



1986

## Polymorphic Arrangement of 5s DNA Orphans in *Xenopus laevis*

Brenda A. Peculis  
*Loyola University Chicago*

Follow this and additional works at: [https://ecommons.luc.edu/luc\\_theses](https://ecommons.luc.edu/luc_theses)



Part of the [Biology Commons](#)

---

### Recommended Citation

Peculis, Brenda A., "Polymorphic Arrangement of 5s DNA Orphans in *Xenopus laevis*" (1986). *Master's Theses*. 3501.

[https://ecommons.luc.edu/luc\\_theses/3501](https://ecommons.luc.edu/luc_theses/3501)

This Thesis is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Master's Theses by an authorized administrator of Loyola eCommons. For more information, please contact [ecommons@luc.edu](mailto:ecommons@luc.edu).



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).  
Copyright © 1986 Brenda A. Peculis

**POLYMORPHIC ARRANGEMENT OF  
5S DNA ORPHONS  
IN XENOPUS LAEVIS**

by

**Brenda A. Peculis**

**A Thesis Submitted to the Faculty of the Graduate School  
of Loyola University of Chicago in Partial Fulfillment  
of the Requirements for the Degree of  
Master of Science**

**December**

**1986**



**This thesis is dedicated in memory of my father,  
Bernard F. Peculis**

## ABSTRACT

In *Xenopus laevis* the major oocyte (Xlo), trace oocyte (Xlt), and somatic (Xls) 5s genes occur as separate multi-gene families, each organized in large clusters of tandemly repeating units. A repeat consists of a 5s RNA gene plus the spacer sequence characteristic of the particular family. I have discovered additional copies of the 5s RNA genes located at a number of dispersed genomic loci, apart from the major clusters. This was done by "null-digesting" total genomic DNA with restriction enzymes that do not cut within the 5s major repeats. The main clusters remain intact as high molecular weight DNA while the dispersed 5s genes can be resolved on gels as lower molecular weight restriction fragments. Probing these smaller fragments with either a 5s gene-specific, Xlo, Xlt or Xls spacer-specific probe has revealed that the organization of the dispersed genes, or orphans, is quite different from the 5s DNAs of the major clusters. The genes appear to be present in single copies and the majority of the fragments have a polymorphic organization in the population. Four types of orphans have been identified on the basis of their structure and incidence of polymorphism. Most orphans contain the 5s gene plus Xlo spacer, while others contain just the gene region. In both cases some of the elements have organizations that appear to be conserved in the population while others appear

polymorphic. Nearly all the gene-coding regions in the 5s orphans are greater than 92% homologous to the Xlo major family. Orphans containing regions homologous to the Xlo spacer contain, at most, traces of sequence divergence with respect to the major family. Totally unique 5s DNA null digestion fragment patterns, containing all four types of orphans, have been observed in each of the individuals examined. Such extensive polymorphism has not been seen in other systems and could be the result of either very high levels of point mutation or genomic rearrangements. Characterizing the structures of single orphans should permit the basis of this genomic instability to be determined.

## ACKNOWLEDGEMENTS

I would like to thank my thesis committee, Dr. Howard Laten, and Dr. John Smarrelli for their support and advice during the course of my research and in preparation of this thesis.

A very special note of appreciation to Dr. Jeffrey Doering, my thesis director, for invaluable critiquing of this thesis, as well as the guidance and encouragement he provided during my research and in my continued studies.

A special thanks to my family for accepting, if not understanding, my leave of absence.

Last, but by no means least, a sincere thank you to Frank Fokta for helping me to get started in the lab, for "Science talks" which often were as enlightening as amusing, and just for being there.

## VITA

The author, Brenda Ann Peculis, is the daughter of Bernard F. Peculis and Barbara (Drochowski) Peculis. She was born in Chicago, Illinois on February 3, 1964.

Her elementary education was obtained in the Chicago Public School System. In 1980 she completed her secondary education, graduating with honors from Lane Technical High School in Chicago.

In September, 1980 Ms. Peculis entered Loyola University of Chicago, receiving the degree of Bachelor of Science in Biology in May, 1984. In September of 1984, she was awarded an assistantship in Biology at Loyola University of Chicago, allowing her to complete the requirements for the degree of Master of Science in 1986.

In 1985, while attending Loyola University, she was elected vice president of the Biology Graduate Student Association and became a member of the American Association for the Advancement of Science.

In 1986 Ms. Peculis gave a presentation at the Graduate Research Forum sponsored by the Loyola University Chapter of Sigma Xi. She also presented some of her work at the annual meeting for the American Society for Cell Biology.



PUBLICATIONS

Doering, J.L., Peculis, B.A., and Fokta, F.F. (1986).  
Dispersed 5s RNA genes in *Xenopus laevis*. *J.Cell Biol.*  
103, 309a

## TABLE OF CONTENTS

	page
ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	v
VITA.....	vi
PUBLICATIONS.....	vii
TABLE OF CONTENTS.....	viii
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
INTRODUCTION.....	1
<i>Xenopus</i> 5s RNA Genes.....	1
5s DNA Orphans in <i>Xenopus laevis</i> .....	8
LITERATURE REVIEW.....	11
Possible Mechanisms for Orphon Dispersal.....	11
Transposons.....	11
Retroposons.....	14
Cellular RNA Intermediates.....	18
Recombinational Events.....	21
Restriction Site Polymorphism.....	22
MATERIALS AND METHODS.....	24
DNA Isolations.....	24
Probe DNAs.....	26
Restriction Digests.....	26
DNA Transfer.....	27
Probes.....	27
Copy Number.....	28
Cloning of a 5s Orphon.....	28
Ligation and Packaging of Phage.....	28
Plating Phage.....	29
Screening Phage.....	30

RESULTS.....	33
5s DNA Orphans.....	33
Copy Number.....	55
Orphon Homology.....	59
Cloning of a 5s DNA Orphon.....	63
DISCUSSION.....	66
Characteristics of the 5s DNA Orphans.....	66
Point Mutations Versus Orphans.....	70
Orphon Dispersal Mechanisms.....	72
Transposable Elements.....	72
RNA Intermediates.....	73
Unequal Crossovers.....	75
Generation of Four Types of Orphans.....	76
Possible Role for Orphans in Development.....	79
Polymorphic Arrangement of Orphans.....	80
Characterization of a 5s Orphon.....	87
REFERENCES.....	93

LIST OF TABLES

Table	page
1. Polymorphic Arrangement of Null-Digestion Restriction Fragments Containing the 5s RNA Gene and the X1o Spacer Sequences.....	37
2. Polymorphic Arrangement of Null-Digestion Restriction Fragments Containing the 5s X1t Spacer Sequence.....	48
3. Polymorphic Arrangement of Null-Digestion Restriction Fragments Containing the 5s X1s Spacer Sequence.....	53

LIST OF FIGURES

Figure	page
1. Three 5s RNA Gene Families in <i>Xenopus laevis</i> .....	3
2. Comparison of Orphon Arrangement in Bam HI Null-Digested DNA from Four Individuals.....	35
3. Comparison of Orphon Arrangement in Eco RI Null-Digested DNA from Four Individuals.....	40
4. Comparison of Orphon Arrangement in Bgl II and Kpn II Null-Digested DNA from Two Individuals.....	42
5. Comparison of the Arrangements of Null- Digestion Restriction Fragments Containing the Xlt Spacer-Specific Region in Four Individuals.....	46
6. Comparison of the Arrangements of Null- Digestion Restriction Fragments Containing the Xls Spacer-Specific Region in Four Individuals.....	51
7. Estimation of Copy Number of 5s Gene-Coding Regions in the 5s Orphons.....	57
8. Determination of Orphon Homology to the Major Family.....	61
9. Recombination Occurring Between Orphons at the Same Locus May Contribute to Polymorphism.....	85
10. Hypothetical Orphon Used to Determine the Basis of the Polymorphic Arrangement Between Individuals.....	89

## INTRODUCTION

### *Xenopus* 5S RNA Genes

Ribosomes are organelles present in the cytoplasmic compartment of all eucaryotes and function as the site of protein synthesis. The typical eucaryotic ribosome is composed of two subunits. The large subunit contains two ribosomal RNAs (28s and 5s) complexed with 30-50 proteins. The small subunit consists of a single ribosomal RNA (18s) with 20-40 proteins. In lower eucaryotes, the genes responsible for the synthesis of the 28s and 18s ribosomal RNAs are transcribed by polymerase I and are usually arranged in tandemly repeated clusters of genes separated by non-transcribed spacer regions. In higher eucaryotes, the 5s RNA genes (5s DNAs) often exist as a separate gene family and are transcribed by polymerase III (Lewin, 1980).

Frogs of the genus *Xenopus* have a dual 5s RNA system (Ford and Brown, 1976). The 5s oocyte and the 5s somatic RNA gene families do not have identical sequences and their expression is under developmental control. Both types of 5s genes exist in clusters of tandemly repeated units consisting of a single 5s RNA gene surrounded by a non-coding spacer region unique to that family (Peterson, Doering and Brown, 1980).

The 5s somatic and oocyte genes begin to be expressed early in oogenesis. The bulk of the 5s RNA in the developing egg is produced by the oocyte genes, of which there are two variations. The major oocyte family (Xlo) is present in 20,000 copies per haploid genome and consists of an approximately 700 bp repeat unit which contains a single functional gene of 120 bases (see Figure 1). This family shows evidence of a duplication that must have occurred long ago. The duplication is in the form of a non-transcribed pseudogene. Each repeat unit contains one pseudogene located downstream of the Xlo functional gene. The pseudogene consists of the first 101 bases of the functional gene and is separated from the functional gene by a sequence of 73 bases. This latter sequence shows striking homology to the 73 bases directly upstream of the functional gene (Miller *et al.*, 1978). Another characteristic of the Xlo family is the presence of a 15 base pair sequence located upstream of the functional gene. This sequence is tandemly repeated anywhere from 6 to 16 times per repeat unit. The variation in the length of the Xlo family is due to variation in the number of times this oligonucleotide sequence is repeated (Fedoroff and Brown, 1978). The Xlo family is found on the long arm telomeres of most of the 18 chromosomes (Pardue *et al.*, 1973).

The second oocyte family, trace oocyte, or (Xlt), is present in 1300 copies per haploid genome. It has a repeat

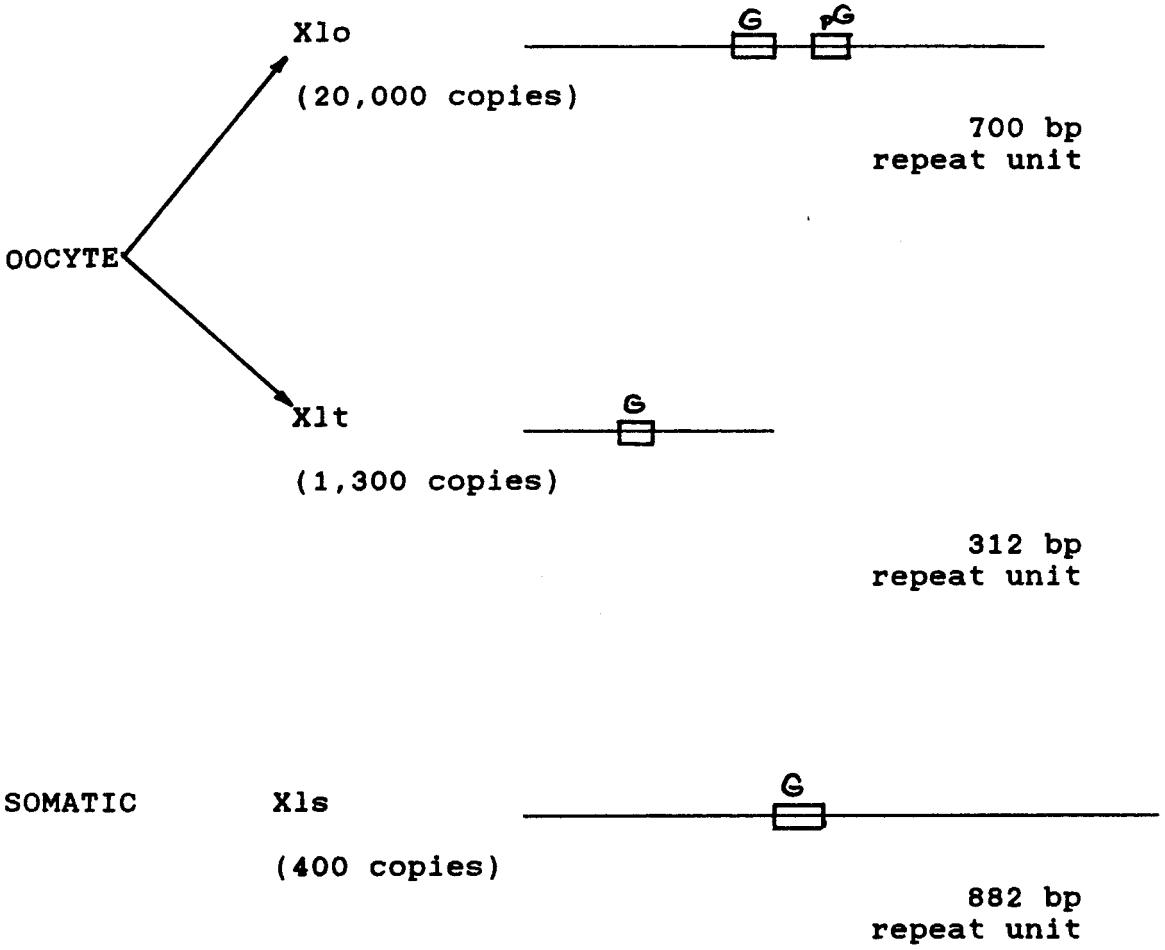
**Figure 1****Three 5s RNA Gene Families in *Xenopus laevis***

The gene coding region in all three gene families is 120 bases long. The nucleotide sequence of the somatic and oocyte genes differs by six bases. None of the spacer regions cross hybridize between families. Copy number and length of repeat unit for each family are as indicated.

G = Gene coding region  
pG = Pseudogene  
|||| = 15 base pair repeated sequence  
X1o = Major oocyte family  
X1t = Trace oocyte family  
X1s = Somatic family



XENOPUS 5s GENES



of 312 bases which contains a single functional gene (Brown *et al.*, 1977) (see Figure 1). The Xlt gene family does not contain a pseudogene nor any internal repeating sequences (Peterson, Doering and Brown, 1980). The Xlt family exists as a single cluster located on the long arm of chromosome number 13 (Harper *et al.*, 1983).

As development progresses and oogenesis is completed, the oocyte genes are turned off. The somatic genes remain active and are responsible for all additional 5s RNA needed by the organism after this time. The somatic 5s RNA genes, (Xls), are present in 400 copies per haploid genome. The repeat unit is 882 bases long and contains a single functional gene surrounded by its characteristic spacer (Peterson, Doering and Brown, 1980). This family is clustered on the long arm of chromosome number 9. (Harper *et al.* 1983).

The dual 5s RNA system in *Xenopus* may be an evolutionary response to the "supply and demand" situation present in the developing organism. The oocyte needs to produce an enormous number of ribosomes which will be stored until needed for protein synthesis later in development. There are two means by which the ribosomal RNAs can accumulate in such large quantities so quickly. One way is through amplification of the gene prior to expression. Extrachromosomal copies of the 18s and 28s RNA genes are synthesized by a rolling circle mechanism and the RNA product continues to

accumulate during the early stages of oogenesis (Rochaix, Bird and Bakken, 1974). These additional copies of the RNA genes are active only during oogenesis. When the egg is mature and cell division begins all RNA synthesis stops. The amplified (extrachromosomal) copies of 18s and 28s ribosomal DNA are no longer functional and are subsequently degraded (Brown and Dawid, 1968).

The other method of accumulating rRNA involves the early induction of a large number of genes which are normally present in the genome. The 5s RNA system turns on the X1o and X1t major families early in oogenesis so as to create and maintain a sufficiently large pool of 5s RNA. When the demand is met the genes are turned off (Ford, 1971). Later in development when RNA synthesis is resumed it is the 5s RNA somatic genes which are active, not the higher copy number oocyte genes.

The developmental control system for the 5s RNA genes has been well characterized. Initiation of transcription of the 5s gene is positively regulated by a transcription factor, IIIA, (TFIIIA), which binds to the gene's internal promoter site. The TFIIIA protein also binds to the 5s RNA after its synthesis to form a 7s storage particle. By binding to its own transcription factor, the 5s RNA can repress additional transcription from the 5s gene by a type of feedback inhibition (Pelham and Brown, 1980). Wormington, et al., (1981), found the somatic genes have a four-fold

greater binding affinity for the TFIIIA than the oocyte genes. The oocyte genes are expressed only when the concentration of TFIIIA is high. After oogenesis the TFIIIA content in the cell is drastically reduced (Wormington and Brown, 1983). When the egg has been fertilized and RNA synthesis resumes, only the somatic genes are active.

The difference in the genes' binding affinities for TFIIIA and the decrease in the concentration of the TFIIIA protein through development partially explains the shift in expression. When it was found that the 5s RNA somatic genes replicate earlier than the oocyte genes in somatic cells, the complete developmental regulation of the 5s RNA gene in *Xenopus* could be explained (Guinta et al., 1986). The concentration of TFIIIA is very high early in oogenesis, and both the somatic and the oocyte genes are active. As the RNA product accumulates and binds the TFIIIA it reduces the amount of free factor available and causes transcription from the oocyte genes to level off (Pelham and Brown, 1982). After oogenesis, the cellular TFIIIA concentration is dramatically decreased and the amount of oocyte-type 5s RNA transcribed drops rapidly (Wormington and Brown, 1983). When cell division begins the somatic-type genes replicate early in the S phase of the cell cycle. These genes have a four-fold higher affinity for TFIIIA than the oocyte genes (Wormington et al., 1981) and bind nearly all of the TFIIIA.

These genes can then assume a dynamic (active) chromatin configuration (Gottesfeld and Bloomer, 1982). By the time the oocyte genes replicate late in the S phase, the majority of the TFIIIA protein has already been sequestered by the somatic genes. What little free factor is present is not enough to reach the threshold factor:DNA ratio required for the oocyte genes to bind TFIIIA (Wormington et al., 1981). If the TFIIIA protein is not bound to the 5s gene, the chromatin assumes a static (non-transcribable) configuration (Gottesfeld and Bloomer, 1982).

#### 5S DNA Orphans in *Xenopus laevis*

Rosenthal and Doering (1983) have found evidence for additional copies of the *X. laevis* 5s genes located at dispersed loci. All of these sequences share homology with the the 5s gene coding region. Most contain sequences which are homologous to the Xlo spacer, but few were found to share homology with the spacer regions from the Xlt or Xls families. These genes, or orphans, appear to be dispersed away from the major family. The majority of orphans show significant population polymorphism with respect to organization; they are not present at the same loci in all individuals.

Recently orphans have been described for the tRNA genes in *X. laevis*. Rosenthal and Doering (1983) identified several loci containing some, but not all, of the tRNA genes

known to comprise a single repeat unit in the major tRNA family. Both the structures and the organizations of these dispersed genes are different from those of the major tRNA families. Like the orphans previously described (Childs *et al.*, 1981), the tRNA orphans exhibit significant population polymorphism in the genomic arrangement of these dispersed genes. Some of the dispersed tRNA genes appear to be present at the same loci in all individuals, but the majority of the orphans have polymorphic organizations (Rosenthal and Doering, 1983).

Orphans were originally discovered in the histone genes of the sea urchin, the ribosomal genes of yeast and the histone and ribosomal genes of *Drosophila*, by Childs *et al.* (1981). The sea urchin genome has approximately fifty histone gene orphans. The structures and organizations of the orphon loci are not conserved in the population. One of the orphans isolated from the sea urchin was found to contain the histone gene plus some spacer region from the major clusters. The orphon is surrounded by non-histone DNA which is moderately repetitive in the genome. The sequences flanking each end of the orphon are not homologous to each other, nor are they palindromes. The histone orphon has less than 2% sequence divergence from the major histone gene clusters (Childs *et al.*, 1981). The extensive population polymorphism in the arrangement of the orphans would tend to indicate, though not prove, that orphans are located in different

sites in the genomes of different individuals. This polymorphism, in addition to the similarity of the orphon sequence to the major cluster, suggests the orphons are dispersed to other genomic locations quite frequently.

I have found that the arrangement of 5s orphons in *Xenopus laevis* generally resembles the histone orphons in the sea urchin. There is an extremely high level of population polymorphism in the arrangement of the 5s orphons in *Xenopus*. This polymorphism exists for the 5s gene region as well as the Xlo, Xlt, and Xls spacer regions. The gene region of the orphons shows greater than 92% sequence homology with the major family, and there are only traces of sequence divergence in the dispersed Xlo spacer regions. I have constructed genomic libraries and screened them in an attempt to isolate a 5s DNA orphon. The orphons, once isolated can be used to determine the method by which the 5s element was dispersed.

## LITERATURE REVIEW

### Possible Mechanisms for Orphon Dispersal

Orphons may be dispersed by several types of genomic rearrangements. There are several mechanisms which have the potential to create the type of rearrangement seen with orphons. Some of the mechanisms suggested involve transposable elements, sequences utilizing RNA intermediates (retroposons or via cellular RNA) and unequal crossing over events.

#### Transposons:

Transposition, the moving of a DNA fragment from one chromosomal location to another by a viral or non-viral mobility element, has been identified and studied extensively in procaryotes. Transposition of fragments into new regions of the genome may generate chromosomal rearrangements including insertions, inversions (Calos and Miller, 1980), deletions, duplications (Farabough and Fink, 1980), and fusions (Calos and Miller, 1980). Most of these events are the result of the mechanisms involved in the insertion process.

The insertion sequence (IS) is the simplest element capable of transposition. These sequences occur naturally in



bacteria and their integration into the genome appears to be independent of DNA sequence homology (Calos and Miller, 1980). All insertion sequences have inverted terminal repeats of 20-40 base pairs. It is thought that the terminal repeats serve as the recognition site for the transposase protein which is essential to the mobility of the fragment (Grindley and Reed, 1985).

Transposons are related to, though more complex than, insertion sequences. Transposons differ from insertion sequences in that transposons are generally larger in size and often terminate in long inverted or direct repeats which appear to be IS or IS-like elements (Calos and Miller, 1980). Many transposons in procaryotes contain stretches of DNA which code for drug resistance genes or fertility factors. These transposons are flanked by IS elements normally found in the bacterial genome (for a review see Grindley and Reed, 1985).

The presence of transposable elements has been described in various eucaryotes. Each transposon carries a sequence which encodes a transposase protein. The transposon is capable of excision from, and reintegration into the genome, independent of other factors. Transposable elements in eucaryotes often contain long terminal direct or inverted repeats. Such transposons may be the result of two copies of an insertion sequence flanking a short stretch of DNA. Upon simultaneous excision and translocation of the insertion se-

quences the intervening DNA is taken along. These composite transposons have both a short inverted repeat (due to the terminal inverted repeat of the insertion sequence) and a longer terminal repeat which may be inverted or direct, depending upon the orientation of the two insertion sequences (Grindley and Reed, 1985). Detailed structural analysis of eucaryotic composite transposon-like sequences, such as those found in *Drosophila*, (Chia et al., 1985) indicates they share many characteristics common to procaryotic transposons. One example of an eucaryotic transposon was described by Spradling and Rubin (1982), who identified a class of transposable elements in *Drosophila*. The P-elements encode all necessary information for self-regulation of transposition. The P-elements may transpose at a high frequency and often induce high rates of mutation and frequent chromosomal aberrations in germ line cells (Spradling and Rubin, 1982).

Recently a family of transposable elements, Tx1, have been identified in *Xenopus laevis* (Garrett and Carroll, 1986). There are approximately 750 Tx1 elements per genome, apparently distributed on all chromosome arms. All of these elements have 19 base pair inverted terminal repeats and generate four base pair duplications at the target site. Detailed examination of the sequence of one of the isolated Tx1 elements revealed several open reading frames. It appears that the Tx1 element is capable of excising as a

discrete element and can regulate its own transposition. Comparisons of genomic DNA from several frogs indicates that the sequences of the Tx1 elements are highly conserved at some loci but are polymorphic at others. The polymorphism is due to a variable number of internal repeats within each Tx1 element. The overall polymorphism of these elements is less than 20%. Because of the sequence heterogeneity within the elements there is little or no interaction between the Tx1 elements at different loci. This results in the isolation of the elements at their respective loci (Garrett and Carroll, 1986). The 5s DNA orphans may represent sequences which were relocated by the imprecise excision of the Tx1 element carrying along a 5s DNA, followed by reintegration at a new locus, or the orphans may be transposons themselves.

#### Retroposons:

Retroposons are a variant of the transposable element. Like the transposon, the retroposons have discrete ends. There are long terminal repeats flanking a stretch of DNA which codes for a full length RNA. Retroposons are capable of mediating their own translocation and do so via an RNA intermediate. The retroposons contain regions which share sequence homology with some of the conserved regions found in the reverse transcriptase genes of retroviruses.

Sequences which utilize an RNA intermediate begin the process by transcribing a sequence of DNA into a molecule of RNA. This RNA molecule may undergo further processing in-

cluding the removal of introns and the addition of a poly-A tail before a reverse transcriptase makes a cDNA copy of the RNA. The cDNA is then reintegrated into the genome, either by insertion into a staggered chromosomal break or via a retrovirus intermediate (Wilde *et al.*, 1982). Since the retroposon involves an RNA intermediate it is a likely method for 5s orphion dispersal in *Xenopus*. The vast quantities of 5s RNA present during oogenesis makes this method of 5s genomic rearrangements possible.

There are several examples of retrovirus activity in eucaryotes. One example is the *mys* family, a group of retrovirus-like transposable elements found in some species of mice. The characteristic direct terminal repeats and target site duplications are present, as in all procaryotic transposons, but the *mys* family also shares some properties common to retroviruses. There is a stretch of uninterrupted polypurines on one strand which can serve as a primer for the reverse transcriptase enzyme. There is also an open reading frame within the sequence which carries several of the conserved regions found in other reverse transcriptases (Wichman *et al.*, 1985). This family of retrovirus-like transposable elements raises some questions concerning how closely these recombinational elements are related. It is possible that the *mys* elements may represent a new family of mammalian retrotransposons. Wichmann *et al.* (1985) suggest that the *mys* elements and retroviruses may have been derived

from early transposons, or conversely, the transposons may have evolved from the retroviruses. Regardless of their origin, these elements all have the potential to create extensive genomic rearrangements.

Another example of retrovirus activity in a eucaryote is seen in the high frequency of restriction site polymorphism between individuals within a single strain of *Drosophila melanogaster*. Some of this polymorphism has been attributed to three unique dispersed repeated gene families. The members of the copia, 412, and 297 gene families are located at widely dispersed sites (Potter et al., 1979) and appear to be capable of rearranging their location in the genome (Strobel et al., 1979). These families were originally identified as genes by the high cellular concentration of poly-A RNA corresponding to each gene family (Potter et al., 1979; Strobel et al., 1979). Although the sequences adjacent to the dispersed genes differ at each site, the coding sequences of the genes themselves are conserved within each family (Potter et al., 1979). The number and location of these elements have been found to change in cultured *Drosophila* cells at a higher rate than expected (Potter et al., 1979).

Strobel et al. (1979) have shown that the copia, 297 and 412 elements have different distributions within the genomes of individuals belonging to different strains. In addition, these elements exhibit polymorphic arrangement

between individuals in a single laboratory stock. Actively transposing elements may cause genomic rearrangements which may be detectable between single generations. Kikuchi *et al.* (1986) recently described *Drosophila* virus-like particles containing elements homologous to copia DNA. They suggest that the copia genes are derivatives of a copia retrovirus-like particle and are capable of retroviral replication (Kikuchi *et al.*, 1986).

Another example of a retroposon is the Ty element in the yeast *Saccharomyces cerevisiae*. This approximately 6 kilobase sequence has terminal direct repeats, called delta elements, of 338 base pairs (Boeke *et al.*, 1985). Although the Ty sequence is repeated about thirty times per genome, the delta regions are present in over one hundred copies, occurring either as a component of the Ty element or as independent units (Farabaugh and Fink, 1980). Excision of the Ty element appears to be through a recombinational event (Farabaugh and Fink, 1980). Following excision, a single delta element is occasionally left behind. It is possible that the approximately fifty delta elements scattered throughout the genome represent previous Ty excision sites (Farabaugh and Fink, 1980).

The Ty elements are believed to transpose via RNA intermediates. Boeke *et al.* (1985) have found that an increased frequency of Ty transposition occurs when the Ty element is linked to a sequence containing an induced pro-

moter. Further evidence for an RNA intermediate is provided by the fact that when a sequence known to contain an intron is inserted into the Ty element, the DNA sequence isolated after transposition has the intron precisely excised (Boeke *et al.*, 1985).

A surprising result of the work by Boeke *et al.* (1985) was the finding that the transposition event appeared to be mutagenic to the Ty element. After transposition the Ty sequences contain several point mutations. The altered nucleotide sequences are probably the result of the reverse transcriptase enzyme which tends to create mutations at high frequencies (Boeke *et al.*, 1985). The high mutation rate of reverse transcriptases may be largely responsible for the structural heterogeneity of the Ty elements, as well as other sequences which transpose via RNA intermediates.

#### Cellular RNA Intermediates:

There is another class of dispersed genes which utilize RNA intermediates, but are distinct from the retroposons. Unlike the retroposons, these sequences are not capable of promoting their own transposition. This class of dispersed sequences is the result of cellular RNA (mRNA) which has been transcribed into a cDNA molecule and subsequently reinserted into the genome. These sequences do not have discrete ends, so the lengths of the repeats are often different for individual members of the family. RNA processing is a common characteristic of these sequences. There are two

classes of these "unintentional transposons". One type includes some of the dispersed gene families, the other is comprised of pseudogenes.

Some gene families have pseudogenes which, by examination of the pseudogene's sequence, are obviously the result of cDNA reintegration. These pseudogenes are very closely related in sequence to the functional gene's coding region, but they are non-functional. Some common maladies of the pseudogenes include truncated genes, deletions, premature stop codons and the absence of promoter sequences or terminator sequences.

Often the pseudogenes are not located on the same chromosome as the major cluster from which they appear to have arisen. The human *B*-tubulin pseudogene (Wilde *et al.*, 1982) the human immunoglobulin pseudogene (Hollis *et al.*, 1982) and the mouse  $\alpha$ -globin pseudogene (Leder *et al.*, 1981) are examples of DNA sequences which appear to have been generated by a cellular RNA intermediate. All three sequences appear to have undergone some degree of RNA processing prior to being acted upon by reverse transcriptase and reinserted into the genome. The extent of the processing ranges from the precise excision of introns to the presence of a 3' poly-A tail in the inserted cDNA sequence (Wilde *et al.*, 1982; Hollis *et al.*, 1982).

Pseudogenes, unlike orphans or transposons, do not demonstrate population polymorphism with respect to organ-



ization. This would indicate that pseudogenes were formed millions of years ago and must have subsequently become stably integrated into the germ line (Lee *et al.*, 1983). The organization of orphans, for example the dispersed histone genes in the sea urchin, is clearly polymorphic within the population indicating recently relocated fragments in a dynamic genome.

The short repetitive DNA sequences abundant in animals, like the human dispersed, repetitive Alu family (Jagadeeswaran *et al.*, 1981) and the small nuclear RNA (snRNA) pseudogenes (VanArsdell *et al.*, 1981) are thought to have been generated via RNA intermediates. There is significant variation in the length of individual members of these dispersed gene families. The Alu family sequences in humans, present in approximately 300,000 dispersed copies, are flanked by direct repeats and are particularly suited to being acted upon by a reverse transcriptase. The internal stretches of poly-A followed by the terminal stretch of poly-U provides the necessary sequence homology for base pairing to produce hairpin loops in RNA. These loops may serve as a priming sequence for the reverse transcriptase (Jagadeeswaran *et al.*, 1981). The resulting cDNA may then be reinserted into the genome. VanArsdell *et al.* (1981) have identified several snRNA pseudogenes which are flanked by direct repeats. The direct repeats in the Alu family show striking homology to those found in the snRNAs. It has been

suggested that both the Alu dispersed family and the snRNA pseudogenes are generated by reintegration of a cDNA into a staggered chromosomal break (VanArsdell et al., 1981).

The terminal stretch of poly-U found at the 3' end of the Alu gene serves as the polymerase III terminator sequence. The 5s genes in *X. laevis*, also transcribed by polymerase III, have similar termination sequences. It is possible that the 5s RNAs are transcribed into cDNAs and reinserted into the genome, generating the 5s orphans.

#### Recombinational Events:

Another dispersal model postulates that orphans are the result of unequal crossing over within a tandemly repeated family due to a mis-matched recombinational event (Childs et al., 1981; McCutchan et al., 1982). This type of recombination leads to sequence homogeneity of the tandemly reiterated family (Smith, 1976). The out-of-register recombinational event may generate orphans by inducing a looped out region in the tandem family. These loops may be randomly excised and subsequently reintegrated into the genome at dispersed locations. The formation of orphans may result if a single repeat unit of a tandemly reiterated gene family is excised and subsequently inserted at a dispersed site in the genome.

### Restriction Site Polymorphism

Once the orphon has been dispersed to a new locus it may remain there and become a stable part of that locus. A rearrangement of this type occurring in a germ line cell would be passed on to progeny, and with time, spread through the population. This type of rearrangement would cause the genomic locations of the orphons to be conserved in the population. The organization of the 5s DNA orphons in *Xenopus laevis* is very polymorphic in the population. There are 5s gene-containing loci present in some individuals, but absent in others. This polymorphism can be the result of frequent genomic rearrangements (utilizing one of the mechanisms discussed previously), or an extremely high frequency of point mutations occurring in and around the locus containing the 5s orphon.

The Tc1 elements in *Caenorhabditis elegans* are an example where the polymorphic organization within a population can be attributed to the presence of an actively transposing element. The genome of the nematode, *C. elegans*, contains a family of 1.7 kilobase elements (Tc1) present as interspersed repetitive sequences. The number of Tc1 elements present per genome differs only slightly between strains. Emmons *et al.* (1983) identified one strain of *C. elegans* which had a ten-fold greater number of Tc1 elements than the other strains. These additional copies appeared to

have been dispersed to new genomic loci, generating new restriction fragments which differed from fragments in other strains by 1.7 kb. In the strain of *C. elegans* identified by Emmons et al. (1983), the Tc1 element is apparently capable of precise excision, transposition, and reintegration into the genome. These events must occur at a high frequency in order to result in the proliferation of the Tc1 elements seen in that strain. This will also induce the generation of increasingly polymorphic restriction fragments as more genomic loci are disrupted by the inserted Tc1 elements.

The polymorphism seen in *Xenopus* 5s orphans could be attributed to a high frequency of point mutations occurring near the dispersed 5s gene. Since orphans are not a part of the major families, they may be free of the correction mechanisms which prevent divergence and maintain homogeneity within the major cluster, i.e. gene conversions (Slightom et al., 1980; Baltimore, 1981) and out-of-register recombinations (Smith, 1976). Without these restraints against sequence divergence, point mutations can accumulate in and around a locus containing a dispersed 5s gene. The point mutations may generate new restriction recognition sites which will contribute to the polymorphism. In order to produce the extensive polymorphism seen in the *Xenopus* 5s orphans there must be an extremely high frequency of point mutations occurring near the orphon's locus.

## MATERIALS AND METHODS

### DNA Isolations

Total genomic DNA was isolated from erythrocytes of adult *Xenopus laevis* females. The animals were anesthetized at 4°C, then double pithed. About two milliliters of whole blood were obtained by cardiac puncture. The blood was immediately diluted to 10 mls with a solution containing 5 units of heparin per ml 0.5X SET (1X SET is 150 mM NaCl, 50 mM Tris, pH 7.9, 1 mM [ethylenedinitrilo]-tetraacetic acid disodium salt (EDTA)).

Cells were washed twice to remove clotting factors, and the red blood cells were pelleted in a clinical table-top centrifuge. The supernatant was decanted and 10 mls of fresh heparin-SET solution was used to completely resuspend the pellet before centrifuging again. The washed cells were resuspended in 40 mls of heparin-SET solution to which proteinase K and sodium dodecyl sulfate (SDS) were added to final concentrations of 0.1 mg/ml and 0.5%, respectively. The tube was tipped to mix well and incubated at 37°C for 16 hours.

In a glass centrifuge tube the lysed cells were extracted with an equal volume of phenol. Phases were separated by centrifugation at 6,000 r.p.m. for 10 minutes in a Sorvall

SS34 rotor. The aqueous phase was harvested and extracted once with an equal volume of phenol:chloroform (1:1), once with chloroform alone and twice with ethyl ether. Following ether extractions the residual ether was boiled off by placing the tube in a 68°C waterbath for about 10 minutes.

When all the ether had been removed the aqueous phase was transferred to a plastic tube. Spermine tetrahydrochloride was added to a final concentration of 1 mM and the DNA was precipitated on ice for 30 minutes. Using a pasteur pipette, the spermine solution was carefully removed without disturbing the DNA and replaced with an equal volume of spermine extraction buffer (0.3 M sodium acetate, 10 mM magnesium acetate, 75% ethanol). The tube was placed on ice for one hour and tipped every 15-20 minutes to aid in the removal of the spermine. The solution around the DNA was carefully removed and the DNA was spermine extracted twice more. After removing as much of the extraction buffer as possible, the DNA was resuspended in 20 mls of 1.33X SET.

When the DNA was completely in solution T1 ribonuclease and ribonuclease A were added to final concentrations of 33.33 units/ml and 53.33 ug/ml, respectively. The solution was incubated at 37°C for two hours. After incubation, an equal volume of distilled water was added along with SDS and proteinase K, to final concentrations of 0.1 mg/ml and 0.5%, respectively. This solution was incubated at 37°C for one hour.

The three organic extractions were repeated (as described above) and the final aqueous phase was ether extracted twice. The DNA was spermine precipitated (as described above) and extracted three times in spermine extraction buffer. The DNA pellet was washed in 70% cold ethanol, and as much of the ethanol was drained away as possible. The DNA was resuspended in 2.0 mls of 0.4X SET. When most of the DNA had gone into solution an equal volume of distilled water was added. When completely in solution, the DNA concentration and purity were determined by ultra-violet absorption spectroscopy.

#### Probe DNAs

Our lab has previously described plasmids containing a segment of *Xenopus borealis* oocyte 5s DNA (pBXbo1), a complete Xlo repeat unit from *Xenopus laevis*, (pXlo31), Xlo spacer specific sequence, (pXlo31-2), and the Xls- and Xlt-spacer-specific sequences (Rosenthal and Doering, 1983). Plasmid DNAs were purified and handled as described (Peterson, Doering and Brown, 1980).

#### Restriction Digests

All restriction enzymes were purchased from Bethesda Research Laboratories and digests were performed in the manufacturer's recommended buffers. Unless otherwise noted, reactions contained 10 ug genomic DNA digested with 40 units

of enzyme. Incubation at 37°C for 2.5 hours yielded complete digestion. Reactions were stopped by the addition of SDS and EDTA to final concentrations of 0.5% and 10 mM, respectively.

### DNA Transfer

The DNA fragments were size fractionated by electrophoresis on 0.9% agarose gels. The DNA was transferred by the alkaline method (Reed and Mann, 1985) onto Gene Screen Plus membranes (NEN).

### Probes

All probes were <sup>32</sup>P-labeled by nick translation (Rigby et al., 1977). Unless otherwise stated, the Gene Screen Plus membranes were pre-hybridized at 37°C for 5 hours in hybridization solution (50% formamide, 1 M NaCl, 50 mM Tris (pH 7.5), 1% SDS, 10 ug/ml denatured *E. coli* DNA). Denatured radioactively labeled probe was added and hybridized at 37°C for 16 hours. Unless otherwise noted, membranes were washed twice for 10 minutes at room temperature in 2X SSC (1X SSC is 150 mM NaCl, 15 mM sodium citrate, 0.1 mM EDTA), twice for 30 minutes at 65°C in 2X SSC plus 1% SDS and twice for 30 minutes at room temperature in 0.1X SSC. Membranes were air dried and exposed to Kodak XAR film with Cronex (DuPont) intensifying screens.



### Copy Number

Plasmid pXlo31, containing the 5s gene plus Xlo spacer was digested with Pst I. The amount of linear plasmid per lane was calculated to contain the same number of inserts present in 10 ug of genomic DNA if the gene occurred in 1, 2, 5, or 10 copies as previously described by Rosenthal and Doering (1983). After Southern transfer, the membrane was hybridized with the 5s gene-specific sequence, pBXbo1.

### Cloning of a 5s Orphon

#### Ligation and packaging of phage:

Genomic fragments for ligation into the lambda gt11 vector were generated by digesting total genomic DNA with Eco RI as described above. Reactions were stopped either by heat denaturing the enzyme at 65°C for 20 minutes or by the addition of EDTA and SDS, as usual. The sample was size-fractionated on a 10-40% linear sucrose gradient and centrifuged at 20°C for 24 hours at 26,000 r.p.m. in an SW41 rotor. Fractions containing fragments of the appropriate size range were pooled and purified (Maniatis et al., 1982).

Prior to ligation into the vector, the integrity and purity of the genomic fragments were analyzed by ligation of the fragments to themselves using conditions identical to those to be used with the vector. Ligation was considered complete when the lower molecular weight fragments shifted to molecular weights greater than 24,000 base pairs. Size

was determined by comparing the ligated sample with unligated material, visualized on an ethidium bromide stained agarose gel.

The genomic Eco RI fragments were ligated into a lambda gt11 vector (Promega Biotec), cut and dephosphorylated at a unique Eco RI site. Vector:insert mass ratios of 1 : 0.25 - 0.5 gave the highest ligation and packaging efficiencies. Ligation reactions contained 0.5 ug lambda gt11 DNA, 0.125 - 0.25 ug Eco RI genomic DNA fragments, 0.5 units of T4 DNA ligase and ligation buffer (1X is 30 mM Tris (pH 7.8), 10 mM MgCl<sub>2</sub>, 10 mM DTT, 0.4 mM ATP) in a final reaction volume of 5.0 ul. The ligation reaction was incubated at 16°C for two hours.

The ligated DNA was packaged in a lambda *in vitro* packaging kit (Packagene by Promega Biotec) under manufacturer's recommended conditions.

#### Plating phage:

The packaging efficiency of a library was determined by transfecting *E. coli* Y1090 host cells and plating on medium containing the indicator 5-bromo-4-chloro-3-indolyl B-D-galactoside (X-Gal). *E. coli* Y1090 host cells were grown in LB medium (10 g NaCl, 5 g Bacto-Yeast Extract, 10 g tryptone in 1 L distilled water, pH adjusted to 7.5) plus ampicillin (50 ug/ml) and 0.02% maltose at 37°C to an A<sub>560</sub> = 0.9. One hundred microliters of host cells were transfected with 100 ul titered phage particles (to give 200-300 plaque

forming units (pfu) per 100 X 15 mm Petri plate) in a sterile tube, mixed gently and incubated at room temperature for 20 minutes to allow phage particles to adsorb. Twenty microliters of isopropyl B-D-thiogalactopyranoside (IPTG) stock (20 mg/ml distilled water) and 50 ul X-Gal stock (20 mg/ml dimethylformamide) were added along with 2.5 ml of melted top agar (LB medium plus 10 mM  $MgCl_2$  and 0.8% agar). The mixture was poured onto a 100 X 15 mm Petri plate containing LB plus ampicillin base agar, and allowed to solidify. Plates were incubated at 37°C for 16 hours. The percentage of phage particles containing inserts was determined from the X-Gal screen. Those phage particles not containing inserts produced blue plaques while the recombinant phage produced colorless plaques.

#### Screening phage:

The library was screened using *E. coli* LE392 host cells grown in LB plus 0.02% maltose to  $A_{560}=0.7$ . The LE392 host cells were infected with phage as described above, except ampicillin was omitted from all media and phage were titered so as to plate at a near-confluent plaque density (approximately 18,000 pfu per 150 X 15 mm Petri plate).

When plaques had achieved a near-confluent lysis of the bacterial lawn the Petri plate was placed at 4°C to harden the top agar. Duplicate plaque lifts for *in situ* plaque hybridization were done for each plate (Benton and Davis, 1977). A dry, sterile (autoclaved 20 minutes on liquid

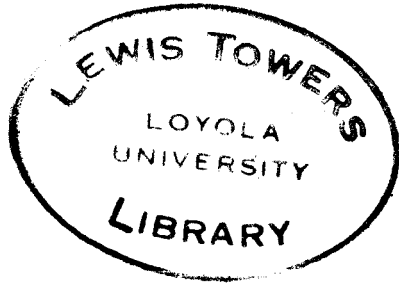
cycle) Colony/Plaque Screen Membrane (NEN) was aseptically placed on the top agar and allowed to gradually wet and adhere to the plate. After 2 - 3 minutes the membrane was keyed to the plate with a sterile needle probe and the membrane was then peeled off the plate. Cells and phage particles were lysed by placing membrane, cell side up, for two minutes on a piece of Whatman 3 MM paper saturated with 0.5 M NaOH, then to fresh 0.5 M NaOH for two minutes. The membrane was then transferred to 1 M Tris, pH 7.5 for two minutes and finally to fresh 1 M Tris, pH 7.5 for two minutes. Membranes were air dried and stored until hybridized.

Prior to pre-hybridization a positive control was added to each bag containing the membranes to be screened. The control consisted of a "plaque's-worth" of insert DNA ( $10^6$  phage particles per plaque, Maniatis et al., 1978) from the 5s gene-specific plasmid (XBo1 purified away from the pBR vector). The amount of XBo1 plasmid DNA needed (calculated as previously described by Rosenthal and Doering, 1983) was spotted on a Colony/Plaque Screen Membrane which was then treated with 0.5 M NaOH and 1 M Tris as described above.

Membranes were pre-hybridized at 37°C for at least three hours in hybridization solution (1 M NaCl, 10X Denhardt's solution (Denhardt, 1966) 1% SDS, 50 mM Tris (pH 7.5), 1 mM EDTA, 50% formamide, 10 ug/ml denatured *E. coli* DNA). Membranes were hybridized in fresh hybridization

solution containing radioactively labeled 5s gene-specific probe (pBXb01). Wash conditions were as described above, except the second set of washes was done at 60°C and the final set of washes was done in 0.5X SSC. Membranes were air dried and exposed to Kodak XAR film with Cronex intensifying screens for at least 24 hours.

Regions corresponding to a positive signal on the autoradiograms from both the original and the duplicate plaque lifts were rescreened. Phage were purified as previously described (Maniatis *et al.*, 1978) except phage particles were eluted in 600 ul plus 25 ul chloroform and the phage suspension was titered so that the plate contained at least 300 plaques.



## RESULTS

### 5S DNA Orphans

The 5s DNA orphans in *Xenopus laevis* are identified by "null-digestion" of total genomic DNA with restriction enzymes that do not cut within the major families. Southern blot analysis using a 5s gene-specific probe allows the 5s gene-containing fragments to be identified. The high molecular weight DNA at the top of the gel represents the intact major families (see Figure 2). The dispersed 5s genes can be resolved as lower molecular weight restriction fragments. The 5s gene-coding regions are highly conserved between two species of *Xenopus* (*X. laevis* and *X. borealis*), but the spacer regions do not cross hybridize (Korn and Brown, 1978). All fragments containing sequences homologous to the 5s RNA gene-coding region can be identified by hybridizing with the *X. borealis* 5s oocyte DNA probe, regardless of which of the three major families the dispersed genes were derived from.

Several restriction enzymes have been identified which yield null-digests of genomic DNA for the 5S RNA gene families. The number of 5s orphans and the size range of the

fragments produced varies, depending upon the enzyme and the individual used (see Figures 2-4). Generally, anywhere from ten to forty orphon-containing fragments are produced.

Comparisons of null-digested DNA from four unrelated individuals using a single restriction enzyme, Bam HI, shows that a unique orphon band pattern is produced in each individual using the 5s gene specific probe (Figure 2). Most of the bands display population polymorphism. These orphon-containing fragments are present in some individuals, but not in all. When an identical membrane is hybridized with a plasmid containing the 5s Xlo spacer-specific sequence, a band pattern very similar, but not identical to the 5s gene-specific pattern is produced (Figure 2).

When the fragments produced by hybridization of null-digested genomic DNA with the 5s gene-specific or Xlo spacer-specific sequences are counted, the data reflects the extensive population polymorphism which exists between individuals. There was only one 5s gene-containing fragment which was conserved in the population (see Table 1). This orphon was present in all individuals examined (see Figure 2). Nearly all fragments which hybridize to the gene-specific probe also hybridize to the Xlo spacer-specific sequence. The few notable exceptions are those which hybridize to only the 5s gene coding region or to only the Xlo spacer-specific sequence. Table 1 shows the total number of bands which have a polymorphic organization among the

Figure 2.

Comparison of Orphon Arrangement in Bam HI Null-Digested DNA from Four Individuals.

Total genomic DNA from four unrelated individuals (1-4) was digested with the enzyme Bam HI. Fragments were resolved on a 0.9% agarose gel and transferred to Gene Screen Plus Membranes. The membranes were hybridized with the 5s gene-specific or Xlo spacer-specific sequences. Some of the fragments containing only the 5s gene or only the Xlo spacer sequences are indicated by arrows.

- 5.4 kb = Gene with a conserved organization in the population
- 4.1 kb = Orphon containing only the 5s gene region (in individual 1)
- 3.3 kb = Orphon containing only the Xlo spacer region (in individual 4)

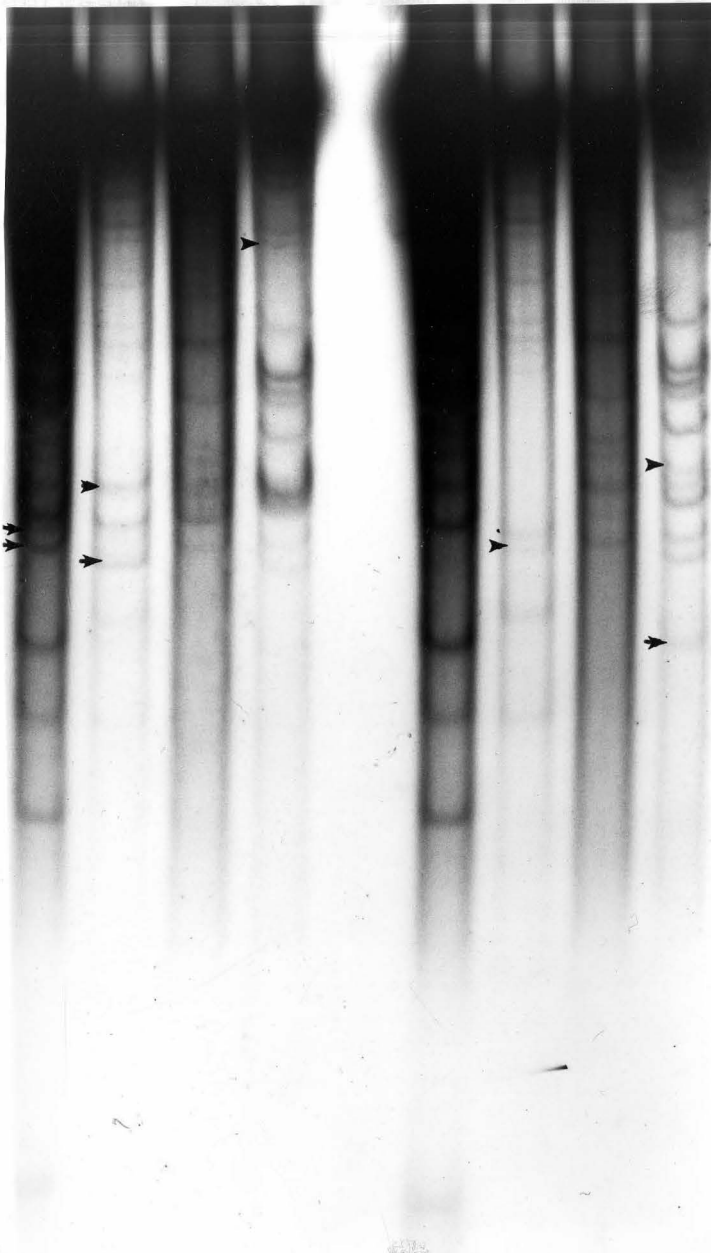


5s Gene

Xlo Spacer

1 2 3 4

1 2 3 4



Kbp

-7.0

-5.4

-4.1

-3.3

Table 1.

Polymorphic Arrangement of Null-Digestion Restriction Fragments Containing the 5s RNA Gene and the Xlo Spacer Sequences

Table 1 is compiled from data derived from Figures 2-4. The numbers represent the total number of bands produced by null digestion of genomic DNA from the individual indicated, using the enzyme indicated. The bands represent the restriction fragments containing regions homologous to the 5s gene and Xlo spacer sequences.

Total bands	=	Total number of fragments containing the 5s gene region.
Poly. bands	=	Number of bands which appear to have a polymorphic organization in the population.
Gene Alone	=	Total number of fragments which hybridize to the 5s gene specific probe, but not to the Xlo spacer probe.
Spacer Alone	=	Total number of fragments which hybridize to the Xlo spacer-specific probe, but not to the gene-specific probe.
% Polymorphic	=	Number of polymorphic bands / total number of bands.

Table 1.

<u>Enzyme</u>	<u>Individual</u>			
	1	2	3	4
<u>Bam HI</u>				
Total Bands	13	12	11	14
Poly. Bands	12	11	10	13
Gene Alone	2	3	0	1
Spacer Alone	1	2	0	1
% Polymorphic	92	92	91	93
<u>Eco RI</u>				
Total Bands	13	17	14	19
Poly. Bands	10	14	11	16
Gene Alone	0	2	1	1
Spacer Alone	1	2	1	0
% Polymorphic	77	82	78	84
<u>Bgl II</u>				
Total Bands		18		21
Poly. Bands		12		15
Gene Alone		2		0
Spacer Alone		0		0
% Polymorphic		67		71
<u>Kpn I</u>				
Total Bands		20		19
Poly. Bands		16		15
Gene Alone		3		2
Spacer Alone		0		0
% Polymorphic		80		79

individuals examined. The minimum estimation of polymorphic arrangement of orphons containing the 5s gene, generated by Bam HI digestion, is 91%.

Different orphon band patterns are seen when other enzymes are used. Total genomic DNA from the four individuals was digested with the enzyme Eco RI and hybridized with the 5s gene-specific or Xlo spacer-specific sequences (Figure 3). There are only three orphon fragments which are conserved in all four individuals (see Figure 3 and Table 1). The minimum estimation of polymorphic fragments containing the 5s gene, produced by Eco RI digestion, is 77% in those individuals examined.

Table 1 shows the number of polymorphic, gene alone, and Xlo spacer alone fragments for two individuals using the enzymes Bgl II and Kpn I (Figure 4). Bgl II null-digestion of total genomic DNA from two individuals produces a total of six lower molecular weight fragments which hybridize with the 5s gene-specific probe whose arrangements appear to be conserved in the population. At least 67% of the orphons produced by Bgl II digestion have polymorphic arrangements (Table 1).

When total genomic DNA is digested with the enzyme Kpn I, four orphon-containing fragments are produced which apparently have identical arrangements in the two individuals examined (see Figure 4 and Table 1). A minimum of 79%

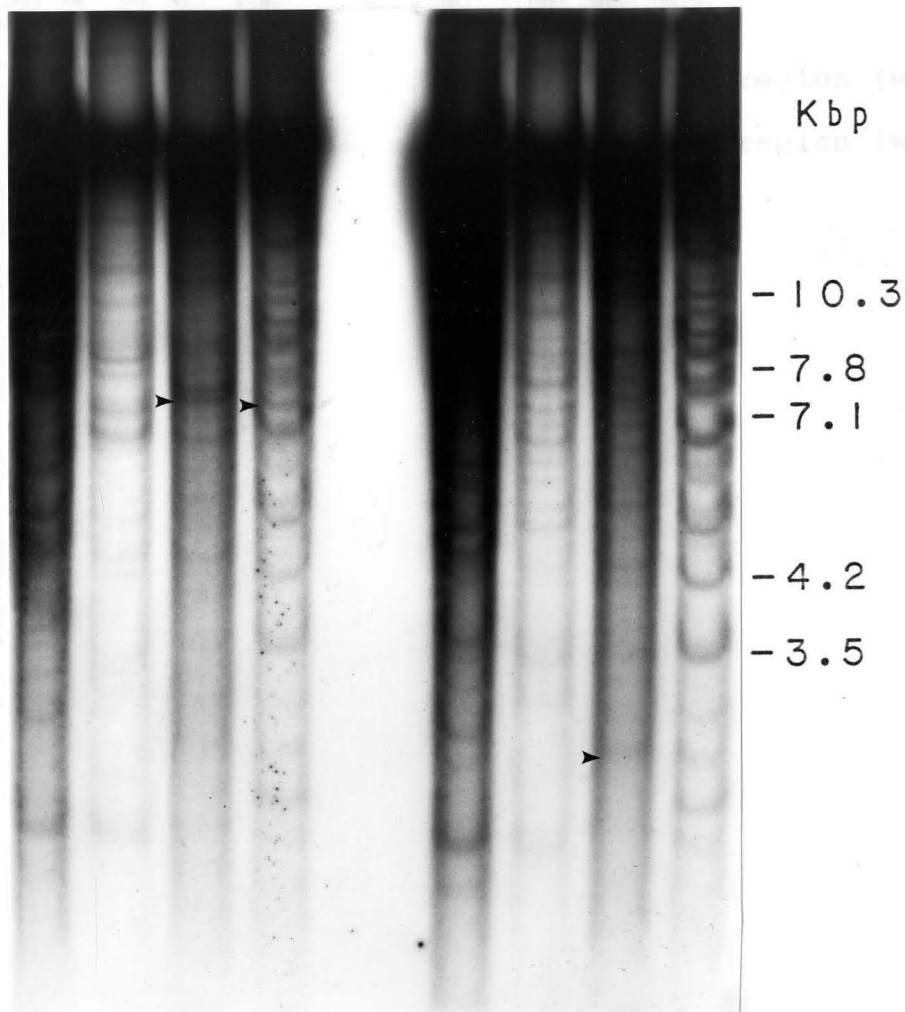
**Figure 3.****Comparison of Orphon Arrangement in Eco RI Null-Digested DNA from Four Individuals.**

Total genomic DNA from four unrelated individuals (1-4) was digested with the enzyme Eco RI. Fragments were resolved on a 0.9% agarose gel and transferred to Gene Screen Plus Membranes. The membranes were hybridized with the 5s gene-specific or Xlo spacer-specific sequences. Some of the fragments containing only the 5s gene or only the Xlo spacer sequences are identified by arrows.

- 10.3 kb = Gene with a conserved organization in the population
- 7.8 kb = Gene with a conserved organization in the population
- 7.1 kb = Orphon containing only the 5s gene region (in individual 4)

5s Gene Xlo Spacer

1 2 3 4 1 2 3 4



**Figure 4.**

**Comparison of Orphon Arrangement in Bgl II and Kpn I Null-Digested DNA from Two Individuals.**

Total genomic DNA from two unrelated individuals (1 and 2) was digested with the enzymes Bgl II or Kpn I. Fragments were resolved on a 0.9% agarose gel and transferred to Gene Screen Plus Membranes. The membranes were hybridized with the 5s gene-specific or Xlo spacer-specific sequences. Some of the fragments containing only the gene, or only the spacer sequences are indicated by arrows.

- 7.0 kb = Orphon containing only the 5s gene region (with Kpn I in individual 2)
- 5.8 kb = Orphon containing only the 5s gene region (with Kpn I in individual 1)

5s Gene

X10 Spacer

Bgl II Kpn I

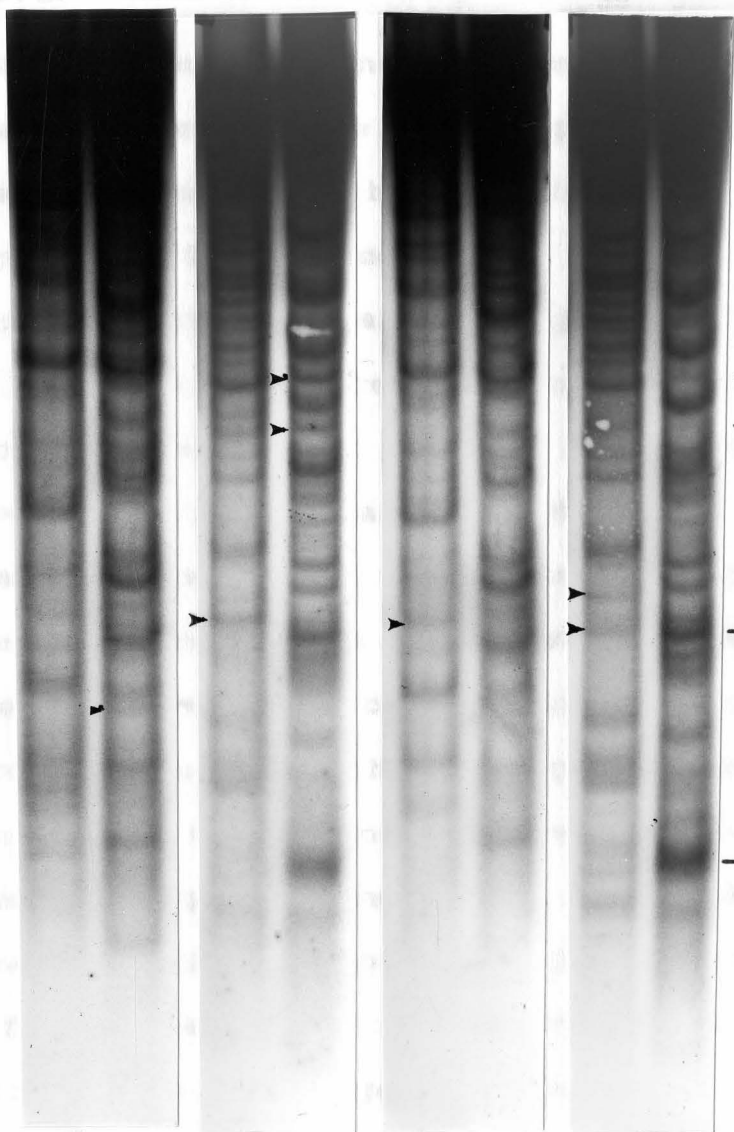
Bgl II Kpn I

1 2

1 2

1 2

1 2



Kbp

-7.0

-5.8

-3.4

-1.9



of the orphon-containing fragments produced in a Kpn I digest have polymorphic arrangements.

Through separate hybridizations using the 5s gene- and Xlo spacer-specific sequences, four types of orphons were identified. Two of the orphon types observed contained regions homologous to both the 5s gene coding region and the Xlo spacer sequences. The bands representing the gene plus Xlo spacer sequences may have a polymorphic organization in the population. These sequences are present in some individuals, but are not identically arranged in all. Alternatively the 5s gene and Xlo spacer sequences may appear non-polymorphic i.e. the fragment size and arrangement appear to be conserved in all individuals studied.

The other two orphon types share homology with only the 5s gene coding region. In these cases, the Xlo spacer probe does not produce a band of hybridization which corresponds to a band produced by the gene-specific probe. Fragments containing only the 5s gene sequences may have a polymorphic or a non-polymorphic arrangement in the population.

The possibility of point mutations occurring within the major family, generating new restriction sites and giving rise to the 5s null-digestion fragments, was considered previously (Rosenthal and Doering, 1983). Since all orphon-containing fragments are large enough to accommodate several 5s gene plus Xlo spacer repeat units, this type of event is possible, though it is not likely to have produced very many

of the 5s fragments (see Copy Number, below). The presence of fragments containing the 5s gene coding region without the Xlo spacer implies these genes are not linked to the major family and have been dispersed to a new genomic location. There are also a few fragments which contain the Xlo spacer, but not the 5s gene-coding region. These Xlo spacer regions must be unlinked to the major families and may have been dispersed to a new genomic location by a mechanism similar to those fragments containing only the 5s gene as well as the gene plus Xlo spacer fragments (see Discussion).

The possibility that some 5s orphans (specifically the fragments containing the 5s gene but no Xlo spacer) contained spacers from either of the other two major families was examined using plasmids containing the spacer region characteristic of the 5s Xls or Xlt gene families (see Materials and Methods). When total genomic DNA from four individuals was null-digested and hybridized to the Xlt spacer-specific probe the extensive population polymorphism is seen. Depending on the individual and the enzyme used, anywhere from one to twenty-three fragments contain regions homologous to the Xlt spacer-specific probe (see Figure 5 and Table 2). There is one fragment produced by Bgl II digestion which hybridizes with the Xlt probe and appears to be conserved in the population. All other bands containing regions homologous with the Xlt spacer specific probe have polymorphic arrangements.

**Figure 5.****Comparison of the Arrangements of Null-Digestion Restriction Fragments Containing the Xlt Spacer-Specific Region in Four Individuals**

Six micrograms of total genomic DNA from four unrelated individuals (1-4) was digested with the enzymes Bgl II, Eco RI or Kpn I, as indicated. Fragments were resolved on a 0.9% agarose gel and transferred to Gene Screen Plus Membranes. The membranes were hybridized with the Xlt spacer-specific sequence.

- 1 = Bgl II digest
- 2 = Eco RI digest
- 3 = Kpn I digest

23.7 kb = Spacer sequence with a conserved organization in the population (with the enzyme Bgl II)

1

2

3

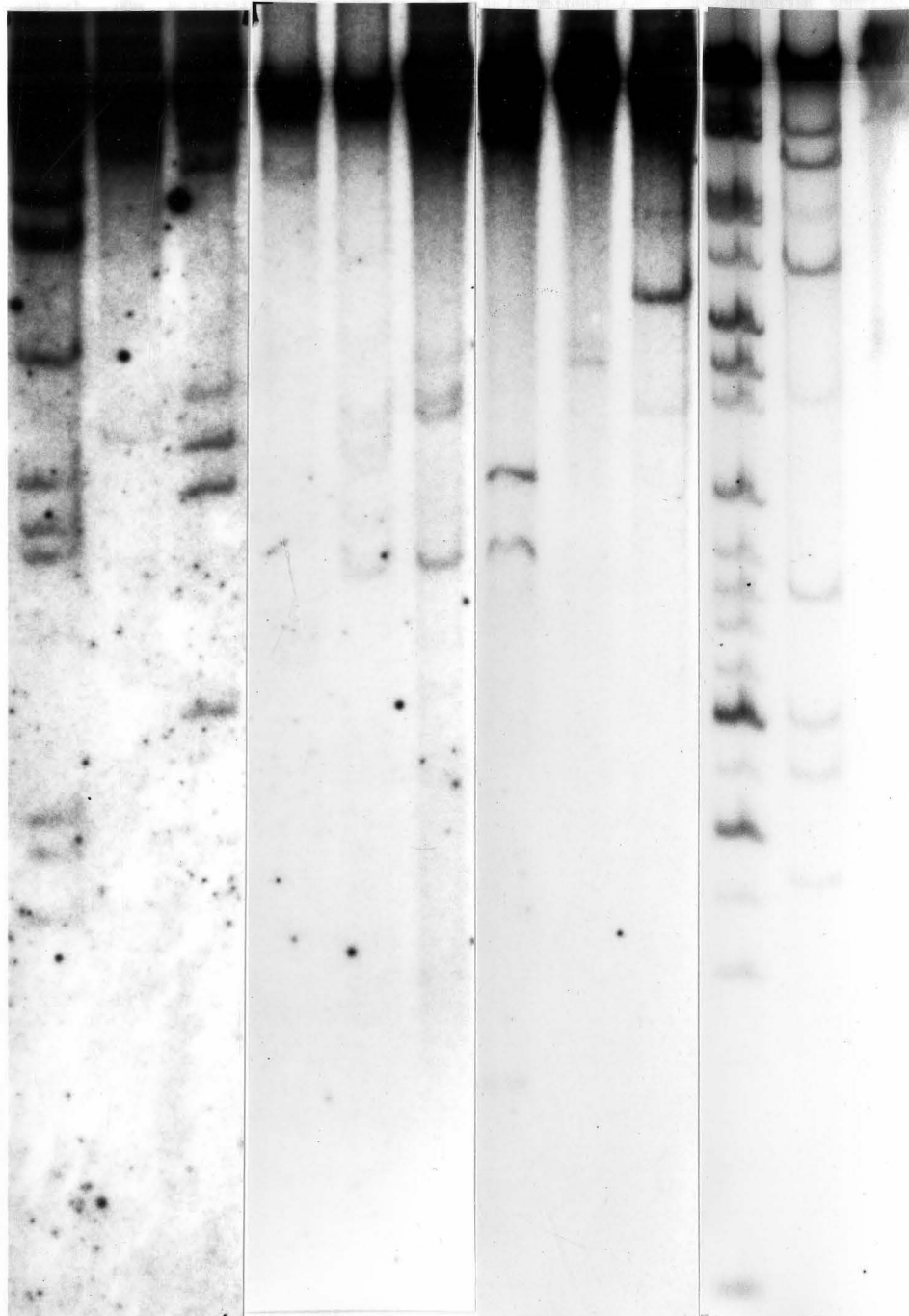
4

1 2 3

1 2 3

1 2 3

1 2 3



Kbp

-23.7

-8.9

-5.1

-2.8

-.68

Table 2.

**Polymorphic Arrangement of Null-Digestion Restriction Fragments Containing the 5s Xlt Spacer Sequence**

Table 2 is compiled from data derived from Figure 5. The numbers represent the total number of bands produced by null digestion of genomic DNA from the individual indicated, using the enzyme indicated. The bands represent restriction fragments containing regions homologous to the Xlt spacer-specific sequence.

Total bands = Total number of fragments containing the 5s Xlt spacer sequence.  
Poly. bands = Number of bands which appear to have a polymorphic organization in the population.  
% Polymorphic = Number of polymorphic bands / total number of bands.

Table 2.

<u>Enzyme</u>	<u>Individual</u>			
	1	2	3	4
<u>Bgl II</u>				
Total Bands	10	1	4	23
Poly. Bands	9	0	3	22
% Polymorphic	90	0	75	96
<u>Eco RI</u>				
Total Bands	3	6	2	9
Poly. Bands	3	6	2	9
% Polymorphic	100	100	100	100
<u>Kpn I</u>				
Total Bands	6	4	4	3
Poly. Bands	6	4	4	3
% Polymorphic	100	100	100	100

The extensive polymorphic organization of the 5s orphans is also seen when genomic DNA is digested and hybridized with the Xls spacer-specific probe. Figure 6 shows the band patterns generated when DNA from four individuals is hybridized with the Xls spacer-specific probe. There is only one fragment in a Bgl II digest which hybridizes with the Xls spacer probe and appears to have a conserved organization in the population. The enzyme Eco RI also produces one fragment which hybridizes with the Xls probe and is present in all individuals examined. All remaining bands have polymorphic arrangements (see Figure 6 and Table 3).

Through separate hybridizations using the 5s gene-specific and Xlo spacer-specific sequences, a few bands were identified as containing only the 5s gene region (see Figures 2-4). These orphans did not hybridize with the Xlo spacer specific sequence. When the Xlt and Xls spacer sequences were used as probes, this group of 5s gene-containing fragments did not hybridize with either of the spacer specific sequences (data not shown). These orphans contain only the 5s gene region and must be unlinked to the major family.

The hybridization patterns produced between different individuals (using the same probe and enzyme) are as diverse as the patterns generated by hybridizing with different probes (Figures 5 and 6). The polymorphism seen using the

**Figure 6.****Comparison of the Arrangements of Null-Digestion Restriction Fragments Containing the Xls Spacer-Specific Region in Four Individuals**

Six micrograms of total genomic DNA from four unrelated individuals (1-4) was digested with the enzymes Bgl II, Eco RI or Kpn I, as indicated. Fragments were resolved on a 0.9% agarose gel and transferred to Gene Screen Plus Membranes. The membranes were hybridized with the Xls spacer-specific sequence.

1 = Bgl II digest  
2 = Eco RI digest  
3 = Kpn I digest.

14.0 kb = Spacer sequence with a conserved organization in the population (with the enzyme Eco RI)  
2.8 kb = Spacer sequence with a conserved organization in the population (with the enzyme Bgl II)



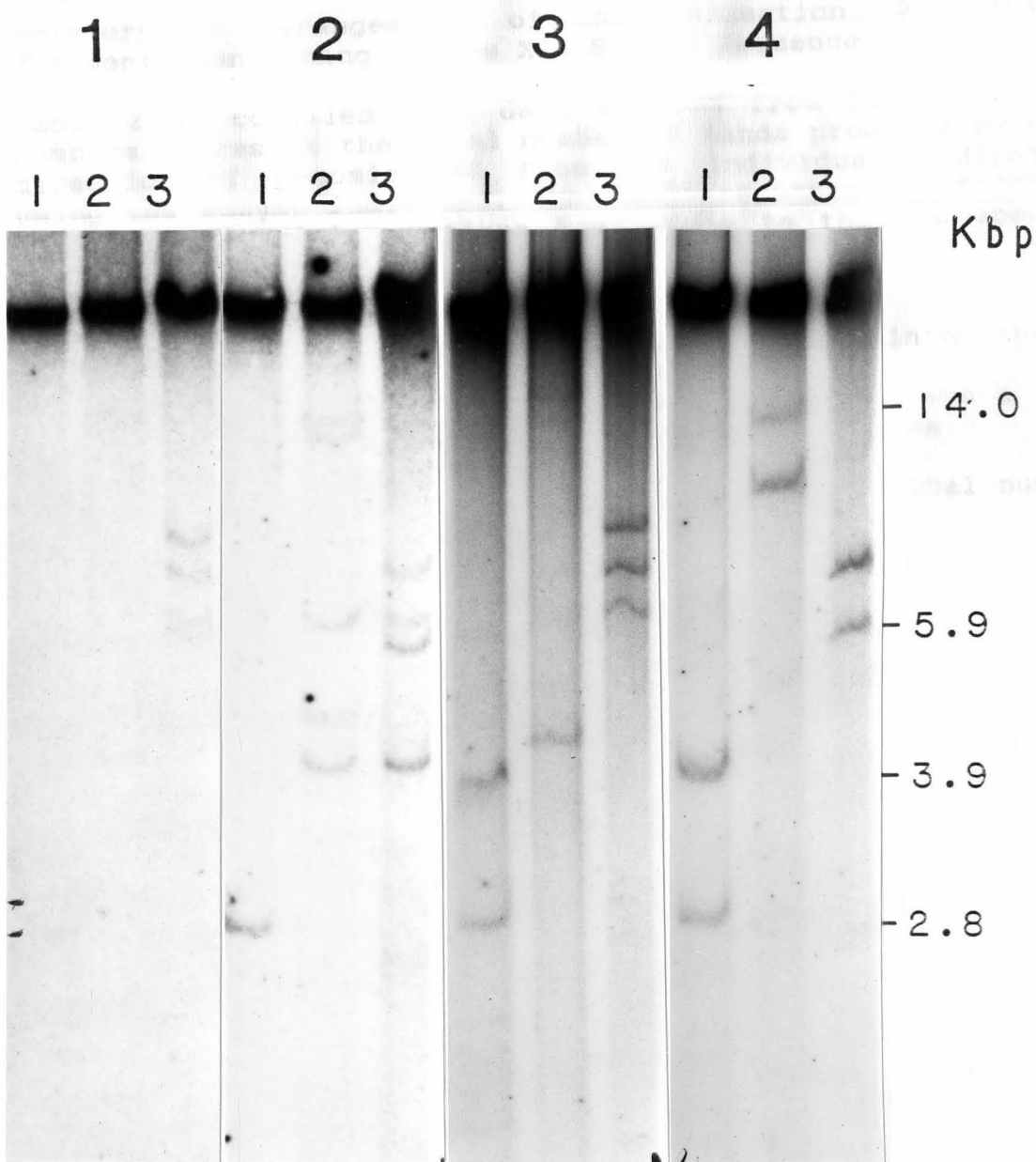


Table 3.

**Polymorphic Arrangement of Null-Digestion Restriction Fragments Containing the 5s Xls Spacer Sequence**

Table 3 is compiled from data derived from Figure 6. The numbers represent the total number of bands produced by null digestion of genomic DNA from the individual indicated, using the enzyme indicated. The bands represent restriction fragments containing regions homologous to the Xls spacer-specific sequences

- Total bands = Total number of fragments containing the 5s Xls spacer sequence.
- Poly. bands = Number of bands which appear to have a polymorphic organization in the population.
- % Polymorphic =  $\text{Number of polymorphic bands} / \text{total number of bands.}$

Table 3.

<u>Enzyme</u>	<u>Individual</u>			
	1	2	3	4
<u>Bgl II</u>				
Total Bands	2	1	2	2
Poly. Bands	1	0	1	1
% Polymorphic	50	0	50	50
<u>Eco RI</u>				
Total Bands	4	5	3	2
Poly. Bands	3	4	2	1
% Polymorphic	75	80	67	50
<u>Kpn I</u>				
Total Bands	3	4	3	2
Poly. Bands	3	4	3	2
% Polymorphic	100	100	100	100

Xlt and Xls spacer-specific probes is even more extensive than the polymorphism seen with the Xlo spacer-specific probe. Comparing the percentage of polymorphic bands produced with each of the three probes (see Tables 1-3) gives an indication of the polymorphic arrangement which exists among the fragments hybridizing with any given probe. Greater than 75% of the spacer-containing regions have polymorphic organizations (see Tables 1-3). However, the polymorphism is more extensive than the numbers show. Using the Xlo spacer-specific probe, several of the orphans are designated as having a polymorphic arrangement. There are one or two individuals which do not have fragments of that size which contain the Xlo spacer-specific sequence (see Figures 2-4). With the Xlt and Xls spacer-specific probes, the majority of the bands are polymorphic because none of the other individuals have fragments of that size homologous to the respective spacer regions (see Figures 5 and 6).

#### Copy Number

When null-digested DNA is hybridized with the 5s gene-specific probe, the uncut major family DNA, present in very high copy numbers, produces an unresolved, intensely hybridizing region at the top of the gel. The lower molecular weight orphan-containing fragments hybridize much less intensely than the major family, though most of these smaller

fragments are approximately of equal intensity to each other. To determine the copy number of the 5s gene coding regions within a given 5s fragment, the intensity of the hybridizing band is calibrated against a known amount of plasmid containing the 5s gene plus Xho spacer (see Materials and Methods). Figure 7 shows that most of the orphon bands appear to contain a single copy of the 5s gene. This is in agreement with the data previously described by Childs *et al.* (1981) in which the histone orphons were reported to be present as unique sequences.

The calibration curve in Figure 7 contains the plasmid vector, pBR322 (see Materials and Methods). When the membrane is hybridized with the gene-specific probe (also in a pBR322 vector) the intensity of the hybridization to the calibration curve is due in part to 5s gene : 5s gene interactions as well as to vector : vector interactions. When the Xho sequence is purified away from the pBR322 vector and used to hybridize the membrane, the intensity of the hybridization to the calibration curve decreases, as expected. The results still verify that most of the orphons contain single copies of the 5s gene (data not shown).

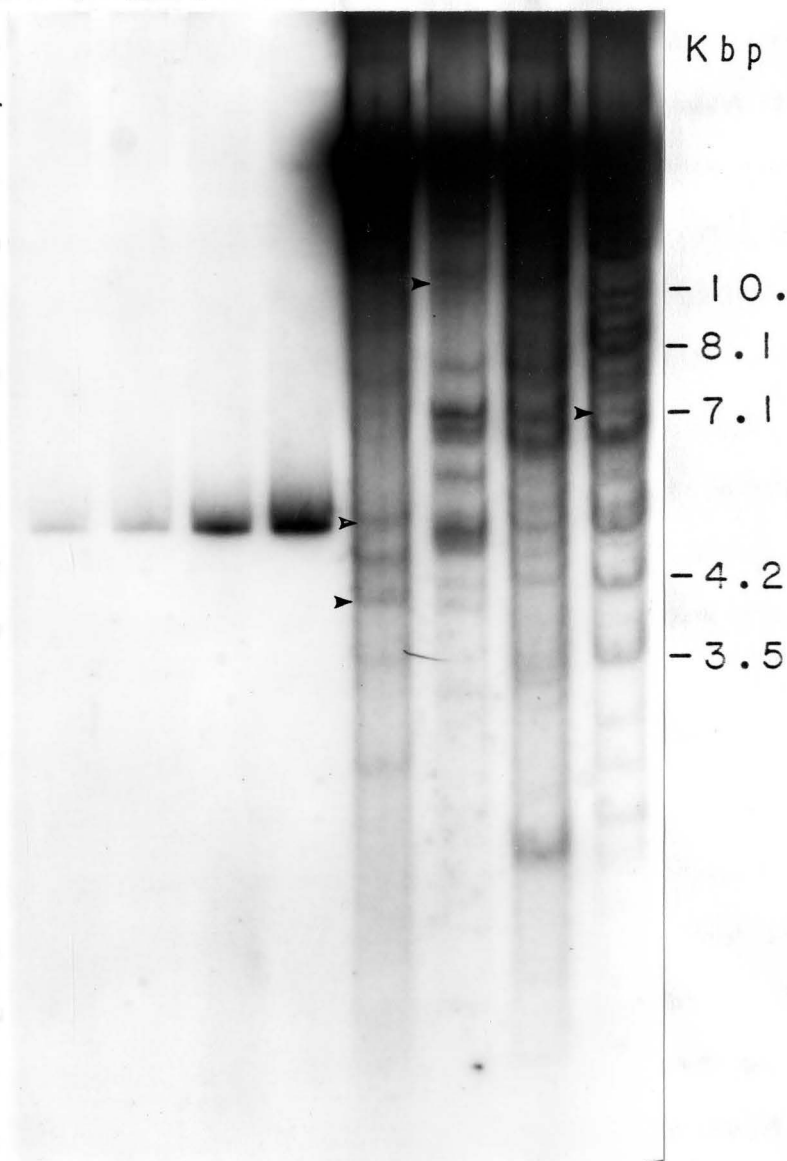
Since the 5s genes appear to be present as a single copy in any given 5s fragment it is highly unlikely that very many of these 5s fragments are the result of point mutations within a major family generating new restriction fragments (Rosenthal and Doering, 1983). For example, if the

**Figure 7.****Estimation of Copy Number of 5s Gene-Coding Regions in the 5S Orphans.**

Six micrograms of total genomic DNA from two unrelated individuals (1 and 2) was digested with the enzymes Bam HI and Eco RI as indicated. The calibration curve was generated from linear plasmids containing the 5s gene plus Xlo spacer sequences (see Materials and Methods). Fragments were resolved on a 0.9% agarose gel and transferred to Gene Screen Plus Membranes. The membranes were hybridized with the 5s gene-specific sequence. Some of the fragments containing only the 5s gene region are indicated by arrows.

Similar results were obtained when the membrane was hybridized with the 5s gene-specific sequence which had been purified away from the pBR322 vector (see Results).

Copy Number      Bam HI      Eco RI  
 1 2 5 10      1 2      1 2



7.1 kb fragment in Figure 7 had been generated by point mutations within the Xlo major family, it would be expected to contain about ten of the 5s RNA major family repeat units, assuming an average repeat size of 700 base pairs (Carroll and Brown, 1976). The number of repeat units which can be accommodated by a given fragment is directly proportional to the fragment size. The 3.5 kb fragment in Figure 7 would be expected to contain about half the number of repeat units present in the 7.1 kb fragment. However, the 3.5 kb fragment does not hybridize at half the intensity of the 7.1 kb fragment. If the dispersed 5s fragments were due to mutations in the major family then the larger fragments should contain more repeat units and thus hybridize more intensely. Generally, this proportionate increase in hybridization with larger fragments is not seen (see Figures 2-4).

### Orphon Homology

The determination of copy number as done here (see Materials and Methods) has its limitations. The intensity of hybridization is as dependant upon the extent of homology as upon copy number. Although the homology must be relatively strong to detect hybridization under the standard stringency washes, the conditions allow up to 23% sequence mismatch with the gene probe and up to 13% sequence mismatch with the Xlo spacer probe (McConaughy et al., 1969). Being able to



account for the homology component allows more weight to be placed upon the accuracy of the estimate of copy number.

Total genomic DNA from one individual was null digested and loaded adjacent to a digested plasmid containing the 5s gene and Xlo spacer regions (see Materials and Methods). After Southern transfer the membranes were hybridized and each gene-specific, spacer-specific probe pair was treated with a different set of stringency hybridization and wash conditions. The lowest stringency used the usual wash conditions while the most stringent conditions used can tolerate only about 8% sequence mismatch with the 5s gene probe and only trace amounts of mismatch with the Xlo spacer probe (McConaughy *et al.*, 1969). Results are shown in Figure 8.

Comparisons between the three stringencies using the gene-specific probe show a few of the bands decrease in intensity slightly, but none of the bands have diverged in sequence to the extent that homology can not be detected even at the extreme stringency of the wash conditions (see Figure 8). These results are in agreement with the data reported by Childs *et al.* (1981). The histone orphans in the sea urchin were estimated to have less than 2% sequence divergence with respect to the major family.

The majority of the bands which hybridize to the Xlo spacer-specific probe at the lower stringencies remain stable as hybrids even at the highest stringencies used.

**Figure 8.****Determination of Orphon Homology to the Major Family.**

Total genomic DNA from a single individual was digested with the enzyme Kpn I. A control plasmid was loaded in alternating lanes (see Materials and Methods). Fragments were resolved on a 0.9% agarose gel and transferred to Gene Screen Plus Membranes. The membranes were hybridized with the 5s gene-specific or Xlo spacer-specific sequence as indicated. Arrows indicate some fragments which were identified as containing only the 5s gene region.

Stringency conditions are as indicated.

- I. Membrane hybridized at 37°C, wash conditions are as usual.
- II. Membrane hybridized at 50°C, wash conditions are as usual except the second set of washes are done in 1% SDS plus 0.5X SSC.
- III. Membrane hybridized at 50°C, wash conditions are as usual except the second set of washes are done in 1% SDS plus 0.1X SSC.

# 5s gene Xlo spacer

I II III I II III

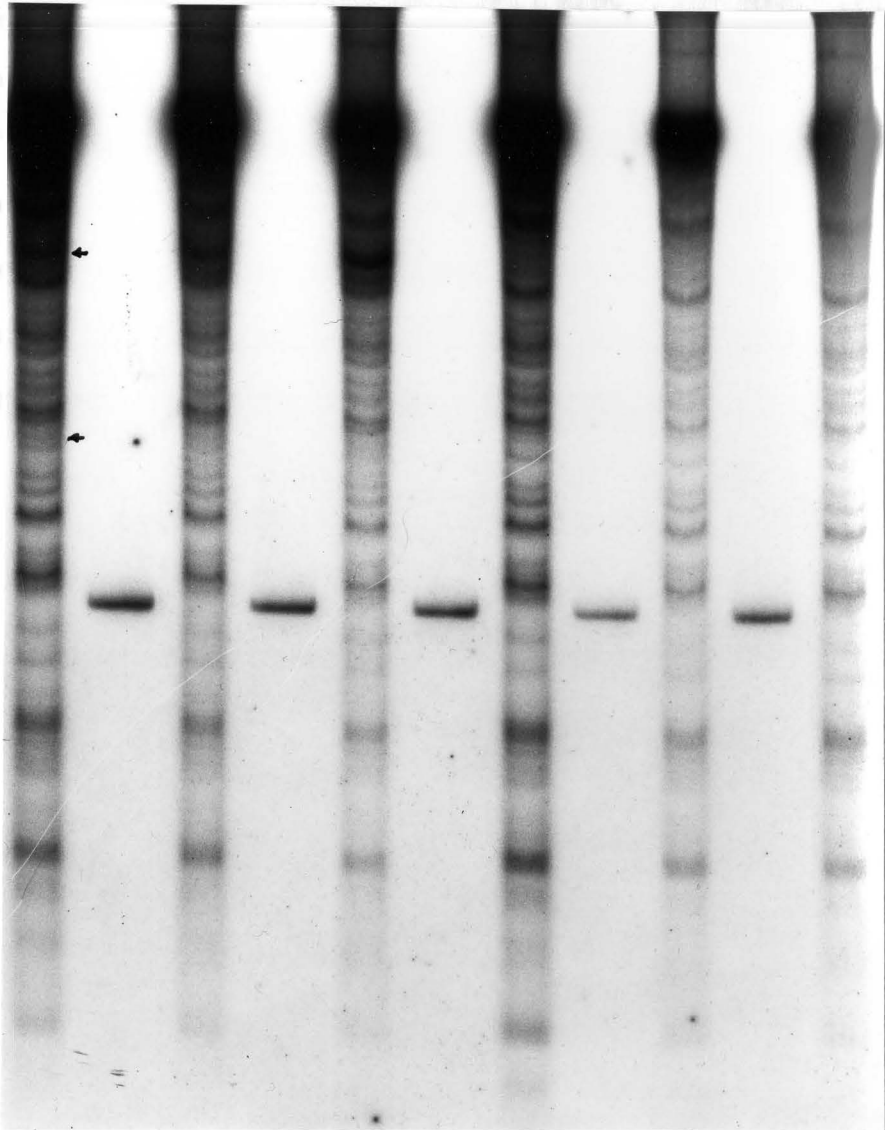
Kbp

13.3 -

6.2 -

4.3 -

2.5 -



These orphans must, at most, contain only traces of divergence from the major family sequence if they are to remain hybridized with the probe (McConaughy *et al.*, 1969).

There are two 5s orphans in the individual shown in Figure 8 which were identified as containing only the 5s gene coding region. Comparisons of the hybridization intensity of this particular type of orphan at the higher stringencies indicates that these dispersed gene regions retain extensive homology with the major family gene-coding region, in spite of any accumulation of point mutations occurring outside the gene region (see Discussion).

#### Cloning of a 5s DNA Orphan

The basis of the extensive polymorphism seen, as well as a possible mechanism of dispersal can be determined by cloning and subsequently characterizing an orphan. A phage vector was used to isolate an orphan because of the high rate and efficiency at which recombinants can be screened. Using a sucrose gradient, a size-selected pool of genomic (insert) DNA ranging from 4 to 8.5 kb was isolated from an Eco RI digest of total genomic DNA (see Materials and Methods). The size-selected fragments were ligated into a lambda gt11 vector (see Materials and Methods). The library of packaged phage particles transfected host cells at an efficiency of  $2.2 \times 10^5$  pfu/ug vector DNA with 90% of the

phage containing inserts. This library was constructed using a vector : insert mass ratio of 1 : 0.25. Total genomic DNA which was not size-selected for was used to generate shotgun libraries. These libraries gave the most efficient ligation and packaging when a vector : insert mass ratio of 1 : 0.50 was used. Transfection of host cells using the shotgun libraries was calculated to occur at an efficiency of  $7 \times 10^5$  pfu/ug vector DNA with 65% of the phage containing inserts.

Plaque lifts onto membranes for *in situ* plaque hybridization (see Materials and Methods) produced good strong signals with little distortion when a control hybridization was done using wild-type lambda DNA as the probe. Sequential lifts (up to four) from the same plate did not appear to decrease signal intensity or resolution significantly. Screening duplicate lifts from plates containing plaques at a near-confluent density increased the rate at which a library can be screened and eliminated unnecessary rescreening due to background hybridization.

Several libraries were screened using the gene-specific probe, but no orphans were found. All controls done indicate the libraries which were constructed do contain the 5s DNA orphon fragments. Positive controls used during the hybridization indicate the screening process described here is sensitive enough to detect the amount of insert (5s DNA) that would be present in a single plaque. Plaques giving positive signals can be rescreened and purified to

homogeneity. The recombinant lambda DNA, once isolated from the phage particles, is readily digested and should yield sufficient insert DNA to proceed with the characterization, once an orthon is isolated.

## DISCUSSION

I have identified four types of orphans in *Xenopus laevis* on the basis of structure and incidence of polymorphic organization. The arrangement of the orphans is quite different from that of the major families. The genes appear to be present in single copies and the majority of orphon-containing fragments have polymorphic organizations. The arrangement of the orphans differs between individuals, regardless of whether the orphans are identified by their homology to the 5s gene-coding region or to any of the three major family spacer-specific probes. This extensive polymorphism could be the result of point mutations occurring within the major family, though the data presented in this thesis argues strongly for genomic rearrangement as the basis of the polymorphism.

### Characteristics of the 5s DNA Orphans

Null-digestion of genomic DNA followed by hybridization to the 5s gene-specific probe produces anywhere from ten to forty 5s RNA orphon-containing fragments. The majority of the 5s null-digestion fragments contain both the 5s gene and the X10 spacer sequence. The total number of orphans which

hybridize with the 5s Xlo spacer specific sequence is very close to the total number of orphans which hybridize with the 5s gene-coding region. There are a few fragments which contain the 5s gene-coding region without the Xlo spacer region. Conversely, there are a few fragments which are homologous to the Xlo spacer, but do not cross-hybridize with the 5s gene-coding region. Most of the orphans appear to contain a single copy of the 5s gene region.

The possibility that the 5s orphans were generated from either the Xlt or Xls gene families was examined. There are anywhere from one to twenty-three fragments produced by null digestion which hybridize with the Xlt- and Xls spacer-specific probes (see Figures 5-6, and Tables 2-3). Some of these spacer-containing fragments also contain the 5s gene region, while others apparently do not (data not shown). None of the fragments identified as containing only the 5s gene region (and no Xlo spacer) hybridized to the spacer sequences from either the Xlt or Xls families. The fragments containing only the 5s gene region must not contain any appreciable amount of the adjacent 5s spacer sequence from any family (or the spacer regions must have undergone significant sequence divergence from that of the major family). These orphans probably contain only the 120 base gene-coding region. The fragments containing only the 5s gene region are unquestionably orphans which have been dispersed away from the major families. Fragments containing



only the spacer sequences from any of the three families must also represent sequences which are unlinked to the major families.

In order for a genomic rearrangement to spread through the population the event must occur in the germ line so as to be passed on to the offspring. Those genes which have higher levels of expression in the germ line have a more "open" chromatin configuration (Guinta et al., 1986; Lewin, 1980). This facilitates the binding of proteins to transcriptional control sites and relieves the topological restraints, which in turn aids the movement of polymerases through the gene coding regions. These "open" regions of chromatin may be more likely to be the site of a random insertion or excision of a transposon. Ford and Mathieson (1976) described an unusual expression pattern of an oocyte-type gene in *X. laevis*. The oocyte gene had undergone a transposition into a region of euchromatin, and remained active in a somatic cell line. Normally, in somatic cells, all of the oocyte-type genes have assumed the non-transcribed conformation of facultative heterochromatin (see Literature Review). The transposition (dispersal) of the 5s genes in *Xenopus* may be dependant upon the level and time of transcription, which in turn depends on the chromatin configuration (see Alu sequences, below).

All three major families are actively transcribed in the oocyte. The number of genes in each of the three major

families is not the same. There are more Xlo genes than Xlt or Xls genes (see Figure 1). The number of 5s orphon-containing fragments for the Xls, Xlt, and Xlo spacer sequences reflects this increase in the number of copies of genes in the major families. However, this increase in the number of orphons does not proportionally reflect the increase in copy number of the major families. The Xls, Xlt, and Xlo gene families are present in copy number ratios of approximately 1 : 3 : 50 (see Figure 1). The 5s spacer-containing orphons for Xls, Xlt, and Xlo are present in approximate ratios of 1 : 2.5 : 6 (Tables 1-3). If the number of orphons generated from a family is a function of the total number of genes present in that family, then there are a relatively low number of Xlo spacer-containing orphons.

Xenopus is not the only system where a non-transcribed spacer region, unlinked to the major family, has been identified. Arnheim *et al.*, (1980) identified a portion of the non-transcribed spacer sequence from the 18s and 28s RNA gene families located outside the major cluster in the mouse. These dispersed spacer sequences appear to be present in relatively high copy number and are scattered throughout the genome, occasionally flanking functional genes. The functions of these dispersed sequences are unknown, but a wide range of possible uses have been suggested. Some of the sequences may function as essential transcriptional control

sequences, or serve as origins of DNA replication. Alternatively, the dispersed sequences may be functionless DNA scattered by transposons or viral integrations.

There apparently has not been much sequence divergence in the 5s orphans. The orphans form stable hybrids with the 5s gene region or the Xlo spacer probes which have been derived from the oocyte major family. These hybrids remain stable under hybridization conditions requiring greater than 92% sequence homology. The homology seen between the orphans and the major family in both the gene region and the Xlo spacer sequence contrasts sharply with the polymorphic arrangements of the orphans. There must be a high frequency of genomic rearrangements and/or point mutations occurring in the regions surrounding the orphans to generate the polymorphism seen, yet the majority of the orphans maintain extensive sequence homology with the major family.

#### Point Mutations Versus Orphans

The 5s orphon-containing fragments could have been generated in either of two ways, through point mutations or dispersal away from the major clusters. Point mutations occurring within the major families could be the source of a few of the fragments, but the data presented here provides a strong argument for genomic rearrangements leading to the dispersal of the orphans. Evidence against point mutations

includes the number of copies of the 5s gene sequence per orphon, the intensity of hybridization of the orphons with respect to each other, and the unusually high frequency at which the point mutations must occur in order to produce the polymorphism seen in the 5s orphons.

One argument against the possibility of point mutations within the major family generating 5s orphons is that there are only a few fragments which appear to be present in slightly higher than single copies (see Figures 3 and 5). Some of these fragments may be the result of point mutations occurring in the major families, although the copy number is still too low to account for the number of repeat units which should be present in a fragment of that size, given the known repeat length of the Xlo major family (Carroll and Brown, 1976). Larger fragments would contain more repeat units, thus they should hybridize at a proportionately higher intensity. Since there is no overall gradual increase in intensity seen in fragments of higher molecular weight, the majority of the 5s orphon-containing fragments are not likely to be due to point mutations within a major family (see Figures 2-4).

Some of the 5s orphons could be generated by a point mutation occurring at the end of a tandem cluster. In this way it would be possible to create a fragment containing only a single repeat unit, however, this type of event could not account for those fragments containing only the 5s gene

region or any of the null-digestion fragments hybridizing with only the spacer sequences. In addition, there are too many 5s fragments to be accounted for by major cluster end fragments.

Finally, if the 5s fragments were generated by point mutations within the major cluster it would require that there be an extremely high mutation rate at the restriction site to see such extensive polymorphic organization. Yet, the 5s gene and Xlo spacer sequences show greater than 92% homology to the major family. If the 5s fragments were the result of point mutations within the major family it would require the gene and Xlo spacer sequences to be immune to the effects of the point mutations which generated the fragments.

#### Orphon Dispersal Mechanisms

Since the properties of the majority of the orphans themselves argue against their generation by point mutation, these orphans were probably dispersed by genomic rearrangements. There are several dispersal mechanisms which have the ability to create elements with characteristics similar to those seen in orphans.

#### Transposable elements:

One method of genomic rearrangement which may result in 5s DNA dispersal involves a viral or nonviral transposable

element. An unusual dispersed histone gene was isolated from the sea urchin by Liebermann et al., (1983) which was found to contain inverted terminal repeats, a characteristic of transposable elements. The 5s orphans in *Xenopus* may represent sequences which were relocated by the imprecise excision of a transposon (such as the Tx1 element in *Xenopus* identified by Garrett and Carroll, 1986). Upon excision the transposon may take along some adjacent 5s DNA and reintegrate at a new locus. Alternatively, the orphans may be transposons themselves. Examination of the region surrounding the gene and spacer sequences of the 5s DNA orphans may help to determine if transposable elements are involved in orphon dispersal (Liebermann et al., 1983; Calos and Miller, 1980).

#### RNA Intermediates:

A second mechanism which may be involved in orphon dispersal involves the reintegration of a cDNA molecule generated from RNA which has been acted upon by a reverse transcriptase enzyme.

There are an abundance of sequences belonging to the dispersed Alu family found interspersed throughout the human genome. These dispersed Alu elements appear to be the result of self-primed reverse transcription of the RNA transcript. This produces a cDNA which may subsequently reintegrate into the genome. These Alu elements are known to serve as active

transcription units for RNA polymerase III *in vitro* (Jagadeeswaran et al., 1981).

In order for the Alu sequences to transpose by self-primed reverse transcription they must contain an internal RNA polymerase III promoter sequence. Transcription would begin at the 5' end and proceed through a poly-A region found near the center. A poly-U polymerase III termination sequence at the 3' end would be transcribed and then used as a priming element. By a folding back upon itself, the poly-A region can hybridize with the stretch of poly-U residues and form the priming element (VanArsdell et al. 1981). Thus, the transposition of the Alu element is directly linked to its transcription. This type of self-priming transposition event is also believed to be the basis for the snRNA pseudogenes found scattered in the genome (VanArsdell et al., 1981).

This mechanism is particularly appealing for explaining the generation of some of the 5s orphans. Since the 5s genes are transcribed by polymerase III they contain the internal promoter and terminal poly-U sequences. The only component the 5s genes lack for a self-primed reverse transcription is a poly-A tail somewhere upstream of the termination sequence. The tRNA genes are also transcribed by polymerase III. It is possible that the tRNA orphans described by Rosenthal and Doering (1983) were also generated by this type of self-primed reverse transcription.

The normal 5s RNA transcript, if used as the template for the reverse transcriptase enzyme, can only be responsible for the generation of those orphans containing only the 5s gene coding region. Normally, only the 120 base gene region, and none of the adjacent spacer is transcribed into RNA.

The Xlo spacer sequence has an abundance of internal repeats. There are a series of a 15 base pair repetitive sequences present in six to sixteen copies upstream of the 5s gene in the Xlo major family (Fedoroff and Brown, 1978). The internal repetitiveness could allow the 5s Xlo sequence to be acted upon by reverse transcriptases, assuming there was an unusual 5s RNA which contained the spacer region as well as the gene sequence, and that there was enough homology to form a hairpin loop.

#### Unequal Crossovers:

The third mechanism which may be responsible for generating orphans involves unequal crossovers. This type of event may occur frequently in the RNA gene families in *Xenopus* (Brown and Sugimoto, 1973) and in yeast (Petes, 1980). It is the homologous and non-homologous crossing over events which maintain the homology of the tandem families and allow each family to evolve as a unit, even though the clusters may be present on different chromosomes. Ironically, it may be the homogenizing events of the major family which are responsible for dispersing the orphon



fragments and generating their polymorphic organization. Unequal crossover events occurring at either meiosis or mitosis may cause looped out regions which can be excised and subsequently reintegrated at a new genomic locus, creating an orphon (Childs *et al.*, 1980).

#### Generation of Four Types of Orphans

The structure of the 5s orphon itself may help to explain the mechanism by which it was generated and possibly give some indication as to when the rearrangement occurred. Genes which have been dispersed away from tandem families are presumably free of the correction mechanisms which tend to keep the major families homologous, i.e. gene conversions (Slightom *et al.*, 1980) and unequal crossovers (Smith, 1976). The dispersed gene coding regions, and especially the non-transcribed spacer sequences, should tend to accumulate point mutations over time. This would eventually result in sequence divergence and loss of sequence homology to the major family.

Assuming the mutation rate has been relatively constant in evolutionary time, the number of base changes between a transposed sequence and the major family is a reasonable reflection of the time that has elapsed since the rearrangement (Fedoroff and Brown, 1978). Using the sequence divergence of the orphans as a type of "evolutionary time clock",

I have proposed several ways in which each of the four types of orphans may have been generated.

One type of orphan contains the 5s gene and the Xlo spacer regions in a fragment whose organization appears to be conserved in the population. Those fragments which appeared to be conserved in smaller sample sizes may not actually be present in all individuals when a larger number of individuals is examined (see below). In the population examined, any fragments which were found to have conserved organizations may have been generated by a rearrangement which occurred long ago in the germ line. This genomic alteration was then transmitted to subsequent generations. With time it spread through the entire population and is now present in all individuals. The sequence homology to the major family is possibly maintained by mechanisms similar to those which keep the major cluster homologous.

A second type of orphan structure consists of the gene and Xlo spacer regions which exists as a polymorphic fragment in the population, being present in some individuals, but not in all. Fragments of this type were probably recently generated. The gene, and particularly the spacer region have not yet had the opportunity to diverge in sequence. These sequences are either novel fragments spontaneously generated in somatic tissues, or have become dispersed in the germ line very recently and have not yet had the opportunity to spread through the population.

Another orphon type consists of only the 5s gene-coding region which exists in a non-polymorphic arrangement in the population. These orphons were probably generated long ago in either of two ways. There may have been a precise excision and dispersal of the 120 bp gene region, with no adjacent spacer sequences. Alternatively, both the gene and the spacer regions may have been relocated together as a unit, but the spacer region has subsequently accumulated point mutations, to the extent that homology can no longer be detected, given the stringency of the hybridization conditions.

The fourth type of orphon displays homology to only the 5s gene-coding region and has a polymorphic organization in the population. This type of orphon may have been generated by the dispersal of a gene and X10 spacer sequence long ago. The sequence of the gene region was conserved while the spacer region has diverged so homology is not detected. Alternatively, the orphon may represent a precisely excised 5s gene-coding region which may have been very recently dispersed. In this situation, it is likely that this sequence is not fixed and may still be moving about the genome.

### Possible Role for Orphans in Development

Since the gene-coding regions are so highly conserved with respect to the major family, it is possible that their sequences are maintained for a specific purpose. After dispersal from the major clusters the orphon's gene(s) may have remained functional, but the coding region was altered slightly, and/or the spacer region diverged dramatically so as to fall under a different developmental control. If this new condition was advantageous to the organism, it may have been maintained and passed on to subsequent generations. This type of slight sequence divergence of the gene-coding region, and dramatic changes in the spacer region, leading to altered developmental control may be very similar to what has already occurred in the 5s gene family, leading to the three structurally and functionally distinct major families. It will be necessary to isolate and characterize an orphon to determine if the sequence is capable of being transcribed and if so, to determine when, during development, the orphon is expressed.

The sea urchin is another system which exhibits a correlation between developmental control of expression and genomic arrangement. The dispersed histone genes in sea urchin have been shown to exhibit an altered pattern of expression, leading to a set of developmentally specific genes. The sea urchin histone genes occur in two forms. The

early histone genes are found in tandem clusters and their organization within the cluster is highly conserved (Kedes and Maxson, 1981). The activity of these genes is reduced as development progresses. After the blastula stage, a new set of histone genes begins to be expressed (Kedes, 1979). The late histone genes differ considerably in both sequence and organization from the early genes (Childs *et al.*, 1979). The late histone proteins are coded for by a small set of dispersed genes. A comparison of the sequence of the late genes between species indicates there has been a great deal of divergence among the late genes as compared to the highly conserved early genes (Maxson *et al.*, 1983).

#### Polymorphic Arrangement of Orphons

The orphons in *Xenopus laevis* display extensive population polymorphism in their organization. The majority of the 5s orphons contain both the 5s gene and the X10 spacer sequence. When comparing four individuals, the lowest estimation of polymorphic fragments is 77% (see Figures 2-4 and Table 1), meaning less than 23% of the orphon-containing fragments produced by the null-digest are conserved in the population examined. Table 1 shows the total number of fragments which hybridize with the 5s gene-specific probe, producing a band with an intensity indicating a copy number equal to (or greater than) one. There are anywhere from

three to five additional bands per individual with each enzyme which produce bands of lesser intensity. These fragments were not included in the count of total fragments because the exact number of these weakly hybridizing fragments could not be determined with reasonable certainty. These weakly hybridizing bands may represent fragments containing gene regions which have diverged in sequence, therefore, they do not form very stable hybrids with the probe.

When comparisons were made using a different enzyme and a smaller number of individuals, the percentage of conserved fragments was larger. The decrease in polymorphism seen when comparisons were made using two individuals and the enzymes Kpn I and Bgl II (Figure 4 and Table 1) could be attributed, in part, to the restriction enzyme's recognition sequence. It is possible that a particular enzyme produces more conserved bands because the recognition sequence for that enzyme is present near many of the loci in which the orphon-containing sequences (preferentially) insert. Alternatively, the recognition sequence may occur somewhere within the transposable element, most likely near each of the terminal ends. In this situation, the orphon may transpose to a new location, but it carries the restriction recognition sites at conserved positions on the flanking sequences. This will cause some orphons to appear to have conserved arrangements in the population (since they are present within fragments

of identical length) even though they may be present at different loci. Tables 1-3, then, show the minimum estimation of polymorphic fragments. However, it is more likely that the lower level of polymorphism is due to the number of individuals studied. Increasing the number of individuals does decrease the total number of bands those individuals have in common (Table 1). Still, the population polymorphism must be quite high if less than 37% of the orphans have conserved arrangements when comparing only two individuals (Table 1).

The polymorphic arrangement of the orphans derived from the Xlt and the Xls families is much more extensive than that seen in the orphans derived from the Xlo family (see Results). The basis for this difference is unknown, but there are many factors which may play a role in regulating orphon dispersal. Differences in both structure and sequence of the spacer regions between the three families may have some effect upon the ability of the elements to become mobile. The mechanism used to disperse the orphans may be partly responsible for the differences in the levels of polymorphism. Several possible dispersal mechanisms were discussed (see Literature Review), any combination of which are possible. If different mechanisms are used by each family, then differences in the frequencies at which the rearrangements can occur, using a particular mechanism, may account for the number and organization seen in orphans from

each family. The time, during development, at which dispersal occurs may contribute to the polymorphism. Rearrangements must occur in the germ line in order to be stably transmitted to progeny, and potentially spread through the population. The stability of the RNA molecule may have an influence upon the number and type of 5s transcripts which are reinserted into the genome. Denis and Wegnez (1977) have shown that the oocyte-type 5s RNAs are more stable than the somatic-type 5s RNAs in oocytes. The 5s oocyte transcripts form stable storage particles. The somatic 5s genes are transcribed, but only about 5% of the somatic 5s RNAs are complexed with proteins to form storage particles. These somatic-type particles are stored for only a few days (Denis and Wegnez, 1977). The difference in RNA stability may play an important role in orphon dispersal, especially if rearrangement is via an RNA intermediate.

Once an orphon has been dispersed away from the major cluster it may no longer be susceptible to out-of-register recombination, but it is not free of other crossover events. Parahomologous recombination, as described by Smith (1976), occurs between two different loci and does not require extensive stretches of homology. Different loci which share short regions of homology may pair, exchange material, and thus contribute to the restriction site polymorphism at each locus.



Normal recombination, occurring within a single locus, may generate polymorphic restriction fragments if there is a high level of point mutations occurring around that locus. For example, a single copy of a dispersed sequence is present in the center of a 2 kb Bam HI fragment at a specific locus. This sequence may pair, during meiosis, with a single copy of a similar dispersed sequence at the same chromosomal locus (see Figure 9). A high frequency of point mutations occurring in the area may delete one of the restriction recognition sites adjacent to the orphon on each chromosome. A subsequent recombinational event in this region will contribute to any already existing polymorphism of these dispersed elements (see Figure 9). Even if the orphons were dispersed to only a few specific loci, this mechanism could account for much of the extensive restriction site polymorphism, assuming the rate of point mutations was high enough to generate the fragment length heterogeneity seen, and the crossovers occurred quite frequently. This would, however, require a mutation rate much higher than normally observed *in vivo*.

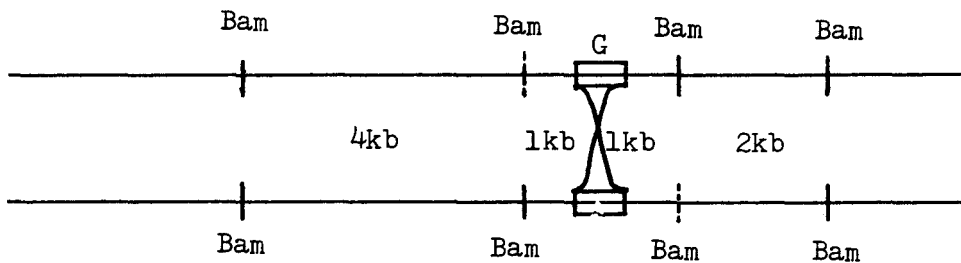
Several mechanisms have been suggested (any combination of which may be responsible) for the generation of orphons and for creating polymorphism once they have been dispersed. The basis of most of the mechanisms is the premise that orphons are single copy genes which have been dispersed away from a tandemly reiterated major family (Childs *et al.*,

**Figure 9.****Recombination Occurring Between Orphans at the Same Locus  
May Contribute to Polymorphism**

Two orphan fragments present at the same locus may undergo recombination. High levels of point mutations occurring in the flanking regions of the orphan may contribute to the already existing polymorphism.

Sizes of fragments containing orphans before point mutations are 2 kb. After point mutation (before recombination) there are two possible fragment sizes. Recombination within the gene region amplifies the polymorphism seen.

**G** = Gene region  
**Bam** = Bam HI recognition sites  
          (generated by high frequencies of point mutations)  
**⋮** = Bam HI sites eliminated by point mutations



**Restriction fragment  
size before  
recombination**

4 kb  
6 kb

**Possible restriction  
fragment size after  
recombination**

2 kb  
4 kb  
6 kb  
8 kb

1980). Orphans then, may represent a transition state in the conversion of a tandem gene family to a dispersed family. The presence of these dispersed genes may enable a cell to survive in the event of a catastrophic deletion of the tandem major family. The dispersed genes, provided they have not been rendered non-functional by sequence divergence, could offset an otherwise lethal deletion event (Childs *et al.*, 1980).

#### Characterization of a 5s Orphon

I have constructed genomic libraries containing the 5s DNA orphans. An orphon could be isolated by screening these libraries with the 5s gene-specific probe. Once an orphon has been purified, it will be necessary to characterize it, and possibly other orphans, to determine the basis of the polymorphism. Additional information may give some clues concerning the mechanism by which the orphans were dispersed.

The orphon-containing fragment will be digested with restriction endonucleases which cut several times within the fragment, generating smaller restriction fragments. These fragments can be separated by gel electrophoresis, and the DNA transferred to a membrane for hybridization. Those restriction fragments which contain the gene region can be identified by hybridizing this membrane with the gene

specific probe. Hybridizing with the Xlo, Xlt and Xls spacer sequences may indicate from which major family the orphon was generated. Those restriction fragments which do not hybridize to any of the major family probes contain the orphon spacer sequences. The orphon spacer sequences can be isolated and used as probes to hybridize back to total genomic DNA from which the orphon was originally cloned. Any fragment which hybridizes only once in total genomic DNA represents a unique orphon spacer sequence (see Figure 10). The unique orphon spacer sequence, then, is a region of DNA which is located outside the 5s orphon region (not in the gene-coding or major family spacer sequence) and occurs only once in the genome. This sequence may be a stretch of the DNA which mediates the transposition, or it may be a portion of the genome adjacent to the site at which the orphon inserted.

This unique sequence can be used as a probe to hybridize null-digested DNA from several individuals. The arrangement of this locus can be examined in other individuals by comparing the sizes of the bands produced. Since the dispersed element, the orphon, represents an extra sequence present at a specific locus, fragments containing the orphon will be of higher molecular weight than fragments at the same locus without the orphon. In this way it may be possible to indirectly determine the size of the orphon (Figure 10).

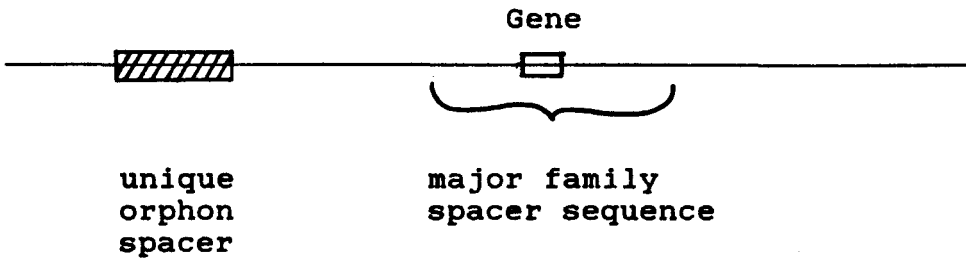
Figure 10.

Hypothetical Orphon Used to Determine the Basis of the Polymorphic Arrangement Between Individuals.

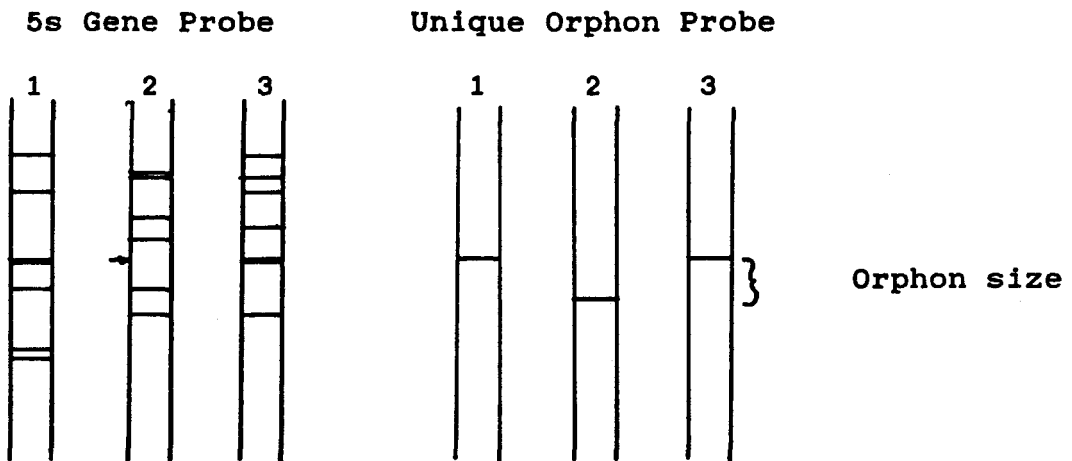
When a unique orphon spacer sequence has been identified it can be used as a probe to compare the arrangement of that locus in other individuals.

DNA samples from three different individuals are numbered 1-3. Hybridizing null-digested total genomic DNA with the 5s gene-specific probe produces a unique orphon band pattern for each individual. Individuals numbered 1 and 3 contain the 5s orphon of interest. Individual number 2 does not contain an orphon-containing fragment of that size (therefore, the orphon has a polyorphic arrangement in the population).

Since the cloned orphon has a polymorphic arrangement in the population, the orphon size may be determined indirectly. Using the unique orphon spacer sequence as a probe a single band will be produced which represents a unique locus present in each individual. Those loci which contain the dispersed sequences will be larger than those which do not contain the orphon. The size of the orphon is the difference between the two fragment sizes.



Hybridizing total genomic digests from different individuals with the 5s gene-specific probe allows for the determination of the polymorphic arrangement of the orphon. Hybridizing total genomic digests from different individuals with the unique orphon spacer sequence allows for the determination of the size of the dispersed sequence.



Comparing the arrangement of the orphon loci in several individuals will allow for the determination of whether the polymorphism is due to a high frequency of point mutations or to genomic rearrangement. If the size and organization of the fragments containing the unique orphon spacer is quite similar in the individuals examined, the polymorphism is probably due to sequence heterogeneity, i.e., point mutations. However, if the arrangement of the fragments containing the unique orphon spacer is very different between the individuals examined, then the polymorphism is probably due to extensive genomic rearrangements.

The extensive population polymorphism present in *Xenopus* 5s orphons has not been seen in other systems (Emmons et al., 1983; Kedes, 1979; Kedes and Maxson, 1981). The organization of the orphons varies so greatly from one individual to the next that it is possible that there are rearrangements occurring within each individual. Preliminary results by other investigators in our lab (F. Fokta, unpublished results) indicate that there are tissue-specific differences in orphon arrangement within a single individual. In addition, the offspring of a controlled mating appear to have orphon-containing fragments which can not be accounted for by either parent. If, indeed, the polymorphism is due to extensive genomic rearrangement it would be interesting to learn how the organism manages to survive with the continual disruption of genomic loci (and



therefore functional genes). Since the arrangement of the orphans is so polymorphic in the population, the genome of *Xenopus* must be extremely elastic to tolerate alterations within many different loci in different individuals.

## REFERENCES

- Arnheim, N., Seperack, P., Banerji, J., Lang, R.B., Miesfeld, R., and Marcu, K.B. (1980). Mouse rDNA nontranscribed spacer sequences are found flanking immunoglobulin C<sub>H</sub> genes and elsewhere throughout the genome. *Cell* 22, 179-185.
- Baltimore, D. (1981). Gene conversion: some implications for immunoglobulin genes. *Cell* 24, 592-594.
- Benton, W.D., and Davis, R.W. (1977). Screening lambda gt recombinant clones by hybridization to single plaques *in situ*. *Science* 196, 180-182.
- Boeke, J.D., Garfinkel, D.J., Styles, C.A., and Fink, G.R. (1985). Ty Elements transpose through an RNA intermediate. *Cell* 40, 491-500.
- Brown, D.D., Carroll, D., and Brown, R.D. (1977). The isolation and characterization of a second oocyte 5s DNA from *Xenopus laevis*. *Cell* 12, 1045-1056.
- Brown, D.D., and Dawid, I.B. (1968). Specific gene amplification in oocytes. *Science* 160, 272-280.
- Brown, D.D. and Sugimoto, K. (1973). 5s DNAs of *Xenopus laevis* and *Xenopus mulleri*: Evolution of a gene family. *J.Mol.Biol.* 78, 397-415.
- Calos, M.P., and Miller, J.H. (1980). Transposable elements. *Cell* 20, 579-595.
- Carroll, D., and Brown, D.D. (1976). Adjacent repeating units of *Xenopus laevis* 5s DNA can be heterogeneous in length. *Cell* 7, 477-486.
- Chia, W., McGill, S., Karp, R., Gubb, D., and Ashburner, M. (1985). Spontaneous excision of a large composite transposable element of *Drosophila melanogaster*. *Nature* 316, 81-83.
- Childs, G., Maxson, R., Cohn, R.H., and Kedes, L. (1981). Orphans: Dispersed genetic elements derived from tandem repetitive genes of eucaryotes. *Cell* 23, 651-663.
- Childs, G., Maxson, R., and Kedes, L.H. (1979). Histone gene expression during sea urchin embryogenesis: Isolation and characterization of early and late messenger RNAs of *Strongylocentrotus purpuratus* by gene specific hybridization and template activity. *Dev.Biol.* 73, 153-173.
- Denhardt, D.T. (1966). A membrane-filter technique for the detection of complementary DNA. *Biochem.Biophys.Res. Commun.* 23, 641-646.

- Denis, D., and Wegnez, M. (1977). Biochemical research on oogenesis. Oocytes of *Xenopus laevis* synthesize, but do not accumulate 5s RNA of somatic type. *Dev.Biol.* 58, 212-217.
- Emmons, S.W., Yesner, L., Ruan, K., and Katzenberg, D. (1983). Evidence for a transposon in *Caenorhabditis elegans*. *Cell* 32, 55-65.
- Farabaugh, P.J., and Fink, G.R. (1980). Insertion of the eucaryotic transposable element Ty1 creates a 5-base pair duplication. *Nature* 286, 352-356.
- Fedoroff, N.V., and Brown, D.D. (1978). The nucleotide sequence of oocyte 5s DNA in *Xenopus laevis*. I. The AT-rich spacer. *Cell* 13, 710-716.
- Ford, P.J. (1971). Non-coordinated accumulation and synthesis of 5s ribonucleic acid by ovaries of *Xenopus laevis*. *Nature* 233, 561-564.
- Ford, P.J., and Brown, R.D. (1976). Sequences of 5s Ribosomal RNA from *Xenopus mulleri* and the evolution of 5s gene-coding sequences. *Cell* 8, 485-493.
- Ford, P.J., and Mathieson, T. (1976). Control of 5s RNA synthesis in *Xenopus laevis*. *Nature* 261, 433-435.
- Garrett, J.E., and Carroll, D. (1986). Tx1: A transposable element from *Xenopus laevis* with some unusual properties. *Mol.and Cell.Biol.* 6, 933-941.
- Gottesfeld, J., and Bloomer, L.S. (1982). Assembly of transcriptionally active 5s RNA gene chromatin *in vitro*. *Cell* 28, 781-791.
- Grindley, N.D.F., and Reed, R.R. (1985). Transpositional recombination in procaryotes. *Ann.Rev.Biochem.* 54, 863-896.
- Guinta, D.R., Tso, J.Y., Narayanswami, S., Hamkalo, B.A., and Korn, L.J. (1986). Early replication and expression of oocyte-type 5s RNA genes in a *Xenopus* somatic cell line carrying a translocation. *Proc.Natl.Acad.Sci.* 83, 5150-5154.
- Harper, M.E., Price, J., and Korn, J. (1983). Chromosomal mapping of *Xenopus* 5s genes: somatic-type versus oocyte-type. *Nuc.Acids Res.* 11, 2313-2323.
- Hollis, G.F., Hieter, P.A., McBride, O.W., Swan, D., and Leder, P. (1982). Processed genes: A dispersed human immunoglobulin gene bearing evidence of RNA-type processing. *Nature* 296, 321-325.
- Jagadeeswaran, P., Forget, B.G., and Weissman, S.M. (1981). Short interspersed repetitive DNA elements in eucaryotes: Transposable DNA elements generated by reverse transcription of RNA pol III transcripts? *Cell* 26, 141-142.
- Kedes, L.H. (1979). Histone genes and histone messengers. *Ann.Rev.Biochem* 48, 837-870.
- Kedes, L., and Maxson, R. (1981). Histone gene organization: Paradigm lost. *Nature* 294, 11-12.

- Kikuchi, Y., Yumiko, A., and Shiba, T. (1986). Unusual priming mechanism of RNA-directed DNA synthesis in copia retrovirus-like particles of *Drosophila*. *Nature* 323, 824-826.
- Korn, L.J., and Brown, D.D. (1978). Nucleotide sequence of *Xenopus borealis* oocyte 5s DNA: comparison of sequences that flank several related eucaryotic genes. *Cell* 15, 1145-1156.
- Leder, A., Swan, D., Ruddle, F., D'Eustachio, P., and Leder, P. (1981). Dispersion of  $\alpha$ -like globin genes of the mouse to three different chromosomes. *Nature* 293, 196-200.
- Lee, M.G., Lewis, S.A., Wilde, C.D., and Cowan, N.J. (1983). Evolutionary history of a multigene family: An Expressed human  $\beta$ -tubulin gene and three processed pseudogenes. *Cell* 33, 477-487.
- Lewin, B. (1980). Gene Expression, Volume 2 Eucaryotic Chromosomes. John Wiley and Sons, Inc., New York.
- Liebermann, D., Hoffman-Liebermann, B., Weinthal, J., Childs, G., Maxson, R., Mauron, A., Cohen, S.N., and Kedes, L. (1983). An unusual transposon with long terminal inverted repeats in the sea urchin *Strongylocentrotus purpuratus*. *Nature* 306, 342-347.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory, New York.
- Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K., and Efstratiadis, A. (1978). The isolation of structural genes from libraries of eucaryotic DNA. *Cell* 15, 687-701.
- Maxson, R., Cohen, R., Kedes, L., and Mohun, T. (1983). Expression and organization of histone genes. *Ann.Rev.Genet.* 17, 239-277.
- McConaughy, B.L., Laird, C.D., and McCarthy, B.J. (1969). Nucleic acid reassociation in formamide. *Biochem.* 8, 3289-3295.
- McCutchan, T., Hsu, H., Thayer, R.E., and Singer, M.F. (1982). Organization of African green monkey DNA at junctions between  $\alpha$ -satellite and other DNA sequences. *J.Mol.Biol.* 157, 195-211.
- Miller, J.R., Cartwright, E.M., Brownlee, G.C., Fedoroff, N.V., and Brown, D.D. (1978). The nucleotide sequence of oocyte 5s DNA in *Xenopus laevis*. II. The GC-rich region. *Cell* 13, 717-725.
- Pardue, M.L., Brown, D.D., and Birnstiel, M.L. (1973). Location of the genes for 5s ribosomal RNA in *Xenopus laevis*. *Chromosoma* 42, 191-203.
- Pelham, H.R.B., and Brown, D.D. (1980). A specific transcription factor that can bind either the 5s DNA gene or 5s RNA. *Proc.Natl.Acad.Sci.* 77, 4170-4174.

- Peterson, R.C., Doering, J.L., and Brown, D.D. (1980). Characterization of two *Xenopus* somatic 5s DNAs and one minor oocyte-specific 5s DNA. *Cell* 20, 131-141.
- Petes, T.D. (1980). Unequal meiotic recombination within tandem arrays of yeast ribosomal DNA genes. *Cell* 19, 765-774.
- Potter, S.S., Brorein, W.J., Dunsmuir, P., and Rubin, G.M. (1979). Transposition of elements of the 412, copia, and 297 dispersed repeated gene families in *Drosophila*. *Cell* 17, 415-427.
- Reed, K.C., and Mann, D.A. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. *Nuc.Acids Res.* 13, 7207-7221.
- Rigby, P.W.J., Dieckmann, M., Rhodes, C., and Berg, P. (1977). Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J.Mol.Biol.* 113, 237-251.
- Rochaix, J.D., Bird, A., and Bakken, A. (1974). Ribosomal RNA gene amplification by rolling circles. *J.Mol.Biol.* 87, 473-487.
- Rosenthal, D.S., and Doering, J.L. (1983). The genomic organization of dispersed tRNA and the 5s RNA genes in *Xenopus laevis*. *J.Biol.Chem.* 258, 7402-7410.
- Slightom, J.L., Blechl, A.E., and Smithies, O. (1980). Human fetal  $\gamma$ - and  $\alpha$ -globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. *Cell* 21, 627-638.
- Smith, G.P. (1976). Evolution of repeated gene sequences by unequal crossover. *Science* 191, 528-535.
- Spradling, A.C., and Rubin, G.M. (1982). Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science* 218, 341-347.
- Strobel, E., Dunsmuir, P., and Rubin, G.M. (1979). Polymorphisms in the chromosomal locations of elements of the 412, copia and 297 dispersed repeated gene families in *Drosophila*. *Cell* 17, 429-439.
- VanArsdell, S.W., Denison, R.A., Bernstein, L.B., Weiner, A.M., Manser, T., and Gesteland, R.F. (1981). Direct repeats flank three small nuclear RNA pseudogenes in the human genome. *Cell* 26, 11-17.
- Wichman, H.A., Potter, S.S., and Pine, D.S. (1985). *Mys*, a family of mammalian transposable elements isolated by phylogenetic screening. *Nature* 317, 77-81.
- Wilde, C.D., Crowther, C.E., Cripe, T.P., Lee, M.G., and Cowan, N.J. (1982). Evidence that a human  $\beta$ -tubulin pseudogene is derived from its corresponding mRNA. *Nature* 297, 83-84.
- Wormington, W.M., Bogenhagen, D.F., Jordan, E., and Brown, D.D. (1981). A quantitative assay for *Xenopus* 5s RNA gene transcription *in vitro*. *Cell* 24, 809-817.

Wormington, W.M., and Brown, D.D. (1983). Onset of the 5s RNA gene regulation during *Xenopus* embryogenesis. *Dev.Biol.* 99, 248-257.

APPROVAL SHEET

The thesis submitted by Brenda A. Peculis has been read and approved by the following committee:

Dr. Jeffrey L. Doering, Director  
Associate Professor, Biology, Loyola

Dr. Howard M. Laten,  
Associate Professor, Biology, Loyola

Dr. John Smarrelli,  
Assistant Professor, Biology, Loyola

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the Committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Science.

12-18-86  
Date

Jeffrey L. Doering  
Director's Signature