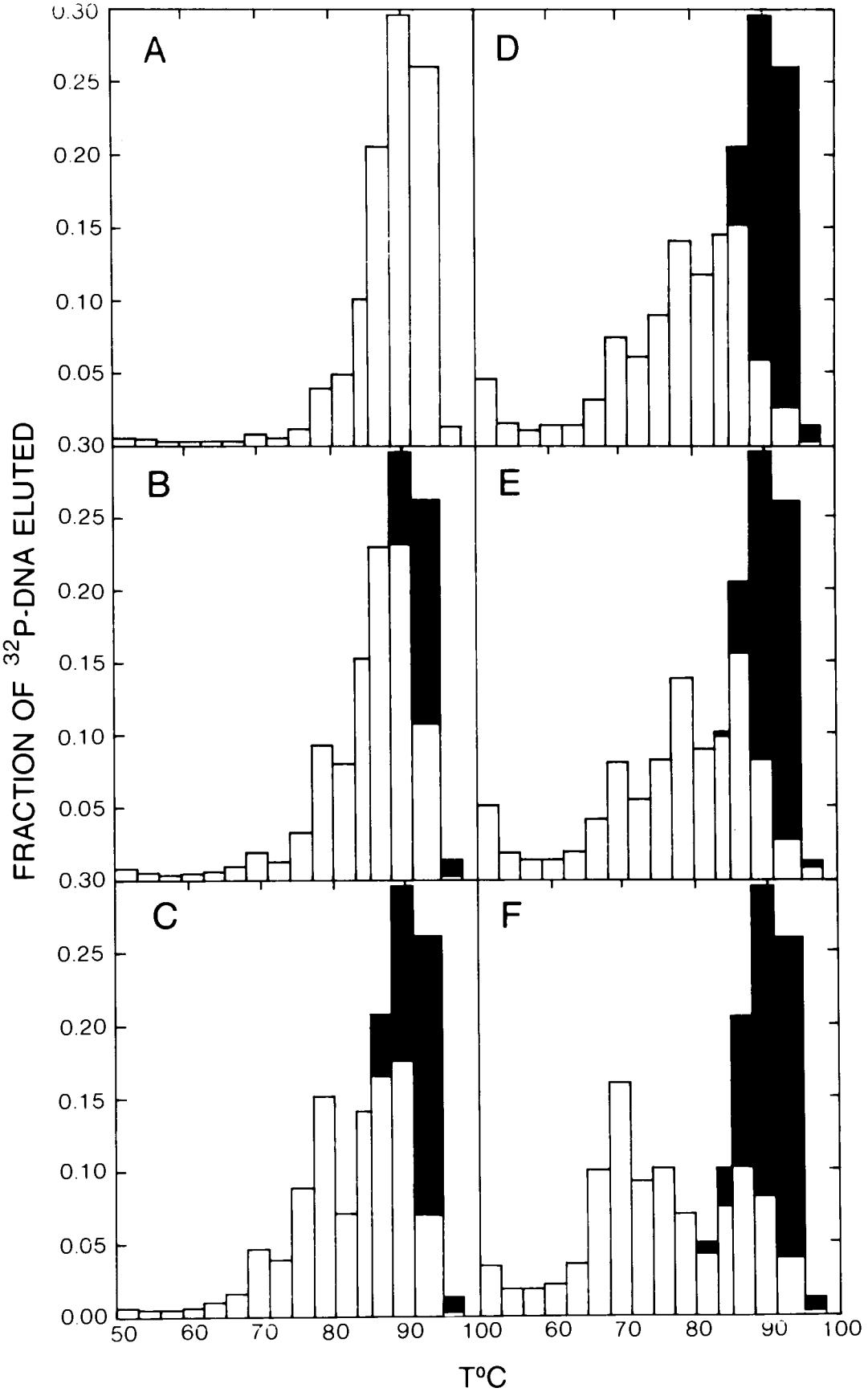
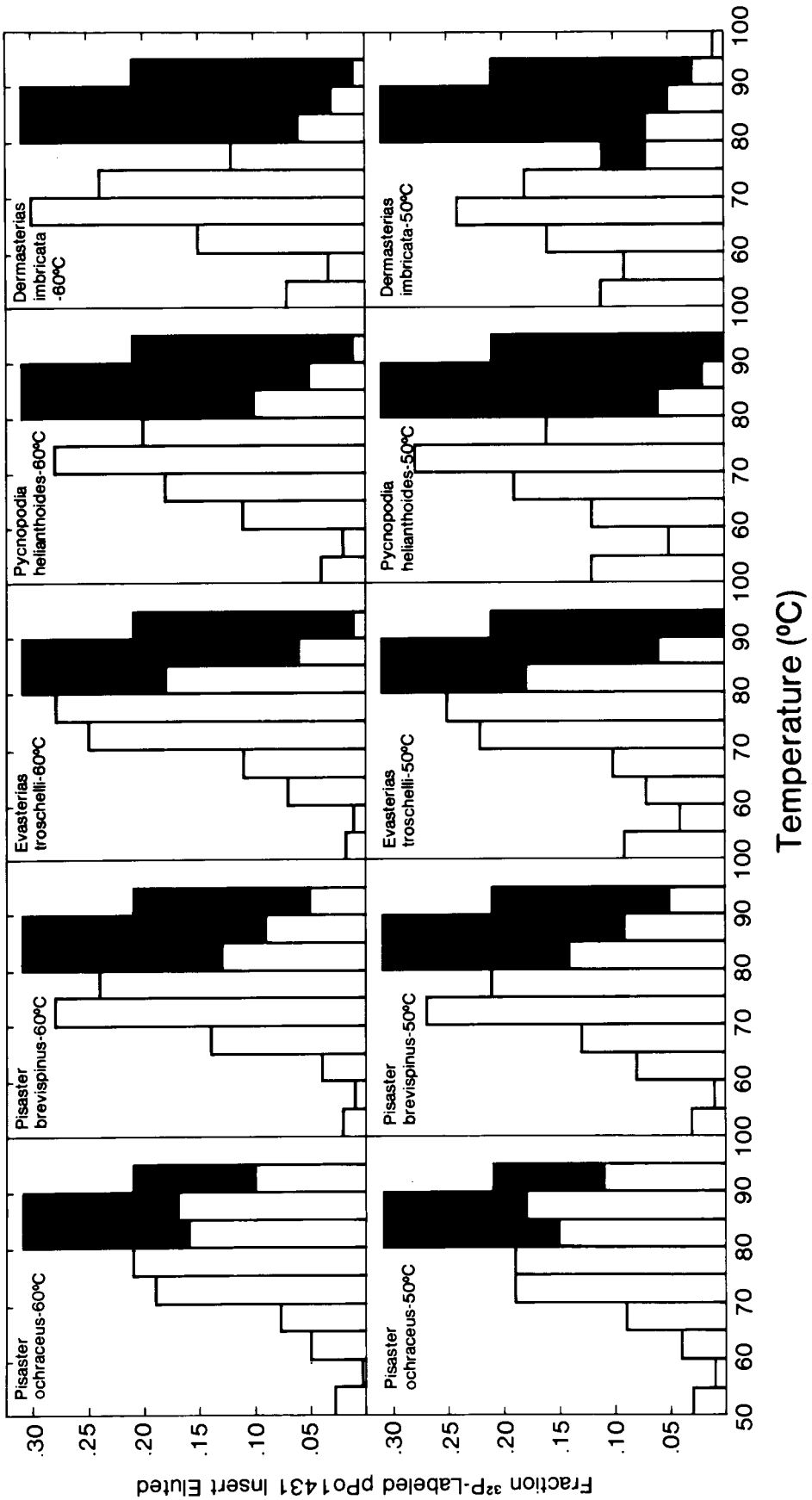

A
B
C
D
E
F
FRACTION OF ^{32}P-DNA ELUTED
0.30
0.25
0.20
0.15
0.10
0.05
0.00
50
60
70
80
90
100
TºC

Figure 14. Thermal denaturation chromatograms of hybrid DNA duplexes between excess sea star genomic DNA and "stripped" ^{32}P-labeled insert from pPol431.

^{32}P-pPol431 repeat insert DNA was heat denatured and passed over a hydroxyapatite column at 60^{o}C in 0.12 M Na phosphate buffer to remove the fold-back portion of the repeat sequence prior to reassociation with the driver DNA. The hydroxyapatite thermal denaturation chromatography was initiated at 50^{o}C as outlined in Materials and Methods. The bars represent the fraction of reassociated ^{32}P-labeled DNA eluted at each 5^{o}C increment. The thermal denaturation profiles are superimposed on the thermal chromatogram of native duplex pPol431 insert (black background). The reassociated duplexes were reacted with genomic driver DNA to various Cot values listed in Table V.


Pisaster ochraceus-60°C
Pisaster brevispinus-60°C
Evasterias troschelli-60°C
Pycnopodia helianthoides-60°C
Dermasterias imbricata -60°C
Pisaster ochraceus-50°C
Pisaster brevispinus-50°C
Evasterias troschelli-50°C
Pycnopodia helianthoides-50°C
Dermasterias imbricata-50°C
Fraction 32P-Labeled pPo1431 Insert Eluted
.30
.25
.20
.15
.10
.05
50 60 70 80 90 100 100 90 80 70 60 100 90 80 70 60 100 90 80 70 60 100 90 80 70 60 100
Temperature (°C)

The reassociation temperature was lowered to a more permissive temperature of 50°C in order to allow hybridization of more divergent members of the repeat family. The thermal chromatograms of the "stripped" probe at both 50 and 60°C are shown in Figure 14. Comparisons of the 50 and 60°C chromatograms demonstrate that the 50 and 60°C criteria did not produce a remarkable difference in the thermal denaturation of hybrids with genomic DNA from any of the species. This suggests that few if any highly divergent copies of the family are present in the genomes.

The amount of sequence divergence intra- and interspecifically was determined by subtracting the reassociated duplex T_m values from the T_m value of pPol431 insert reassociated with P. ochraceus genomic DNA (Table VI). The dT_m from native duplex measures the % base pair mismatch in reassociated DNA sequences. For every % base pair mismatch there is a corresponding decrease of 1°C in the melting temperature of reassociated duplex (Davidson, 1976). The single copy T_m values (Smith et al., 1982) for the five starfish species and their observed dT_m values obtained from the reactions with the repeat DNA are listed in Table VII. The thermal denaturation depression observed in the homologous reaction with P. ochracreus has been subtracted from the other species to demonstrate the divergence between the species and not the total

Table VI footnotes.

The length corrected T_m values determined for the S1-nucleased duplexes (Table IV) and for the hybrid duplexes formed with untreated and stripped tracers at 60° and 50°C (Table V) were subtracted from the length corrected T_m value (91.5°C) of the precise duplex, ^{32}P-labeled pPol431 insert.

Table VI. Thermal denaturation depressions from native duplex T_m.

Driver DNA	S1-nuclease Digested Duplex	Tracer		Stripped Tracer	
	60°C	60°C	50°C	60°C	50°C
P. ochraceus	1.0	8.4	8.4	8.3	9.4
P. brevispinus	3.6	10.6	12.1	11.1	12.6
Evasterias	8.2	13.4	14.4	13.4	14.4
Pycnopodia	8.2	15.7	17.6	15.7	17.7
Dermasterias	13.1	17.8	19.3	18.3	19.8

Table VII footnotes.

To estimate the amount of sequence divergence between species, the thermal denaturation depressions listed in Table VI of the heterologous species were subtracted from the thermal depressions of the parent species, P. ochraceus. The values listed in this table represent the percent of interspecific sequence divergence since for each 1°C depression in T_m there is an equivalent 1% sequence diversity.

a; Single copy DNA from P. ochraceus was reacted with other sea star genomic DNAs and the thermal denaturation depressions are listed for comparison ot the repeat DNA family represented by pPol431 (Smith et al., 1982).

Table VII. Extent of sequence divergence within the repeat DNA family determined by thermal denaturation depressions.

Driver DNA	S1-nuclease Digested Duplex	Tracer from pPol431		Stripped Tracer from pPol431		Single Copy Tracer	
	60°C	60°C	50°C	60°C	50°C	60°C	50°C
P. brevispinus	2.6	2.2	3.7	2.8	3.2	5.2	4.3
Evasterias	7.2	5.0	6.0	5.1	5.0	9.5	14.6
Pycnopodia	7.2	7.3	9.2	7.4	8.3	9.7	13.1
Dermasterias	12.1	9.4	10.0	10.0	9.9	10.2	16.1

divergence from the native duplex represented by the labeled insert of pPol431.

The repeat DNA family is not transcribed in P. ochraceus oocyte nor is it ribosomal.

Dot blots containing increasing amounts of oocyte RNA, ribosomal DNA representing the 18S and 28S gene and spacers or pPol431 DNA as a control, were used to determine if the the repeat sequence was transcribed. No reaction was observed with RNA or the rDNA blots but strong reaction was observed in the control. The dot blot technique is reliable and it appears that the repeat family is not found in any of these transcribed sequences. To confirm the rDNA result, a Southern blot of an EcoRI digested lambda clone of the ribosomal sequences and hybridized with ^{32}P-labeled repeat insert and no reaction was observed in the autoradiogram.

DISCUSSION

Characteristics of the repeat family in Pisaster ochraceus DNA.

Restriction enzymes have been a convenient tool in isolating middle repetitive DNA from the starfish genome. The size and intensity of electrophoretic bands observed in agarose gels can

be estimated and from this information we can assess whether or not the fragments within the band are from a repeat class of interest. The principle advantage with this technique of selecting a repeat family is that it is also easy to clone since the same restriction enzyme used to digest the genomic DNA may also cleave pBR322 in the drug resistant genes facilitating selection for recombinants. Sea urchin (Scheller et al; 1977) and chicken (Eden, 1980) repeat DNA was cloned into plasmid vectors by first reassociating denatured genomic DNA to low Cot values, isolating the duplexes, adding artificial linkers containing the sequence of an appropriate restriction enzyme for cloning and then inserting the repeat fragments into the vector. This technique provides an opportunity to select for many different kinds of cloned repeat families although if a particular family is required then it would be tedious to screen for the correct positive clones. In the starfish experiments we were interested in the evolution of a family of moderately repetitive DNA which was also from the long repeat class. Digestion of P. ochraceus DNA with BamHI, a convenient enzyme for cloning into the pBR322 tetracycline site, produced an electrophoretic band 3.9 kb in length of moderate intensity when stained with ethidium bromide. This band of fragments therefore met the requirements for studying long moderately repetitive DNA although we did not know at that point if there was more than one repeat family within the band.

Pisaster ochraceus genomic DNA contains 35% repeat DNA which is found in both the long and short interspersion patterns of the genome (Smith and Boal, 1978). Long repeat sequences greater than 1000 kb in length represent 7.9% of total genomic DNA. The repeat family represented by pPol431 cloned repeat is reiterated 500 times in the genome and at a length of 5.3 kb this is equivalent to 0.5% of the genome, or 1.7% of total repeat DNA or 3.4% of long repeat DNA. This is equivalent to 2.8×10^6 bp. Therefore, the family is a significant amount of genomic DNA.

The repeat family found within P. ochraceus genome is a well conserved, homogeneous and cytosine-guanosine rich (54%) class of repetitive DNA. P. ochraceus genomic DNA has 41% GC content. The melting temperature of native ^{32}P-labeled insert from pPol431 is 91.5^{o}C indicating a CG rich sequence. When the labeled inserts are reacted with sheared genomic DNA the melting temperature of reassociated duplexes measured by hydroxapatite chromatography is 86.0^{o}C. This 5.5^{o}C depression in thermal denaturation temperature from native pPol431 insert relative to genomic driver hybrid indicates the range of sequence divergence in the repetitive family. A value of 5% intraspecific sequence variation attributed to polymorphism was observed in single copy DNA using the same reassociation criteria and assay system (Smith et al., 1982). This suggests that the members of the repeat family are at least as well conserved as single copy DNA

within the species. The reassociation temperature was lowered to 50°C in order to measure the amount of duplex formation which may have occurred at a less stringent criterion. This is a test for the presence of sequences which are less homologous than those formed at 60°C. The T_m values and the shift in thermal histograms are not sufficiently different from the 60°C data to demonstate that other family members with more sequence divergence are present in the genome.

The thermal denaturation properties and the existence of a fold-back sequence in the repeat DNA sequence suggests that it is possible that the regions within the basic elements of the repeat family are heterogeneous. A shoulder is observed in the thermal histogram of P. ochraceus which suggests that perhaps there is more than one melting component which can obscure the assessment of the T_m. Low melting components or shoulders can be due to tracer self-reaction, to segments of the repeat element having a lower T_m, or they may also be caused by duplex length heterogeneity since the melting temperature is lower if the length of duplex is shorter. Furthermore, others have shown (Moyzis et al., 1981; Preisler and Thompson, 1981b) that this kind of heterogeneous melting can result from the reassociation of rearranged subelements in a large fragment of repeat DNA. The presence of fold-back sequences in the pPol431 family correlates with the possibility that the repeat family

may contain subelements within the 3.9 kb element. Lowering the reassociation temperature from 60° to 50°C and thermally eluting the duplexes from hydroxyapatite columns should have demonstrated sequence heterogenity between repeat family members since the more permissive the temperature the greater the tolerance for sequence mismatch. A difference in thermal denaturation profiles at 50° and 60°C were not observed in any of the reactions. Therefore, the amount of duplex formation did not increase by lowering the temperature to 50°C allowing more divergent members to reassociate. There appears then to be a homogeneous family of repeats within each of the five starfish genomes that displays equivalent sequence divergence between rehybridized driver DNA and ^{32}P-labeled pPol431 insert at either 60° or 50°C. The optimal reassociation temperature for duplexes is 25°C below the native T_m value which in this case would be 67°C (the T_m of native pPol431 insert was 91.5°C). Perhaps if 60° and 70°C reassociation temperatures had been used, it may be possible to detect differences in duplex stability and sequence heterogeneity.

Restriction enzymes MspI and HpaII are used commonly to assay for DNA methylation at cytosine residues in CG pairs. The sequence 5'-CCGG-3' is recognized by both these enzymes but if the second cytosine residue is methylated then only MspI can cleave DNA. However if both cytosine residues are methylated

then neither enzyme can cleave DNA. Therefore, the assay procedures are limited in that there is no available test to determine if both cytosine residues are methylated. However, in eukaryotes, ordinarily only a CpG dinucleotide pair is methylated.

Aside from the drawbacks of this assay system for DNA methylation some interesting methylation properties do appear in the repeat family of the starfish genomes (Fig. 10). There is extensive methylation of P. ochraceus, P. brevispinus and Dermasterias sperm DNA observed by the extent of HpaII digestion relative to that of MspI digestion shown in the electrophorograms. However, the autoradiograms of these Southern transferred DNAs reacted with labeled repeat DNA show that only a few sites are methylated in the P. ochraceus repeat DNA family, represented by pPol431, generating a 1.6 kb band in the HpaII digest that is not visible in the MspI digest. This may be a newly acquired property in the P. ochraceus repeat family since methylated repeats homologous to pPol431 are not detected in the other two species.

These results obtained from the digestion of genomic DNAs by the isoschizomers, MspI and HpaII, also demonstrate that there may be more observable conservation of genomic sequence organization between smaller regions of the 3.9 kb repeat elements than that

found with EcoRI, BamHI or PstI genomic digests of the three sea star species. For example, the P. ochraceus 1.5 kb autoradiogram band is the same in P. brevispinus with some residual hybridization in Dermasterias; there is a 1.1 kb band in the autoradiograms of all three species; and, 0.8 kb fragments hybridize in both P. ochraceus and Dermasterias but not in P. brevispinus genomic digests. If the 0.8 kb sequences in Dermasterias are conserved in P. ochraceus then perhaps P. brevispinus has simply lost the restriction site throughout its repeat family whereas P. ochraceus retained the sequence during the speciation process.

There was no evidence that the repeat family is transcribed in either oocyte RNA or ribosomal RNA. However, it is still possible that the 5.3 kb repeating elements represent genes whose transcripts could not be detected in the oocyte RNA populations.

Evolution of the repeat family in sea stars.

Two important features of the evolution of this particular family have been revealed by studying the prevalence and sequence homology among five related species and by determining the repeat organization in three related species, P. ochraceus, P. brevispinus and Dermasterias imbricata. Reassociation

kinetic analyses have shown that there is a decreasing number of copies of the family with increasing phylogenetic distance (Table I). This property of repeat DNA has been revealed many times before (Moore et al., 1978) and it was not surprising to find it in this case. What is significant about these results is that the repeats appear to be maintained essentially in a clustered arrangement. Another important feature of the repeat family represented by pPol431 is that it is more conserved than single copy DNA between sea star species (Table VII). This fidelity of a repeat sequence family requires some mechanism intra- and interspecifically for insuring retention of specific sequence and rejection of divergent members. Mechanisms such as gene conversion and unequal crossing-over have been proposed to explain such homogenization processes (Jeffreys, 1981; Zimmer et al., 1980, Dover et al., 1982). In repeat families that are dispersed it has been observed (Galau et al., 1976; Preisler and Thompson, 1981a) that in fact the sequence divergence may be about the same as that for single copy. Without special mechanisms of sequence conservation which prevent sequence variation, mutations would accumulate in the family until the repeat sequences no longer form stable duplexes with other members under standard reassociation conditions. These sequences would then become single copy DNA.

Pycnopodia has more copies in the genome than Evasterias although according to rate of sequence divergence in single copy DNA (Smith et al., 1982) Pycnopodia is further removed than Evasterias from Pisaster ochraceus. There could be two reasons for this discrepancy. The two species are closely related and frequency estimates by kinetic analysis may be not be precise enough (RMS of less than or equal to 5%). Also, Pycnopodia has a larger genome than Evasterias (Hinegardner, 1974) as well as more copies of pPol431 homologs suggesting that genome size and frequency of a repeat DNA family may be interrelated.

Southern blot analysis of the repeat family in Pisaster and Dermasterias reveals that the family organization is conserved over 850 myr of divergence time. The restriction sites for BamHI, EcoRI and PstI in Pisaster ochraceus and Pisaster brevispinus are identical over 10-20 myr of divergence time. Some restriction sites may also be conserved in Dermasterias since the multiple bands in the EcoRI, PstI or BamHI Southern blots could be due to loss or a deletion of restriction enzyme sites. Furthermore, there are similar MspI sites amongst the repeat family elements in Pisaster and Dermasterias. This extent of conservation has not been observed in any other organism yet studied and it raises the question as to why the organization of the repeat family would be so well conserved in the starfish genomes. Perhaps genomic organization appears well

conserved amongst these sea stars because the repeat family elements are slowly amplified and dispersed compared to other families of repeats in other organisms discussed below. It may be that repeats are slowly dispersed in the starfish but to address this question would require comparisons with other clustered and dispersed families of repeats in these organisms and analysis of many other types of organisms.

Upon longer exposure of the autoradiograms of Southern blots of genomic DNAs more bands were observed. This indicates that although the members are retained in a large deposit in the two Pisaster and in the Dermasterias genomes, some members are dispersed. There are a number of explanations for the presence of these bands. They may be repeat members appearing elsewhere in the genome, as clusters or individually. They may be fragment ends of large clusters. The bands may also represent other sequences which have some weak homology to the pPol431 sequence. They could result from some members in the cluster losing or gaining restriction sites. Finally, intense and dispersed clusters without common restriction sites found in the main cluster, may be from reamplification of family members which had lost some sequence homology and dispersed in the genomes.

Extensive reviews have been recently published (Long and Dawid, 1980; Britten, 1981; Britten, 1982; Jelinek et al., 1982; Doolittle, 1982) about repeat DNA in eukaryotes which describe the enormous complexity of the properties of repetitive DNA. In the following discussion repeat DNA from the sea urchin, chicken, Drosophila and plants are compared to the family found in starfish since they all contain the long repeat class of DNA. The sea urchin, a closely related animal to the sea star, is representative of many animal genomes and the repeat DNA has been well characterized. The chicken genome has also been probed for individual repeat families but the genome has unusually long regions of repeat and single copy DNA arranged in a long-period interspersion pattern similar to Drosophila. Drosophila, a Protostome, contains only long repeats in its genome and these are often tandemly arranged in very large clusters. The plant kingdom has provided us with a very interesting viewpoint about repeat DNA since within many species the repeat DNA may represent up to 80% of the genome and its rate of repeat DNA amplification is very high compared to that in animals.

Britten and Davidson and coworkers have used the sea urchin as a model system to determine the possible function of repeat DNA in the eukaryotic genomes. They have suggested a model whereby the dispersed repeat members of a family may provide a means of gene

regulation since homologous sequences could connect distant regions of the genome (Britten and Davidson, 1969; 1971). They also suggest that new families of repeats could provide new relationships between genes resulting in perhaps new ontogenic regulatory regions (Davidson, 1982).

The genus Strongylocentrotus diverged from Lytechinus 150 to 200 myr ago whereas the two species, S. franciscanus and S. purpuratus total divergence is only 10 to 20 myr. These species were used to study the evolution of repeat DNA in sea urchins. It was found (Moore et al., 1978) that repeat DNA with a reiteration frequency greater than 1000 copies in the Strongylocentrotus genome could not reassociate with any of the repeat DNA from Lytechinus nor could the reciprocal reaction occur. Furthermore, individually cloned members of repeat families in S. purpuratus could only hybridize to a few copies in Lytechinus or not at all. Therefore, apparently the dominant repetitive sequence families in one genome are more prevalent than in a related species. Copies of repeats are being added or eliminated in the two species at a sufficient rate to dominate the S. purpuratus genome compared to Lytechinus or even to the closer relative, S. franciscanus. Comparisons between repeat families in these genomes also showed that there was no correlation between sequence divergence and loss of copy number with evolutionary time. Thermal denaturation characteristics

between homologous and heterologous species showed that repeat DNA which lost copy number so rapidly with evolutionary time maintained sequence homology better than single copy DNA in any of the species.

The genomic organization of three different kinds of individual repeat families from S. purpuratus was determined in the sea urchin species (Klein et al, 1978; Anderson et al., 1981; Scheller et al., Posakony et al., 1981). The CS2109 family contains 1000 members which are 200 to 300 base pairs in length and are interspersed amongst single copy DNA. The T_m of native duplex is 83.4°C and that of duplex reassociated at 50°C with S. purpuratus and S. franciscanus genomic DNAs is 56.8 and 58.8°C, respectively. CS2108 is a more complex family of repeats with 4.5 kb long members, interspersed throughout the genome and with a native T_m of 86.7°C. There are 20 copies of this family found under stringent conditions of reassociation but many more are observed when these conditions are relaxed. The long repeats were found to consist of scrambled arrangements of subelements between family members. Thirdly, the CS2034 family of long repeats (greater than 2.0 kb) is reiterated 2500, 160 and 10 times in S. purpuratus, S. franciscanus and L. pictus, respectively. Furthermore, thermal denaturation of reassociated duplex at 50°C between the repeat family and the homologous genome, S. purpuratus show that there

is less sequence divergence than that found in single copy DNA. The native T_m is 79.4^oC and with driver DNAs from S. purpuratus, S. franciscanus and L. pictus are 76.0, 70.5, and 57.0^oC, respectively. The family is found in S. purpuratus in one large deposit, some small clusters that are dispersed and even some single occurances.

Of the three distinct repeat families described in the sea urchin genome, CS2034 is somewhat similar to the pPol431 family of repeats found in Pisaster ochraceus with important differences. The native T_m of the repeat duplex in the sea star is 11^oC higher than in sea urchin indicating a much higher GC content. There are many more copies of the CS2034 family in S. purpuratus than the number of copies of pPol431 in P. ochraceus. Copy number is better conserved in starfish since there are still 80 copies found in Dermasterias which is 500 myr removed from P. ochraceus whereas in sea urchin there are only 10 copies in L. pictus which is 200 myr removed from S. purpuratus. Furthermore, the clustered arrangement is more conserved in starfish than in the sea urchin. Most of the family members and perhaps all are found in either one large cluster or a few large clusters whereas in sea urchin even single family members are found interspersed although a larger number of CS2034 repeats are clustered. Lastly, the sequence of the repeat family has diverged 22% from S. purpuratus to L.

pictus but only 16% sequence divergence was observed in the Dermasterias genome compared to P. ochraceus. Therefore, it appears that the sea urchin repeat DNA represented by CS2034 is amplified and dispersed faster and its sequence and organization are less conserved than the pPol431 family in sea stars.

The members of the sea urchin repeat families are all found in RNA populations in various tissues or in different developmental stages of sea urchin. For example, the CS2108 repeat family transcripts have been found in oocyte RNA, gastrula and intestinal nuclear RNA with the most found in the oocyte. CS2034 was represented in intestinal nuclear RNA. However, unless there are members of the sea star pPol431 family which are transcribed in other stages than oocyte then this repeat DNA is not represented in any RNA populations of the organisms.

The three families of repeats described in the sea urchin can be used as an illustration of one model of repeat DNA evolution. In an early stage of growth of a new family of repeats a sequence is amplified in a tandem arrangement in the genome. An example could be CS2034 which has one large cluster of repeats and appears to be recently amplified. This is probably true since the very large cluster of repeats is not found in S. franciscanus. By some mechanism copies of the cluster begin to disperse throughout the genome. The family disperses to such an

extent that individual and isolated copies of the repeat sequence in the genome can no longer amplify by crossing-over mechanisms. These repeats which are found dispersed amongst single copy DNA eventually become so divergent as to become single copy. For example, CS2109 family is similar to this situation since its family members are short and completely interspersed amongst single copy DNA. Furthermore, its sequence homology among family members is much less conserved (35% sequences divergence in S. franciscanus) than those in CS2034 (9% sequence divergence in S. franciscanus). Therefore CS2109 may represent a very ancient family in the sea urchin genome.

Two major issues should be addressed if this model is applicable to the evolution of repeat DNA in eukaryotes. The first is that the genome would continue to become larger as the repeat families amplify, disperse and evolved into single copy sequence. Genome size may vary depending on the amount of repeat DNA present and that as the repeat classes grow, so will the genome at a similar rate. This appears to be fatal to the cell since if there is no loss of repeats during some stage of evolution and if the repeat families are dispersing throughout the genome then eventually they will interrupt important coding regions. Meanwhile repeats may also insert into other repeats or stay in large clusters resulting in all organisms eventually possessing a long period interspersion pattern of sequence

organization in their genomes.

A second issue which the model does not explain is why there is a decreasing number of repeats of a given family in related species with increasing phylogenetic distance from P. ochraceus. It may be possible that some of the repeats which begin to disperse may be lost at the time of recombination or during cell division. If there are repeats which are dispersed and lost randomly then it would be difficult to explain how this phylogenetic relationship could be maintained during evolution. However, if the repeats are lost or simply stop amplifying during speciation while other repeat families are retained and continue to amplify, then perhaps the phylogenetic relationship can be explained. In any case, during the process of changing the dominant repeat DNA families within the genome some repeat sequences must be lost since the genome size would continue to increase as new families are created.

On the other hand, the behavior of the repeat families described in sea urchin and starfish may be explained by the model of repeat amplification and dispersion described above. According to the model, new families of repeat DNA can be generated by recent amplification events of a pre-existing sequence and are therefore found in clustered arrangements within the genome. These families usually possess a high percent of sequence

homology. Depending on the repeat DNA family and perhaps even the species, the rate of amplification can be variable. The family members begin to disperse elsewhere in the genome, by some unknown mechanism, reducing the size of the cluster. These dispersed repeats may then lose sequence homology with other family members and eventually become single copy. The rate of dispersion may again depend on the nature of the repeat DNA family and on the species.

Repeat DNA dispersion or perhaps loss of newly amplified members of CS2108 or CS2109 and loss of sequence homology in the S. purpuratus sea urchin may have occurred so rapidly that homologous sequences are not found in the closely related S. franciscanus genome. Evidently, repeat families evolve at different rates since the CS2034 family is still present as a tandem cluster in the parent S. purpuratus genome and yet is found in S. franciscanus (160 copies) and in the most distantly related species, L. pictus (10 copies). The repeat family from starfish could also be evolving in a manner described by this model except that the members must be dispersing at an extremely slow rate since Dermasterias still contains a large number of its repeats in a clustered arrangement. Therefore, there appears to be a difference in the rate of evolution of repeat families where some are perhaps selectively maintained in a clustered arrangement during evolution.

When family members begin to disperse there seems to be a corresponding rearrangment of subelements in clusters of the repeats. It would be important to know if either Pisaster or Dermasterias repeat members are scrambled within their large genomic clusters to ascertain if scrambled rearrangement of subelements is indicative of dispersion. If this were so, then Dermasterias should contain dispersed members and clusters which are scrambled and certainly different compared to P. ochraceus clusters. The Pisaster species would be expected to be fairly homologous if most of its members were recently amplified. The amount of sequence homology among repeat members and the few Southern blot bands found in the Pisaster genome indicates that the family is indeed homogeneous. Dermasterias Southern blots reveal many more secondary bands with long exposure and the restriction enzyme sites are not the same in the genomic cluster as were in found in the Pisaster species. This suggests that the clusters in Dermasterias are very different from the Pisaster species and may even be in some rearranged form of subelements.

Eden (Eden, 1980; Eden and Hendrick, 1978) studied the chicken genome in order to define the properties of long and moderately repetitive DNA in a eukaryotic genome. The chicken genome sequence organization can be described as containing the long-period interspersion pattern of repeat DNA throughout

single copy DNA. The repeat DNA is found in long continuous regions greater than 20kb in length which are adjacent to single copy DNA that is usually at least 17 kb before the next repeat cluster. Therefore, this genome is similar to Drosophila except that its long interspersion pattern contains excessively long regions of repeat and single copy DNA. A repeat DNA class was cloned (Eden et al., 1980) and individual families were isolated. Furthermore, some of the repeat clones (for example, pGg132) were used to screen the chicken genomic library to locate these large clusters in the genome. Analysis of the genomic clones showed that the repeat DNA is found in large clusters which contain smaller subelements which are rearranged with respect to other members of the family (Musti et al., 1981). Furthermore, regions which were not rearranged were found to be methylated in 5'-CG-3' sites detected using the isoschizomers MspI and HpaII (Eden et al., 1981). The cluster found in the chicken genome represented 10% of the the total repeat DNA. This kind of rearranged clustered repeat elements is very similar to the class of sea urchin repeat represented by CS2108. It is interesting that when comparing these two organisms the chicken repeat cluster is not dispersed and contains methylated sites whereas the sea urchin repeats are found in 4.5 kb regions interspersed throughout the genome. The urchin was not assayed for the presence of methylation in the repeat family.

The conservation of the repeat family within related species was determined by Southern blot analysis of DNA from birds in the Order Galliformes and from the duck and ostrich which are from different orders. The birds within the order all show reaction with the cloned repeat and furthermore it was found that some of the methylation sites were still conserved. On the other hand, the specific cloned repeat could not cross react with the distantly related species, the duck or the ostrich, although total repeat DNA from the chicken could hybridize with their total genomic DNA. It was suggested that the organization of the repeat DNA in a genome has more evolutionary constraints than the sequence of any particular repeat family itself. The evolutionary distance between or within the avian orders is not known so it is not possible to know whether the chicken repeat family is evolving at the same rate as any of the sea urchin families. It may be possible that the family of repeats found in chicken is old since it consists of scrambled subelements found in related species but it cannot disperse to other regions of the genome. The family may be newer than the repeat families which are dispersed in the genome although it could be that this group of repeats will not disperse because there is a genome size constraint on its mobility. Therefore it remains in a cluster which becomes rearranged with time as was observed in the related species. However, the sea urchin repeat family CS2034 is not as ancient as the CS2108 or the chicken family

since it does not show any rearrangement of subelements nor is it dispersed completely in the genome.

The large deposit of repeats from the chicken is similar to that found in the P. ochraceus repeat family. There are few detectable copies of pPol431 repeats found elsewhere in the genome. If the cluster is one large tandem array then it would cover 25 kb of the genome. There is also a methylated site in the repeat family although none were observed in the sister species or in Dermasterias. The starfish repeat family is unlike the chicken clustered repeats in many ways. The pPol431 family represents only about 2% of the total repeat DNA in the genome, it does not appear to be in rearranged forms of subelements, its sequence and genomic organization are more conserved than that observed in the avian species, although the divergence times between the avian orders is unknown, and the starfish repeats are not extensively methylated.

The evolution of repeat DNA in sea urchin, chicken and starfish have some common properties. The starfish is unusual in that it has large repeat clusters that are conserved over 850 myr of divergence time but are not rearranged in any apparent manner. The sea urchin and chicken genomes contain large clusters also but some of the chicken cluster is extensively rearranged like the CS2108 dispersed family in the sea urchin but unlike the

large cluster of the CS2034 family. This CS2034 has no subelemental reorganization amongst members and is similar to starfish. This family may have been newly amplified in the sea urchin genome since large deposits were not found in the closely related species, S. franciscanus.

A number of different kinds of long repeat DNA have been studied in Drosophila. Some of these are found in very large tandemly duplicated deposits of repeat DNA, others are dispersed about the genome, are mobile and transcribed. It has been shown that this latter class of repeats have some functional and structural properties which are used by the cell. Transposable elements in prokaryotes have been well characterized and understood for some time and more recently they were also found to exist in Drosophila. Some families of repeats are called copia, 412 and 297 (Potter et al., 1979) which are about 3 kb in length, reiterated 100-fold in the genome and are transcribed. Other kinds of possible transposable elements which are not transcribed but are found dispersed throughout the genome have also been characterized in Drosophila. Clustered repeats are also commonly found in the Drosophila genome. There is a large cluster of repeats in the nontranscribed spacer region in the 18S rDNA locus. The majority of spacers in D. melanogaster are 3 to 5 kb in length consisting of short 200 to 250 bp long tandem repeats. Interspecies comparisons demonstrate that there

is length heterogeneity within the short tandem elements although there is enough sequence homology conserved that spacer rDNA from various Drosophila genomes can cross hybridize in Southern blot analyses (Rae et al., 1981). Conservation of repeat DNA in the nontranscribed spacers of rDNA may not be universal since the rDNA of the amphibian sibling species Xenopus laevis and X. borealis contain no detectable sequence homology (Brown et al., 1972). Other clusters, containing repeat elements 500 bp long and reiterated about 100 times, are found in the genome. There are approximately 52 different kinds of these repeat families which are rearranged in different orders in different clusters (Wensink et al., 1979).

The mechanism for transposition of repeat elements in prokarotes is analogous to that for copia, 297 and 412 families of repeats since they have similar structure to the prokaryotes (Calos and Miller, 1980; Kleckner, 1977). However, this similarity is not observed in other dispersed repeats in Drosophila. What have been found however, are episomes in the fly nucleus of cultured cells which are circularized clusters of repeats (Stanfield and Lengyel, 1979). This repeating unit may find homologous sequences in the genome to which they can base pair and recombine inserting in many different regions resulting in the scrambled rearrangement of repeat elements in the clusters. The discovery of these repetitive episomes may provide an

explanation for the mechanism of rearrangement of repeat sequences in clusters and for their dispersal about the genome. Krolewski et al. (1982) have characterized small polydisperse circular DNAs containing AluI sequences found in African Green Monkey kidney cells. There are no short direct repeats flanking the AluI sequences so that the mechanism of genomic excision and perhaps reinsertion would be unlike that described for transposable elements (Shapiro, 1979). It would be interesting if a similar mechanism for movement of repetitive elements about the genome could be applied to the animals described above (Wigler et al., 1979).

Genomic DNA in plants have an incredible amount of repeat DNA, up to 80%, compared to most organisms whose repeat DNA has been assessed. The telomeric repeat DNA alone constitutes 10% of the genome in Secale cereale (Bedbrook et al., 1980a). This heterochromatin consists of four different classes of repeat families of which 45-66% are not found in closely related species such as S. silvestre. The S. cereale families are 100 to 500 bp long and are tandemly arranged in a scrambled organization. A 120 bp family of repeats found in the telomere of S. cereale is also found in interstitial DNA and in an uniform, tandem arrangement. This same family is conserved and found in about equal proportions in S. silvestre and in the wheat, Triticum dococcum (Bedbrook et al., 1980b).

Approximately, 30 to 50% of the repeat DNA from rye, wheat, barley and oats consists of short repeats which are species specific but also interspersed with other families of short repeats which are common to the five different species (Bedbrook et al., 1980a).

Priestler and Thompson (1981a,b) had made some interesting comparisons between the repeat DNA of two legumes, the mung bean and the pea. The pea has a haploid DNA complement nine times greater than the mung bean. The pea also has about twice the amount of repeat DNA in mung bean which is interspersed throughout single copy DNA generating single copy regions no greater than 1000 bp. The mung bean on the other had has long single copy stretches in its genome, about 6700 bp in length.

The properties of the repeat DNA from these plant genomes were determined using thermal denaturation techniques. From the results a mechanism of repeat DNA evolution was proposed. They suggested that repeat DNA was amplified at a higher rate in the pea than in the mung bean which then disperses into other regions of the genome. If amplification is continued then eventually repeat elements from distant regions will insert into other repeat clusters in the genome generating a scrambled arrangement of repeat elements in clusters. Eventually, if one genome is not amplifying and dispersing at a similar rate then

the thermal denaturation characteristics of repeat DNA would indicate a more homogeneous melt. This was observed in the pea and the mung bean where more homogeneous melts were found in the mung bean compared to the pea. With increasing evolution time the dispersed repeats evolve into single copy DNA.

An interesting aspect of the evolution of repeat DNA in the pea is observed when the thermal renaturation temperature is lowered permitting more duplex formation to occur. At a criterion of 35^{o}C below the optimal T_m virtually all the single copy DNA behaved like repeat DNA. The mung bean was treated similarly and found to contain about one-third of its single copy DNA in fossil repeats. Therefore, it appears at least in this system that the rate of repeat DNA turnover, amplification and loss of repeat family homology, is very rapid (Preisler and Thompson, 1980b, Murray et al., 1981).

The evolution of repeat DNA in plants may be analogous to that in animals. All eukaryotes have been shown to contain repeat DNA elements that amplify and disperse in the genome. The scrambled arrangement of subelements in clusters may be a key step in the process of evolution. For example, rearrangement of subelements may reflect a stage of repeat DNA evolution in that if scrambled repeats are observed in a particular family then it is often found dispersed throughout the genome. It should be

stressed that there are many different kinds of repeat families in the animal and plant genomes and that those discussed here were selected to illustrate a possible model described for the evolution of repeat DNA. It appears that the differences between repeat families within and between species are due to differences in rates of amplification and dispersion of repeat family members. This mechanism may be ultimately controlled by genome size.

A very elegant use of repeat DNA has been made by the cell. Telomeric sequences contain tandemly repeating units which are 3 kb in length. These telomeres are found to mediate associations of chromosomes by preventing fusion and allowing some transient, end-to-end association of nonhomologous chromosomes during interphase and meiotic prophase in plants, insects and mammalian nuclei (Rubin, 1977). The telomeres of the five major chromosomal arms in Drosophila contain common but not identical repeat families. A cloned repeat, Dm356 was found in the telomeric sequences and also could hybridize with members from the copia family. The clone was made from RNA so it could be that transposable elements which are transcribed are also found in the telomeric sequences. The telomere regions of the chromosomes also have thin fibres of DNA extending from them which connected to other chromosomes including the polytenes. These fibres also have 3 kb repeats and could react with the

cloned repeat from Dm356.

The centromere contains very short tandem repeats which are characterized as satellite DNA and may be structually important to the cell during replication or division (Rubin, 1977). Satellite sequences are not usually well conserved (Lee, 1981; Lohe, 1981) although some mouse satellite DNAs have been conserved within the Genus Apodemus and the Genus Mus (Brown and Dover, 1979; 1980). Restriction enzyme digestion of satellite DNA results in the formation of multimeric banding patterns observed in agarose gels due to loss or gain of restriction enzyme sites. Frequent reamplification of rearranged repeat families generates sequence heterogeniety commonly found in centromeric satellite DNA.

The repeat DNA in starfish and represented by pPol431 may be from the hetochromatin regions of chromosomes. However, both the centromeric and telomeric sequences are rarely very well conserved over evolutionary times. Restriction enzyme digestion of sea star DNA does not result in these multimeric banding patterns found in satellite DNA nor do the cloned repeats digested by a variety of restriction enzymes produce a ladder effect in agarose gels. Therefore, the only common feature of the starfish DNA with heterochromatic repeat DNA is that both are tandemly arranged in the genome.

There are many different kinds of repeat families found in advanced organisms and it may be imagined that during speciation, specific families are perhaps selectively amplified while others are stopped or members are lost from an ancestral cluster. If there is a limited tolerance for repeat DNA in the genome then only a restricted number of repeat DNA families may expand, disallowing others to amplify in that genome. Why or how repeat families could be selected for, especially during the speciation process, can only be speculated upon, although from these starfish studies of the pPol431 family it is clear that when the starfish orders diverged, the family was no longer amplified except in the Pisaster lineage.

Summary.

1. Pisaster ochraceus genomic DNA isolated from sperm was restriction enzyme digested with BamHI which produced fragments 3.9 kb in length observed as a band in an electrophoretic gel.

2. The band of fragments were cloned into the tetracycline resistance gene of pBR322. Positive transformants were selected, their plasmid DNA amplified and isolated.

3. Plasmids containing 3.9 kb inserts were designated pPol431, pPol289, pPol675 and pPo328. Restriction enzyme analysis showed

that pPol431 and pPol675 3.9 kb inserts were in the reverse orientation to that of pPol289 and pPo328 inserts.

4. EcoRI, HincII and AvaI restriction enzyme sites were mapped in the repeat DNA of pPol431.

5. The four cloned repeats have identical EcoRI restriction sites but HaeIII digestion shows that these repeats are different since the banding patterns varied in polyacrylamide gel electrophoresis.

6. Reassociation kinetics at 60^{o}C and 0.12 M Na phosphate buffer show that there are 500 copies of the repeat members in P. ochraceus, 300 copies in P. brevispinus, 130 in Evasterias, 190 in Pycnopodia, and 80 copies in Dermasterias.

7. Thermal denaturation chromatography indicated that the repeat family sequence is more conserved than single copy DNA.

8. At the 60^{o}C reassociation criterion there is approximately 2.6, 7.2, 7.2 and 12.1% sequence divergence in the repeat family in the related species P. brevispinus, Evasterias, Pycnopodia and Dermasterias, respectively. Lowering the temperature of reassociation to allow hybridization of more divergent members in the family in all five species did not produce a measurable

difference between the T_m values observed at 60°C criterion.

9. There is a fold-back component in the 3.9 kb repeat which reassociates at low Cot values and binds 10-15% of ^{32}P-labeled pPol431 insert to hydroxyapatite columns.

10. The total length of the repeating element in the Pisaster ochraceus and Pisaster brevispinus genomes is 5.3 kb demonstrated by Southern transfer of restriction enzyme digested genomic DNA reacted with ^{32}P-labeled repeat DNA from pPol431.

11. Southern blot analysis also showed that the repeat members are clustered in P. ochraceus, P. brevispinus and Dermasterias with perhaps some dispersion indicated by autoradiograph bands appearing away from the main cluster of repeats.

12. Methylation of '5-CG-3' sites at the cytosine residue was detected in P. ochraceus only after digestion of genomic DNA from P. ochraceus, P. brevispinus and Dermasterias with MspI and HpaII. Southern transfer and hybridziation with repeat probe from pPol431 produced the same banding patterns in P. brevispinus and Dermasterias but there was one 1.6 kb band in the HpaII digest of P. ochraceus which was not observed in the Msp I digest.

13. The repeat family is not found in egg RNA nor is it represented in ribosomal sequences.

REFERENCES CITED

Anderson, D.M., R.H. Scheller, J. W. Posakony, 1.B. McAllister, S.G. Trabert, C. Beall, R.J. Britten and E.H. Davidson. 1981. Repetitive sequences of the sea urchin genome. J. Mol. Biol. 145:5-28.

Angerer, R.C., E.H. Davidson and R.J. Britten. 1975. DNA sequence organization in the Mollusc Aplysia Californica. Cell 6:29-39.

Bedbrook, J. R., J. Jones, M. O'Dell, R.D. Thompson and R.b. Flavell. 1980a. A molecular description of telomeric heterochromatin in Secale species. Cell 19: 545-560.

Bedbrook, J.R., M. O'Dell and R.B. Flavell. 1980b. Amplification of rearranged repeated DNA sequences in cereal plants. Nature 288:133-137.

Benton, W.D. and R.W. Davis. 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196:180-182.

Blattner, F.R., H.E. Faber, L. Furlong, D.J. Grunwald, D.O.

Kiefer, D.D. Moore, J. W. Schumm, E.L. Sheldon and O. Smithies. 1977. Charon phages: safer derivatives of bacteriophage lambda of DNA cloning. Science 196:161-169.

Blin, N. and D.W. Stafford. 1976. A general method for isolation of high molecular weight DNA from eukaryotes. Nucleic Acids Research 3:2303-2308.

Bolivar, F., R.L. Rodriguez, P.J. Greene, M.C. Betlach, H.L. Heyneker and H.W. Boyer. 1977. Construction and characterization of new cloning vehicles. Gene 2:95-113.

Britten, R.J. 1981. DNA sequence organization and repeat sequences. Chromosomes today 1:9-23.

Britten, R.J. 1982. Genome evolution in "Evolution and Development", J.T. Bonner (ed.), pp.41-64, Dahlem Konferenzen, Berlin.

Britten, R.J. and E.H. Davidson. 1969. Gene regulation for higher cells:a theory Science 165:349-357.

Britten, R.J. and E.H. Davidson. 1971. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. The Quarterly Review of Biology

46:111-133.

Britten, R.J., D.E. Graham, F.C. Eden, D.M. Painchaud and E.H. Davidson. 1976. Evolutionary divergence and length of repetitive sequences in sea urchin DNA. J. Mol. Evol. 9:1-23.

Britten, R.J., D.E. Graham and B.R. Neufeld. 1974. Ananlysis of repeating DNA sequences by reassociation in "Methods in Enzymology", L. Grossman and K. Moldave (eds.) Vol. 12B, pp. 195-206, Academic Press, New York.

Britten, R.J. and D.E. Kohne. 1968. Repeated sequences in DNA. Science 161:529-540.

Britten, R.J. and J. Smith. 1970. A bovine genome. Carnegie Inst. Yearbook, 68:378-386.

Brown, D.D., P.C. Wensink and E. Jordan. 1972. A comparison of the rDNAs of X. laevis and X. mulleri: the evolution of tandem genes. J. Mol. Biol. 63:57-73.

Brown, S.D.M. and G.A. Dover. 1979. Conservation of sequences in related genomes of Apondemus: constraints on the maintenance of satellite DNA sequences. Nucleic Acids Research 6:2423-2434.

Brown, S.D.M. and G.A. Dover. 1980. Conservation of segmental variants of satellite DNA of Mus musculus in a related species: Mus spretus. Nature 285:47-49.

Brown S.D.M. and G. Dover. 1981. Organization and evolutionary progress of a dispersed repetitive family of sequences in widely separated rodent genomes. J. Mol. Biol. 150:441-466.

Calos, M.P. and J.H. Miller. 1980. Transposable elements. Cell 20:579-595.

Chamberlain, M.E., R.J. Britten and E.H. Davidson. 1975. Sequence organization of Xenopus DNA studied by electron microscopy. J. Mol. Biol. 96:317-333.

Clewell, D.B. 1972. Nature of ColE1 plasmid replication in E. coli in the presence of chloramphenicol. J. Bacteriology 110:667-676.

Cochet, M., F. Gannon, R. Heu, L. Maroteaux, F. Perrin and P. Chambon. 1979. Organization and sequence studies of the 17-piece chicken conalbumin gene. Nature 282:567-574.

Crain, W.R., F.C. Eden, W.R. Pearson, E.H. Davidson and R.J. Britten. 1976. Absence of short period interspersion of

repetitive and non- repetitive sequences in the DNA of Drosophila melanogaster. Chromosoma 56:309-326.

Curtiss III, R., M. Inoue, D. Pereira, J.C. Hsu, L. Alexander and L. Rock. 1977. Construction and use of safer bacterial host strains for recombinant DNA research in "Molecular Cloning of Recombinant DNA", W.A. Scott and R. Werner (eds.), pp. 99-114, Academic Press, London.

Davidson, E.H. 1976. "Gene Activity in Early Development", pp. 187-200, Academic Press, New York.

Davidson, E.H. 1982. Evolutionary change in genomic regulatory organization: speculations on the origins of novel biological structure in "Evolution and Development", J.T. Bonner (ed.), pp. 41-64, Dahlem Konferenzen, Berlin.

Davidson, E.H., B.R. Hough, C.S. Amenson and R.J. Britten. 1973. General interspersion of repetitive with nonrepetitive sequence elements in the DNA of Xenopus. J. Mol. Biol. 77:1-24.

Deininger, P.L. and C.W. Schmid. 1979. A study of the evolution of repeated DNA sequences in primates and the existence of a new class of repetitive sequences in primates. J. Mol. Biol. 127:437-460.

Denhardt, D.T. 1966. A membrane-filter technique for the detection of complementary DNA. Bioch. Biophys. Res. Comm. 23:641-646.

Doolittle, W.F. 1982. Middle repetitive DNAs in "DNA and Evolution: Natural Selection and Genome Size", T. Cavalier-Smith (ed.), John Wiley and Son, Ltd., Chichester, in press.

Doolittle, W.F. and C. Sapienza. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284:601-603.

Dover, G.A. 1980. Ignorant DNA? Nature 285:618-620.

Dover, G.A., S. Brown, E. Coen, J. Dallas, T. Strachan and M Trick. 1982. The dynamics of genome evolution and species differentiation in "Genome Evolution", G.A. Dover and R.B. Flavell (eds.), pp. 342-372, Academic Press, London.

Dover, G.A. and W.F. Doolittle. 1980. Modes of genome evolution. Nature 288:646-647.

Eden, F.C. 1980. A cloned chicken DNA fragment includes two repeated DNA sequences with remarkably different genomic organization. J. Biol. Chem. 255:4854-4863.

Eden, F.C., R.T. Burns and R.F. Goldberger. 1980. Complicated organization of a single repeated DNA sequence in the chicken genome is revealed by cloning. J. Biol. Chem. 255:4843-4853.

Eden, F.C. and J.P. Hendrick. 1978. Unusual organization of DNA sequences in the chicken. Biochemistry 17:5838-5844.

Eden, F.C., A.M. Musti and D.A. Sobieski. 1981. Clusters of chicken repeated DNA sequences are extensively methylated but contain specific undermethylated regions. J. Mol. Biol. 148:129-252.

Enea, V., G.F. Vovis and N.D. Zinder. 1975. Genetic studies with heteroduplex DNA of bacteriophage fl. Assymmetric segregation base corrections and implications for the mechanism of genetic recombination. J. Mol. Biol. 96:495-509.

Flavell, R.B., M. O'Dell and J. Hutchinson. 1981. Nucleotide sequence organization in plant chromosomes and evidence for sequence translocation during evolution. Cold Spring Harb. Symp. quant. Biol. 45:501-508.

Galau, G.A., M.E. Chamberlain, B.R. Hough, R.J. Britten and E.H. Davidson. 1976. Evolution of repetitive and nonrepetitive DNA in "Molecular Evolution", F.J. Ayala (ed.), pp. 200-224, Sinauer Assoc.

Goldberg, R.B., W.R. Crain, J.V. Ruderman, G.P. Moore, T.R. Barnett, R.C. Higgins, R.A. Gelfand, G.A. Galau, R.J. Britten and E.H. Davidson. 1975. Sequence organization in the genomes of five marine invertegrates. Chromosoma 51:225-251.

Graham, D.E., B.R. Neufeld, E.H. Davidson and R.J. Britten. 1974. Interspersion of repetitive and nonrepetitive DNA sequences in the sea urchin genome. Cell 1:127-137.

Hinegarder, R. 1974. Cellular DNA content of the echinodermata. Comp. Biochem. Physiol. 49B:219-226.

Jeffreys, A.J. 1981. Recent studies of gene evolution using recombiant DNA in "Genetic Engineering 2", R. Williamson (ed.), pp. 1-49, Academic press, London.

Jelinek, W.R. and c.W. Schmid. 1982. Repetitive sequences in eurkaryotic DNA and their expression in "Annual Reviews of Biochemistry", E.S. Snell (ed.), vol. 51, pp. 813-844, Annual Reviews Inc., Palo Alto, California.

Kleckner, N. 1977. Translocatable elements in procaryotes. Cell 11:11-23.

Klein, W.H., T.L. Thomas, C. Lai, R.H. Scheller, R.J. Britten

and E.H. Davidson. 1978. Characteristics of individual repetitive sequence families in the sea urchin genome studied with cloned repeats. Cell 14:889-900.

Krolewski, J.J., A.H. Bertelsen, M.Z. Humayun and M.G. Rush. 1982. Members of the alu family of interspersed repetitive DNA sequences are in the small circular DNA population of monkey cells grown in culture. J. Mol. Biol. 154:399-415.

Lee, C.S. 1981. Restriction enzyme analysis of a highly diverged satellite DNA from Drosophil nasutoides. Chromosoma 83:367-379.

Lohe, A.R. 1981. The satellite DNAs of Drosophila simulans. Genet. Res. 38: 237-250.

Long, E.O. and I.B. Dawid. 1980. Repeated genes in eukaryotes in "Annual Reviews of Biochemistry", E.E. Snell, P.D. Boyer, A. Meister, C.C. Richardson (eds.), vol. 49, pp. 727-757, Annual Reviews Inc., Palo Alto, California.

Mandel, M. and A. Higa. 1970. Calcium dependent bacteriophage DNA infection. J. Mol. Biol. 53:159-162.

Maniatis, T., R.C. Hardison, E. Lacy, J. Lauer, C. O'Connell and D. Quon. 1978. The isolation of structural genes from libraries

of eukaryotic DNA. Cell 15:687-701.

Manning, J.E., C.W. Schmid and N. Davidson. 1975. Interspersion of repetitive and nonrepetitive DNA sequences in the Drosophila melanogaster genome. Cell 4:141-155.

Marmur, J. 1961. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J. Mol. Biol. 3:208-218.

McDonnell, M.W. 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. 110:119-146.

Moore, G.P., F.D. Constantini, J.W. Posakony, E.H. Davidson and R.J. Britten. 1980. Evolutionary conservation of repetitive sequence expression in sea urchin egg RNAs. Science 208:1046-1048.

Moore, G.P., W.R. Pearson, E.H. Davidson and R.J. Britten. 1981. Long and short repeats of sea urchin DNA and their evolution Chromosoma 84:19-32.

Moore, G.P. R.H. Scheller, E.H. Davidson and R.J. Britten. 1978. Evolutionary change in the repetition frequency of sea urchin

DNA sequences. Cell 15:649-660.

Moyzis, R.K., J. Bonnet, D.W. Li and P.O. Ts'o. 1981. An alternative view of mammalian DNA sequence organization. J. Mol. Biol. 153:841-864.

Moyzis, R.K. J. Bonnet, D.W. Li and P.O. Ts'o. 1981. Short repetitive sequences are organized into scrambled tandem clusters in Syrian hamster DNA. J. Mol. Biol. 153:871-896.

Murray, M.G., D.L. Peters and W.F. Thompson. 1981. Anceint repeated sequences in the pea and mung bean genomes and implications for genome evolution. J. Mol. Evol. 17:31-42.

Musti, A.M., D.A. Sobieski, B.B. Chen and F.C. Eden. 1981. Repeated DNA clusters in the chicken genome contain homologous sequence elements in scrambled order. Biochemistry 20:2989-2999.

Orgel, L.E. and F.H.C. Crick. 1980. Selfish DNA: the ultimate parasite. Nature 284:604-607.

Orgel, L.E., F.H.C. Crick and C. Sapienza. 1980. Selfish DNA. Nature 288:645-646.

Pearson, W., E.H. Davidson and R.J. Britten. 1977. Program for

least squares analyses of reassociation and hybridization data. Nucleic Acids Research 4:1727-1737.

Posakony, J.W., r.H. Scheller, d.M. Anderson, R.J. Britten and E.H. Davidson. 1981. Repetitve sequences of the sea urchin genome. J. Mol. Biol. 149:41-67.

Potter, s.S., W.J. Brorein, Jr., P. Dunsmuir and G.M. Rubin. 1979. Transposition of elements of the 412, copia and 297 dispersed repeated gene families in Drosophila. Cell 17:415-427.

Preisler, R.S. and W.F. Thompson. 1981a. Evolutionary sequence divergence within repeated DNA families of higher plant genomes. J. Mol. Evol. 17:78-84.

Preisler, R.S. and W.F. Thompson. 1981b. Evolutionary sequence divergence within repeated DNA families of higher plant genomes. Analysis of reassociation kinetics. J. Mol. Evol. 17:85-93.

Rae, P.M.M., T. Barnett and V.L. Murtif. 1981. Nontranscribed spacers in Drosophila ribosomal DNA. Chromosoma 82:637-655.

Rigby, P.W.J., M. Dieckmann, c. Rhodes and P. Berg. 1977. Labelling DNA to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. 113:237-251.

Roberts, R.J. 1980. Directory of restriction enzymes in "Methods in Enzymology", S.P. Colowick and N.O. Kaplan (eds.), vol. 68, pp. 1-15, Academic Press, New York.

Rubin, G.M. 1977. Isolation of a telomeric DNA sequence from Drosophila melanogaster. Cold Spring Harb. Symp. quant. Biol. 42:1041-1051.

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. J. Mol. Biol. 149:15-39.

Scheller, R.H., F.D. Costantini, M.R. Kozlowski, R.J. Britten and E.H. Davidson. 1978. Specific representation of cloned repetitive DNA sequences in sea urchin RNAs. Cell 15:189-203.

Scheller, R.H., T.L. Thomas, A.S. Lee, W.J. Klein, W.D. Niles, R.J. Britten and E.H. Davidson. 1977.Clones of individual repetitive sequences from sea urchin DNA constructed with synthetic EcoRI sites. Science 196:197-200.

Shapiro, J.A. 1979. Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. Proc. Nat. Acad. Sci. U.S.A. 76:1933-1937.

Sharp, P.A., B. Sugden and J. Sambrook. 1973. Detection of two restriction endonuclease activities in Haemophilus parainfluenza using analytical agarose-ethidium bromide electrophoresis. Biochemistry. 12:3055-3063.

Smith, G.P. 1974. Unequal crossover and the evolution of multigene families. Cold Spring Harb. Symp. quant. Biol. 38:507-513.

Smith, M.J. and R. Boal. 1978. DNA sequence organization in the common Pacific starfish Pisaster ochraceus. Canad. J. Biochem. 56:1048-1054.

Smith, M.J., R.J. Britten and E.H. Davidson. 1975. Studies on nucleic acid reassociation kinetics: reactivity of single-stranded tails in DNA-DNA renaturation. Proc. Nat. Acad. Sci. U.S.A. 72:4805-4809.

Smith, M.J., A. Lui, K.K. Gibson and J.K. Etzkorn. 1980. DNA sequence organization in the starfish Dermasterias imbricata. Canad. J. Biochem. 58:352-360.

Smith, M.J., R. Nicholson, M. Stuerzl and A. Lui. 1982. Single copy DNA homology in sea stars. J. Mol. Evol. 18:92-101.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

Stanfield, S.W. and J.A. Lengyel. 1979. Small circular DNA of Drosophila melanogaster: chromosomal homology and kinetic complexity. Proc. Nat. Acad. Sci. U.S.A. 76:6142-6146.

Tartof, K. 1974. Unequal mitotic sister chromatid exchange and disproportional replication as mechanisms regulating ribosomal RNA gene redundancy. Cold. Spring Harb. Symp. quant. Biol. 38:491-500.

Thayer, R.E. 1979. An improved method for detecting foreign DNA in plasmids of E. coli. Anal. Biochem. 98:60-63.

Thomas, P.S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Nat. Acad. Sci. U.S.A. 77:5201-5205.

Weiss, B., A. Jacquemin-Sablon, T.R. Live, G.C. Fareed and C.C. Richardson. 1968. Enzymatic breakage and joining of deoxyribonucleic acid. J. Biol. Chem. 243:4543-4555.

Wensink, P.C., S. Tabata and c. Pachl. 1979. The clustered and

scrambled arrangement of moderately repetitive elements in Drosophila DNA. Cell 18:1231-1246.

Wigler, M., R. Sweet, G.K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, s. Silverstein and R. Axel. 1979. Transformation of mammalian cells with genes from prokaryotes and eukaryotes. Cell 16:777-785.

Zimmer, E.A., S.C. Maxtri, S.M. Berverly, Y.W. Korn and A.C. Wilson. 1980. Rapid duplication and loss of genes coding for the alpha chains of hemoglobin. Proc. Nat. Acad. Sci. U.S.A. 77:2158-2162.

Zuker, C. and H.F. Lodish. 1981. Repetitive DNA sequences cotranscribed with developmentally regulated Dictyostelium discodeum mRNAs. Proc. Nat. Acad. Sci. U.S.A. 78:5386-5390.