ORGANISATION OF SEQUENCES HOMOLOGOUS TO THE MAJOR rDNA INSERTION OF DROSOPHILA MELANOGASTER

Simon John Kidd

Department of Biochemistry
Imperial College of Science and Technology
London SW7

Submitted for the degree of Ph.D. Summer 1980

ABSTRACT

The isolation and characterisiation of a cloned segment of D.melanogaster DNA (Dm219) which contains sequences homologous to the major class of ribosomal DNA (rDNA) insertion, but which originates from outside the rDNA is described. By characterising subcloned restriction fragments of Dm219 using restriction mapping and blot hybridisation techniques it was found to consist of five tandemly arranged units homologous to the rDNA insertion. Two of the units have identical deletions of 1.4Kb close to a position corresponding to the right end of the major class of rDNA insertion. Another unit contains an 8Kb substitution (the zeta element) for 2Kb of insertion sequences. This is located near sequences homologous to the left end of an rDNA insertion. In addition some variation was found between the restriction maps of the Dm219 units. This is more than the variability found between independently cloned examples of the major rDNA insertion. Thermal melt experiments indicate that the units of Dm219 are still closely related to rDNA insertions.

The arrangement of Dm219 sequences in the D.melanogaster genome was examined by blot hybridisation experiments and using DNA fractionated on cesium chloride actinomycin D equilibrium gradients. The results of these experiments was consistent with the structure of Dm219. The zeta sequences were found to occur 30 to 40 times per haploid genome, organized in seven major sequence arrangements. It was found by hybridising insertion and zeta sequences to a library of segments of D.melanogaster DNA cloned in bacteriophage λ, that at least 12 to 23% of the zeta sequences were linked to insertion sequences in the genome.

ACKNOWLEDGEMENTS

I want to thank my supervisor David Glover for his help and assistance over the three year course of this project.

I wish to thank Mike Lovett for his gradient polyacrylamide gels and I am indebted to the various people who made enzymes and DNA available to me. My thanks go to Heli Roiha, Debbie Smith and Mike Lovett for proof reading the thesis; and to Sarah MacGowan for her typing. Thanks go also to Sue Hayman for typing the tables in the figures of Appendix II.

The work was supported by an M.R.C. studentship to myself and a C.R.C. grant to David Glover.

CONTENTS

CHAPTER I

CHAPTER II

MATERIALS AND METHODS

CHAPTER III

(Chapter III continued)

CHAPTER IV

CHAPTER V

DISCUSSION

APPENDIX I

ISOLATION AND INDENTIFICATION OF SUBCLONES

APPENDIX II

RESTRICTION MAPPING cDm103 AND cDm219

(Appendix II continued)

TABLE CAPTIONS

FIGURE CAPTIONS

ABBREVIATIONS

bp	- Base pair, a nucleotide or nucleotide pair of single or double stranded nucleic acid.
BSA	- Bovine Serum Albumin.
DBM paper	- Diazobenzyloxymethyl-paper.
EDTA	- Diaminoethanetetra acetic acid, disodium salt.
ETS	- External Transcribed Spacer.
Kb	- Thousand base pairs, a thousand nucleotides or nucleotide pairs of single or double stranded nucleic acid.
L-plates	- Lauria agar (10g tryptone, 5g yeast extract, 5g NaCl and 15g agar/L) plates.
NO	- Nucleolus Organiser.
NTS	- Non Transcribed Spacer.
PM	- 0.02% ficoll, 0.02% BSA and 0.02% polyvinylpyrrolidone.
rDNA	- DNA sequences coding for rRNA.
rRNA	- ribosomal RNA.
SDS	- Sodium lauryl sulphate.
SSC	- 0.15M NaCl, 0.015M tri-Sodium citrate.
TE_2	- 10mM tris pH8.0, 2mM EDTA.
TEMED	- N N N' N' - Tetramethylethylenediamine.
TNE_2	- 10mM Tris pH8.0, 10mM NaCl, 2mM EDTA.
Tris	- Tris(hydroymethyl)methylamine.

CHAPTER I

1. INTRODUCTION

Eukaryotic ribosomes in general contain four RNA molecules (rRNA). These have sedimentation coefficients of approximately 28s, 18s,5.8s and 5s. The genes coding for the larger ribosomal RNA molecules have been extensively characterised in Xenopus laevis and Drosophila melanogaster. Their organisation and expression in X.laevis appears to be typical of that in many other eukaryotes. Some eukaryotes though, differ from X.laevis due to the gene for the large rRNA (21s to 28s) being interrupted by a DNA sequence not present in the final RNA molecule. The expression of these genes appears to be similar to that of other interrupted genes in eukaryotes.

The organisation and expression of the rDNA of D. melanogaster is more complex than other eukaryotes. After a short review of the rDNA of X.laevis and some other eukaryotes, that of D.melanogaster is discussed in this chapter.

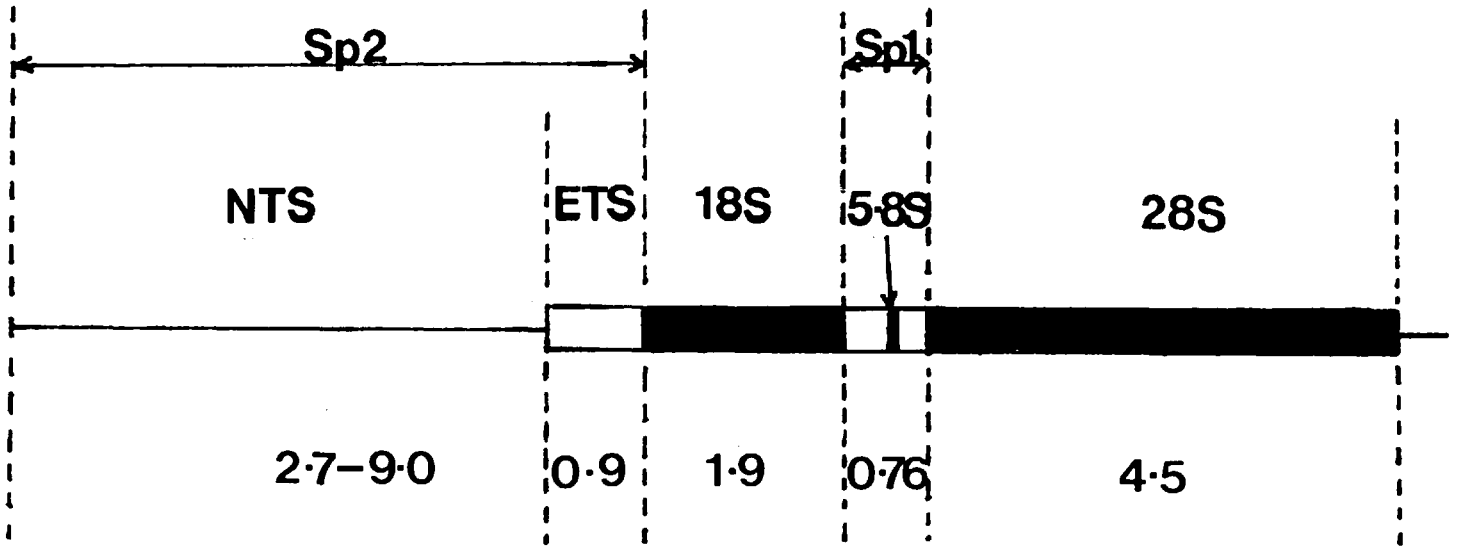
2. THE rDNA OF Xenopus laevis

Xenopus laevis contains 450 copies of the genes for 28s, 18s, and 5.8s rRNA (Brown and Weber, 1968). These are arranged in tandem in a single nucleolus organiser on one of the autosomes (Wallace and Birnstel, 1966). The genes are transcribed as a single 8.1Kb precursor, which is processed to give the mature rRNA's of the cytoplasm (Loening, 1969; Wellauer and Dawid, 1974; Reeder et al 1977). The organisation of the rDNA unit of X. laevis is shown in Figure 1. Within the transcription

Figure 1 The rDNA unit of Xenopus laevis

The non transcribed sequences of the rDNA unit are indicated by the thin line. The open boxes represent transcribed sequences which are removed during processing, and the position of the mature rRNAS in the rDNA unit are indicated by the filled in boxes (redrawn from Fedoroff, 1979).

Figure 1.



unit the sequences coding for the cytoplasmic rRNA's are separated from each other by spacer sequences. These are transcribed in the 8.1Kb precursor, but are not conserved during its processing. The positions of the rRNA's in the precursor has been reviewed by Reeder (1974), except that the polarity of the transcript is opposite to that described (Reeder et al., 1976; Dawid and Wellauer, 1976). The cytoplasmic 18s rRNA is located towards the 5' end of the precursor. It is separated from the 5' end by the external transcribed spacer (ETS). Between the 18s and the 28s rRNA coding regions is the internal transcribed spacer (spacer 1), which contains the 5.8s rRNA gene (Boseley et al., 1978). The 28s rRNA sequences are found at the extreme 3' end of the precursor (Sollner-Webb and Reeder, 1979). When actively transcribing rRNA genes are visualised in the electron microscope transcription matrices arranged in a head to tail fashion can be seen (Miller and Beatty, 1969). These are separated from each other by a non-transcribed spacer (NTS - see Figure 1). The rDNA units of X.laevis are heterogeneous in length, due to variation in the size of the NTS (Wellauer et al., 1976a). The NTS is internally repetitious, variation in the number of the repeating units leads to heterogeneity in the size of the NTS (Botchan et al., 1977).

The rDNA of X.laevis can also be extrachromosomal. This occurs in oocytes where the rDNA is amplified to a level greater than 2,000 times the diploid rDNA content. The final level reached is independent of the number of nucleolar organisers present in the cell (reviewed by Reeder, 1974). The extrachromosomal rDNA is found in nucleoli which contain several circular molecules. These contain many rDNA units in a tandem head to tail arrangement (Miller and Beatty, 1969). Heteroduplex experiments show that the units in long strands of amplified rDNA are homogeneous in size, indicating that the circular molecules are derived from single rDNA units (Wellauer et al., 1976b). Comparison of the restriction patterns of

amplified and chromosomal rDNA indicates that only a small proportion of the chromosomal rDNA units are amplified in oocytes (Bird, 1977).

3. THE rDNA OF SOME OTHER EUKARYOTES

Although the organisation of Xenopus laevis rDNA outlined above is typical of many eukaryotes, variations can occur. The location of the ribosomal RNA genes varies considerably. Species can have more than one nucleolus organiser per haploid genome, and these can be located either on the autosomes or sex chromosomes (Pardue et al., 1970; Hsu et al., 1975). Variation can also occur in the sizes of the regions shown in Figure 1. In eukaryotes the size of the 18s rRNA remains relatively conserved, being 2.01Kb in mammalian Hela cells and 2.06Kb in yeast. But in contrast, in more complex organisms the 28s rRNA increases in size, in Hela cells it is 5.03Kb, while in yeast it is only 3.7Kb (Loening, 1968). Hybridisation experiments give similar results, specific sequences in rRNA are conserved over a long evolutionary time span (Birnstiel and Grunstein, 1972). Of these sequences two thirds are located in 18s rRNA, indicating that both its sequence and size are conserved (Gerbi, 1976). In higher eukaryotes such as mammals the rDNA unit and the transcription unit are larger than in X.laevis. For example, in man the unit is 43Kb and the primary transcript 12Kb (Wellauer and Dawid, 1979). This means that a larger part of the rDNA unit is either not transcribed or is removed during the processing of the primary transcript.

In X.laevis the genes for 5s rRNA are not associated with the rDNA. Instead multiple copies of the 5s genes are found on probably all the chromosomes near the telomeres (Pardue, 1973). This lack of linkage between 5s and the other rRNA's appears to be typical of

most eukaryotes. Two exceptions to this are yeast and Dictyostelium where the 5s genes are located at the 3' end of the rRNA transcription unit, but at least in yeast are separately transcribed (Udem and Warner, 1972; Rubin and Sulston, 1973; Cramer et al., 1976; Maizels, 1976).

Amplified extrachromosomal rDNA is found in the oocytes of a wide variety of organisms (reviewed by Gall, 1969). In Tetrahymena another possible form of amplification occurs. Here the micronucleus, which acts as the 'germ line', contains a single integrated chromosomal rDNA unit (Yao and Gall, 1977). In contrast the macronucleus, which carries out metabolic functions, has approximately 30,000 extrachromosomal copes of rDNA. There are 19Kb palendromic molecules, each containing two rDNA units in a tail to tail arrangement (Karrer and Gall, 1976; Pearlman et al. 1974). Extrachromosomal rDNA in the form of palendromes has also been found in Physarum where the molecules are approximately 60Kb, and in Dictyostelum where they are 88Kb (Vogt and Braun, 1976; Cockburn et al., 1978).

4. THE rDNA OF MITOCHONDRIA AND CHLOROPLASTS

Both mitochondria and chloroplasts contain coding sequences for rRNA (reviewed by Chua and Luck, 1974). Chloroplasts contain several copies of rDNA organized in a similar fashion to E.coli, with the coding sequences for 5s RNA at the 3' end of the transcription unit (see Bedbrook et al., 1977 for references). The chloroplasts of Zea mays contain two copies of rDNA in an inverted orientation. It is possible that like E.coli, the 5s RNA is included in the primary transcript, which contains both the 23s and 16s rRNA's (Bedbrook et al., 1977).

Mitochondria contain only one copy of the rRNA genes. In fungal

and animal mitochondria only the larger rRNA's (21s to 25s and 15s to 19s) have been observed (Chua and Luck, 1974). The coding sequences for the large and small rRNA's are widely separated on the mitochondrial genome, and discrete precursors containing both rRNA's have not been detected (Chua and Luck, 1974; Borst and Grivell, 1978; Grimm and Lambowitz, 1979). Yeast petite mutants can still produce one rRNA when the coding sequences for the other are missing. This suggests that each rRNA gene has its own separate promotor (Borst and Grivell, 1978). This is very unlike the situation in X.laevis chromosomal rDNA.

5. SPLIT rRNA GENES IN EUKARYOTES

Many genes in eukaryotes have been found to be split into non-contiguous segments, by DNA sequences which are not represented in the final mature transcript (see Crick, 1979, and Abelson, 1974 for reviews). (These sequences, where they occur in the rDNA of Diptera will be referred to as 'insertions'. Otherwise the terminology of Gilbert (1978) will be used. Here the interrupting sequence is referred to as an "intron" and the surrounding coding sequences as "exons". An alternative name for the non-coding sequence is intervening sequence (IVS).) Introns have been found in the genes coding for the large rRNA (21s to 28s) of a variety of species. In some strains of yeast the mitochondria contain a 1Kb intron in the 21s gene (Bos et al., 1978). The mitochondrial 25s rRNA gene of Neurospora crassa has a 2.3Kb intron approximately two thirds in from the 3' end of the gene (Hahn et al., 1979). A 0.94Kb intron occurs in a similar position in the 23s genes of Chlamydomonas reinhardii chloroplasts (Rochaix and Malnoe, 1978). Some strains of Tetrahymena pigmentosa have a 0.39Kb intron near the 5'

end of the 25s gene (Wild and Gall, 1979). An intron of similar size, and in the same position has also been found in T. thermophila (Cech and Rio, 1979). A different situation is found in Physarum polycephalum where most of the 26s genes are split by two introns of 0.68 and 1.21Kb (Campbell et al., 1979). In all the above examples, with the exception of Physarum, all the rRNA genes are split, and as mature rRNA is still found, the genes must be functional. In D.melanogaster and some other Diptera the situation is different, since only some of the 26s genes are interrupted. The organization of Drosophila rDNA is described in the following sections.

6. THE rDNA OF Drosophila melanogaster

The rRNA of D. melanogaster

The largest precursor to rRNA detected in Drosophila melanogaster is 8Kb. The 5' end of this transcript has not undergone processing as it still contains a 5' triphosphate. This suggests that it is the primary transcript (Levis and Penman, 1978). It is processed into the cytoplasmic 26s and 18s rRNA's. The 26s rRNA consists of four separate molecules hydrogen bonded together. These are 5.8s, 2s and two of approximately 18s in size. The two 18s molecules result from a central cleavage of 26s rRNA in the nucleus (Jordan et al., 1976). This also removes approximately 120 bp from the centre of 26s (Wellauer and Dawid, 1977). The 5.8s and 2s are then produced by later processing events (Jordan et al., 1976).

The Location of the Genes for rRNA in D. melanogaster

Drosophila melanogaster has one nucleolar organiser (NO) on both the X and Y chromosomes. The NO are located on the long arm of the X

Figure 2 The X chromosome of Drosophila melanogaster

The thin and thick lines represent the euchromatin and the heterochromatin respectively of the X chromosome. NO is the nucleolus organiser and K the centromere. The diagram shows the approximate positions of various inversions of the X chromosome of cr, Xh^{abo} and some non-rDNA Type I sequences (see text, redrawn from Procunier and Tartof, 1978).

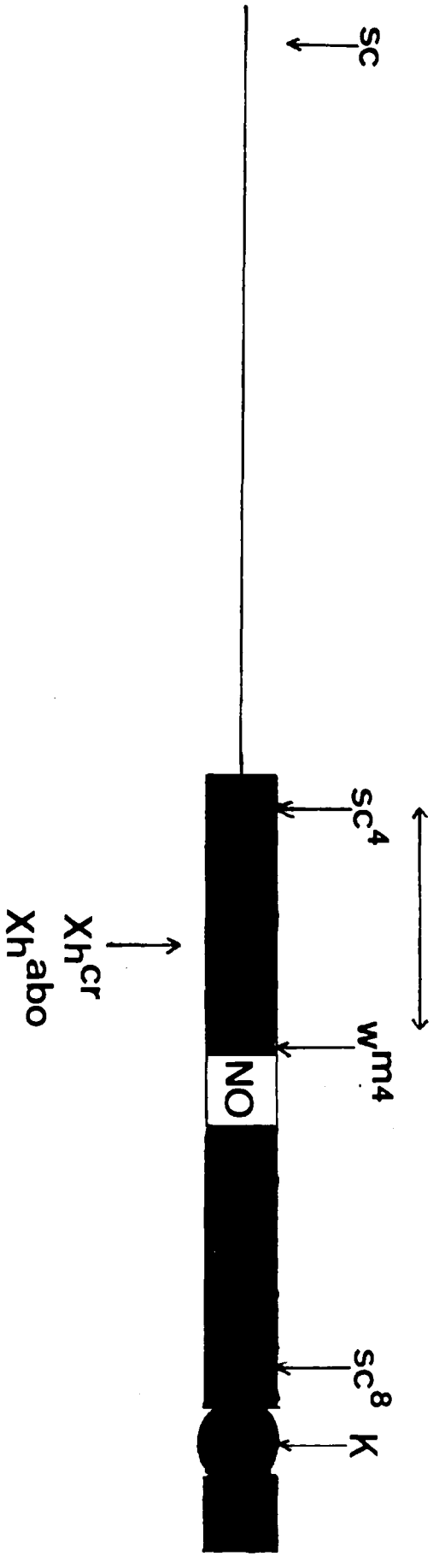


Figure 2.

chromosome, in the heterochromatin surrounding the centromere (see Figure 2), and on the short arm of the Y chromosome (Ritossa, 1976). This chromosome is completely heterochromatic in most tissue (Williamson, 1976). The NO of the X chromosome lies between the proximal breakpoints of the sc^4 and sc^8 inversions (Figure 2). Recombination between X chromosomes carrying these inversions result in the region between the proximal breakpoints either being duplicated in the sc^8-sc^4 chromosome ($X^{sc^8-sc^4}$) or deleted in the sc^4- sc^8 chromosome ($X^{sc^4-sc^8}$). Ritossa and Spielgelman (1965) have used these chromosomes to construct flies that have between one and four NO's. They carried out saturation hybridisation of rRNA to DNA extracted from adult males and females of the Urbana strain of D.melanogaster, and found 0.27% of the genome, in both sexes, coded for rRNA. This indicated that the NO's of the X and Y chromosomes contain similar amounts of rDNA. In similar experiments with male flies with one NO($X^{sc^4-sc^8}/Y$) half the amount of rDNA was found. Flies with three ($X^{sc^8-sc^4}/Y$) and four ($X^{sc^8-sc^4}/X^{sc^8-sc^4}$) NO's had approximately 1.5 and 2 times the normal amount of rDNA. This shows that the amount of rDNA correlates with the number of NO's in the genome (Ritossa and Spielgelman, 1965). Confirmation of this was obtained by the in situ hybridisation of radioactive rRNA or rDNA to the nucleolus of salivary gland polytene chromosomes (Pardue et al., 1970 ; Glover et al., 1975). The amount of rDNA in the Urbana strain is the equivalent of about 100 copies per NO (Ritossa and Spielgelman, 1965). Similar experiments with Oregon R strain adult females shows that the X chromosome contains about 250 copies of rDNA (Tartof, 1973). This indicates that strains of D. melanogaster differ in their rDNA content. Other experiments have shown that the Y chromosome in Oregon R flies contains about half the rDNA found in the X chromosome (Spear, 1975).

Like X. laevis the 5s RNA genes of D. melanogaster are not linked to the rDNA. They are found at 56F on the right arm of the second chromosome (Prensky et al., 1973).

The Structure of the rDNA units of D. melanogaster

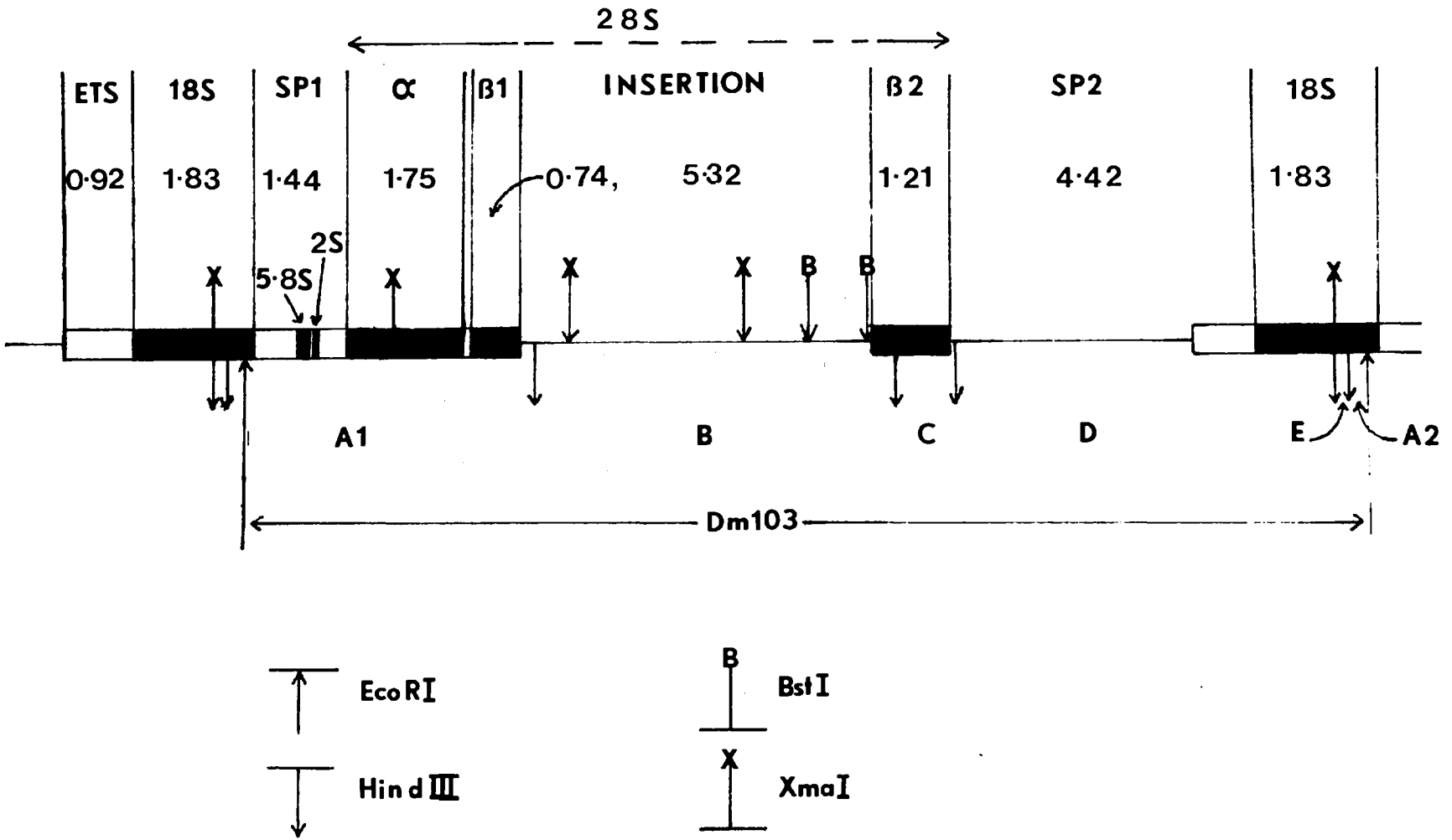
When radioactive rRNA is hybridised to EcoRI cleaved genomic DNA by the technique of Southern (1975), four major fragments can be seen. The two most strongly hybridising fragments are 17 and 11.6Kb, and the two minor fragments are 7.4 and 5.4Kb (Tartof and Dawid, 1976). The 17 and 11.6kb fragments result from the rDNA units that contain a single EcoRI site in the 18s gene. Hence these two fragments contain an entire repeating unit (Glover and Hogness, 1977; White and Hogness, 1977; Wellauer and Dawid, 1977; Pellegrini et al., 1977).

The organization of the rRNA coding sequences in these fragments was first determined in the cloned 17Kb EcoRI fragment Dm103 (Glover et al., 1975). This has been analysed both by R-looping (White and Hogness, 1977) and saturation hybridisation of rRNA to its subcloned Hind III fragments (Glover and Hogness, 1977). The 26s rRNA gene of Dm103 was found to be split into two coding regions by a 5.3Kb insertion (intron), 2.5Kb from the 5' end of the gene. In the map of Dm103 in Figure 3, it can be seen that the insertion does not correspond to the position of the central cleavage in 26s rRNA. With the exception of the insertion the organisation of the rRNA genes in Dm103 is similar to that of X. laevis (compare figures 1 and 3). Both the 5.8s and 2srRNA coding sequences are located in the internal transcribed spacer close to the 26s gene (Pellegrini et al., 1977; Jordan and Glover, 1977).

Figure 3 The 17Kb EcoRI rDNA fragment Dm103 from Drosophila melanogaster

Dm103 is shown together with its surrounding rDNA, the coding and non-coding sequences are indicated in the same manner as in Figure 1. The restriction endonuclease cleavage sites for EcoRI, Xma I, Bst I and Hind III are indicated. The letters below the diagram are the Hind III and Hind III/Eco RI fragments of Dm103 (Glover and Hogness, 1977). The position of the external transcribed spacer (ETS) is from Wellauer and Dawid (1978) (Redrawn from White and Hogness, 1977).

Figure 3.


28S
ETS
18S
SP1
α
ß1
INSERTION
ß2
SP2
18S
0·92
1·83
1·44
1·75
0·74,
5·32
1·21
4·42
1·83
2S
5·8S
X
B
A1
B
C
D
E
A2
Dm103
EcoRI
HindIII
BstI
XmaI

Other examples of the 17Kb EcoRI fragments have been cloned (Dawid and Botchan, 1977). The insertion in these clones was shown not to be internally repetive, either at the level recognised by restriction endonucleases, cross hybridisation experiments, or DNA:DNA reassociation kinetics (Wellauer and Dawid, 1978; Dawid and Long, 1979). EcoRI fragments containing 26s genes with smaller insertions have also been cloned. The shorter insertions in these fragments are 1.0 and 0.5Kb, they were found to be homologus to the right hand side of the insertion in the 17Kb fragments. The 1.0Kb insertions contain two BstI sites similar to those of Dm103 (see Figure 3), while the 0.5Kb insertions have only one. Although on this basis, insertions of 1.0Kb might be a repeat of those of 0.5Kb, it does not appear to be so (Dawid et al., 1978; Wellauer and Dawid, 1978).

The characterisation of the 11.6Kb EcoRI fragments shows that the 26s gene in these rDNA units is contiguous. Other than the presence of the insertion, the 17 and 11.6Kb fragments appear to be identical (White and Hogness, 1977; Wellauer and Dawid, 1977; Pellegrini et al., 1977).

Insertions such as that of Dm103 which lack EcoRI sites have been called Type I sequences. Another class of insertion called Type II sequences has been described which contain EcoRI sites. These split the 26s gene in the same position as Type I insertions but show no sequence homology with them. EcoRI cleavage of rDNA units with Type II insertions produces two fragments of 7.4 and 5.4Kb, which contain the right and left halves of the unit respectively. These EcoRI fragments have been cloned and analysed (Wellauer et al., 1978; Dawid et al., 1978; Wellauer and Dawid, 1978). Complete rDNA units with Type II insertions have been examined using cloned segments of randomly sheared rDNA inserted into the bacterial plasmid by the dA:dT

tailing technique. The Type II insertions of four cloned rDNA units have all three identically arranged EcoRI sites. Cross hybridisation experiments show that the insertions of the four clones are homologous. But variation in the positions of cleavage sites for restriction endonucleases other than EcoRI can be seen (Roiha and Glover, manuscript in preparation). Some clones containing the 5.4Kb EcoRI fragment from the left half of rDNA units with Type II insertions have been isolated where the interrupting sequences are not homologous (Wellauer and Dawid, 1978). This suggests there might be more than one class of Type II insertion.

The rDNA of the X and Y chromosomes

Drosophila melanogaster contains both split and contiguous 26s genes. The distribution of these classes of rDNA unit between the X and Y chromosomes can be determined using flies lacking either the X or Y nucleolus organiser (NO). Restriction analysis of DNA from adult females (X/X) shows that they have the 17, 11.6, 7.4 and 5.4Kb EcoRI fragments, corresponding to contiguous genes, and those split with Type I and II insertions. In contrast the rDNA of flies with only a Y NO ($X^{sc^4-sc^8}/X^{sc^4-sc^8}/Y$ females and $X^{sc^4-sc^8}/Y$ males) almost completely lacks the 17Kb EcoRI fragment, suggesting that they only have contiguous genes and those interrupted by Type II insertions (Wellauer et al., 1978). As exchange can occur between the X and Y NO's (Ritossa, 1976), and the rRNA's transcribed from them are probably identical (Maden and Tartof, 1974), a mechanism presumably exists which restricts Type I insertions to the X chromosome.

The chromosomal arrangement of contiguous and non-contiguous genes has been examined using purified rDNA longer than one repeating unit. This can be used to form R-loops or DNA:RNA hybrids with rRNA and viewed in the electron microscope (Wellauer and Dawid, 1977;

Pellegrini et al., 1977; Wellauer et al., 1978). In DNA extracted from male and female embryos carrying a Y chromosome partially deleted for the NO (Y^{bb}), 86% of the rDNA is from the X chromosome. It contains three major size classes of insertion, these are less than 1.4Kb, 1.5 to 4.0, and 4.1 to 6.5Kb. The classes were found in 14, 16 and 35% of the rDNA units respectively. EcoRI cleavage removes the middle class (1.5 to 4.0Kb) indicating that they are Type II insertions. The largest class then, probably corresponds to rDNA units such as Dml03 with 5Kb insertions, while the smallest contains the small Type I insertions (Wellauer et al., 1978). Contiguous and split genes are randomly interspersed on the X chromosome. One exception to this are insertions between 0.5 and 3.5Kb, which show some clustering (Wellauer and Dawid, 1977), these probably correspond to Type II insertions. In rDNA extracted from flies with only the Y NO, 16% of the units were interrupted, and these were found to be clustered. The size of the majority of the insertions corresponds to Type II sequences. Only 3% are similar in size to Type I insertions, and might be due to contaminating X chromosomal material (Wellauer et al., 1978).

The Non Transcribed Spacer

Like the rDNA of X. laevis, spacer 2 of D. melanogaster is heterogeneous in size. Heteroduplexes between different EcoRI fragments contain deletion loops in varying positions within this spacer, indicating that it is internally repetetive. The loops do not extend into a region adjacent to the 5' end of the 18s gene (Wellauer et al. 1978). R-looping experiments using the 34s rRNA precursor show that the region extending 0.92Kb from the 5' end of the 18s gene contains the external transcribed spacer (ETS, see Figure 3) (Dawid et al., 1978). So the region where the deletion loops occur corresponds to the nontranscribed

spacer (NTS). The NTS of cloned rDNA units from the X and Y chromosomes are closely homologous in thermal denaturation and heteroduplex experiments (Wellauer et al., 1978). Restriction mapping with the enzyme Alu I has shown that the NTS contains a 250 bp repeating sequence. This repetetive region extends from the 5' end of the ETS to, in most rDNA units, 1.9Kb from the 3' end of the 26s gene (Long and Dawid, 1979a).

The Chromosomal Location of rDNA Insertion Sequences

The chromosomal location of cloned segments of DNA can be found by in situ hybridisation to salivary gland polytene chromosomes. As was expected for a plasmid containing rDNA, sequences homologous to Dm103 are found at the nucleolus, but in addition hybridisation was also found to other regions of the salivary gland chromosomes (D. Finnegan and W.J. Peacock, cited by White and Hogness, 1977). The sequences giving the non-nucleolar hybridisation are found in Hind III fragment Dm103/B (see Figure 3, Peacock et al., manuscript in preparation). Thus sequences homologous to the Type I insertion are found elsewhere in the Drosophila genome..

A similar conclusion was reached by fractionating D.melanogaster genomic DNA on cesium chloride gradients containing various dyes (Dawid and Botchan, 1978; Dawid and Wellauer, 1978). Sequences homologous to Type I insertions were found in two regions of the gradient. The first at lower density corresponds to Type I rDNA insertions. The second of similar size occurs midway between the density of rDNA and pure insertion sequences. This shows that insertion sequences can be found linked to sequences of a different density to rDNA.

There are about the same amount of Type I insertion sequences linked to rDNA and to different sequences, this is about 400Kb in

each class. The non rDNA insertion sequences appear to be arranged in blocks of less than 20Kb but greater than 5Kb. They have a similar G:C content to rDNA insertions (55%) and form well matched hybrids with them. Restriction analysis though, shows that the two sequences have very different restriction patterns (Dawid and Botchan, 1978; Dawid and Wellauer, 1978). Non rDNA Type II insertion sequences have not been detected (I.B. Dawid, cited by Wellauer and Dawid, 1978).

rDNA in Diploid and Polytene Tissue of D.melanogaster

In the polytene chromosomes of 3rd instar larval salivary glands the euchromatin, where most genetic loci have been located is replicated to give on average 1024 copies per chromosome (Rasch, 1970). In contrast the heterochromatin which contains repetetive DNA is replicated only a few times or not at all (Gall et al., 1971). The rDNA although it is localised in the heterochromatin, is transcriptionally active and so might have different replication properties to the surrounding DNA. This can be verified in 3rd instar larvae by comparing the rDNA content of imaginal disks and brains (diploid tissue) with that of salivary glands (polytene tissue).

In diploid tissue for Oregon R females with one nucleolus organiser ($X^{sc^4-sc^8}/X$) the rDNA content was half (0.3%) that found in X/X females with two NO's (0.512%), while in males (X/Y) only 0.39% was found (Spear, 1975). This indicates that in diploid tissue, the rDNA content is dependent on genotype, and that the X and Y chromosomes contain different amounts of rDNA.. The rDNA of the X chromosome constitutes 0.256% of the genome, while that of the Y chromosome is only 0.134%. In contrast to this, polytene tissue from females with one or two NO's from the X chromosome has an rDNA content of 0.1%.

This indicates that the rDNA of the X chromosome is disproportionately replicated if the homologous chromosome is missing, or deficient for the nucleolus organiser. It is also possible that if the amount of rDNA is less than that found in X/X tissue, disproportionate replication occurs. For example, the rDNA content of X/Y polytene tissue (0.11%) is similar to that of X/X tissue. The nucleolus of the Y chromosome will also undergo disproportionate replication, as polytene X_{NO-}/Y tissue has an rDNA content of 0.1% which is similar to that of X/Y tissue. Similar rDNA contents of approximately 0.1% are obtained in female polytene tissue with three or four NO's (Spear, 1975). This shows that in salivary gland polytene tissue the rDNA content is not dependent on genotype. It replicates independently of the surrounding heterochromatin, but is under-replicated relative to the euchromatin (Spear, 1975).

Differences between the rDNA of diploid and polytene tissue has also been examined using the technique of Southern (1975). This allows the independent replication of the various classes of rDNA unit in D. melanogaster to be analysed. In the stock of Oregon R used in these experiments both diploid tissue from males and females contained 17 and 11.6Kb EcoRI fragments. These represent genes interrupted with 5Kb Type I insertions and contiguous genes respectively. In addition to small amounts of the 11.6Kb fragment, males contained a major 12Kb band. This was present in only small amounts in females, suggesting it was derived from the Y chromosome. Its presence on the Y chromosome was confirmed using tissue from larvae deficient for the nucleolus of the X chromosome. In addition, as expected larvae with only the Y nucleolus lacked the X chromosome specific 17Kb EcoRI fragment. The additional sequences on the 12Kb fragment were found to lie in the equivalent of Hind III fragment D of Dm103 (see Figure 3

Endow and Glover, 1979). This contains the NTS, plus a portion of the 18s gene. As insertions have not been found in the 18s gene, the additional sequences probably lie in the NTS. It suggests that in this stock of Oregon R there is a specific size class of NTS on the Y chromosome.

By using these EcoRI fragments as chromosome specific markers, differential replication of the X and Y chromosomes in polytene tissue could be analysed. In polytene tissue from males (X/Y), the major hybridisation is to the Y specific 12Kb fragment. Less hybridisation was seen to the 11.6Kb fragments of both the X and Y chromosomes, and none to the X chromosome specific 17Kb fragment. The proportion of this fragment in X/X polytene tissue, relative to the contiguous 11.6Kb units, is one third of that in diploid tissue , indicating that contiguous genes were over-replicated. None of these comigrate with the Y specific 12Kb EcoRI fragments (Endow and Glover, 1979). This suggests that during chromosome polytenisation specific classes or rDNA units are replicated. In males this is the Y chromosome specific contiguous repeat, while in females the contiguous repeats are replicated in preference to the interrupted rDNA units (Endow and Glover, 1979). This is similar to the amplication of specific rDNA units in the oocytes of X.laevis (Bird, 1977), although the structure of the differentially replicated repeats in D.melanogaster is unknown. In the tandemly arranged rDNA the 11.6 and 17Kb units are randomly interspersed on the X chromosome (Wellauer and Dawid, 1977). To obtain the specific replication of the contiguous repeats, only short regions of rDNA, with a small number of units can be replicated. This suggests that the disproportionate replication of rDNA in X_{NO}-X polytene tissue is not due to the nucleolus organiser undergoing another round of replication

(Spear, 1976) but to the replication of small regions of it.

The independent replication of rDNA in polytene tissue has been suggested to be the cause for the dosage compensation effect. This occurs in adult X_{NO}-/X females and X/O males, which have an rDNA content only slightly lower than X/X flies. The increased amount of rDNA on the X chromosome is however not inheritable, and does not occur on the Y chromosome in X_{NO}-/Y flies (Tartof, 1973). Polytene tissue has been found in adults (Ashburner, 1970). They also have a lower rDNA content than diploid tissue (0.45% in X/X adults to 0.512% in X/X diploid tissue, Tartof, 1973; Spear, 1975). This might be due to the under-replication of rDNA due to the presence of polytene tissue in adults. There are several objections to this hypothesis. Firstly, the rDNA content of X/O polytene tissue would have to be higher than that in diploid tissue to account for the amount in X/O adults. This does not accur in larvae, where it is under-replicated relative to diploid tissue (Spear, 1975). But since the rDNA can replicate independently (Spear, 1975; Endow and Glover, 1974) it does not necessarily have to under-replicate in the polytene tissue of adults. Secondly, the rDNA of the Y chromosome does not dosage compensate in adults (Tartof, 1973) but it undergoes disproportionate replication in polytene tissue (Spear, 1975). A comparison of the rDNA content of diploid tissue with that of adults suggests that the Y chromosome might undergo dosage compensation. The rDNA of the Y chromosome in X_{NO}-/Y adults constitutes 0.175% of the genome (Tartof, 1973), which is higher than that in the Y chromosome in diploid tissue (0.134%, Spear, 1975), the difference might be due to dosage compensation. It would be interesting to see if specific classes of rDNA repeat are replicated during dosage compensation, in the same way they are in polytene tissue.

7. Insertions in the rDNA of Other Diptera

By analysing different species of Diptera, it is possible to find if various features of the rDNA are conserved over long evolutionary time spans. Any conservation might indicate that the feature has (or had) some functional significance.

The rDNA of a variety of species of Drosophila has been examined. Some species within the D. melanogaster subgroup also have the 17 and 11.6Kb EcoRI fragments. Tartof (1979) has found the proportion of 'putative' interrupted 17Kb rDNA units decreases relative to the 11.6Kb fragments in species with low amounts of rDNA in the genome.

The D. melanogaster subgroup is a member of the subgenus Sophopora (Patterson and Stone, 1952). Insertions have been found in the 26s genes of D. hydei and D. virilis. These are members of different subgroups of the subgenus Drosophila. In D. hydei only 5.5Kb insertions are found, and like those of D.melanogaster, they are limited to the X chromosome (W. Kunz, cited by Jordan, 1979). Hybridisation experiments indicate that the rDNA insertions of D.hydei and D.melanogaster are partially homologous (Barnet and Rae, 1979).

In D.virilis the major insertion is 9.6Kb. Smaller insertions of 0.3Kb occur less frequently, and it is thought that contiguous 26s genes are found in few if any rDNA units (Barnet and Rae, 1979). The restriction map of the rRNA coding sequences in D.virilis is similar to that of D.melanogaster, and the insertions occur in the same position in the 26s gene. The 9.6Kb insertion appears to consist of two diverged repeating sequences. Its restriction map is different to that of Dml03, but hybridisation and thermal denaturation experiments show that the two contain closely related sequences.(Barnet and Rae, 1979). Like the Type I insertions of D.melanogaster, sequences homologous to the 9.6Kb insertion can be found

linked to non rDNA sequences (Barnet and Rae, 1979).

Figure 4 shows the phylogenic relationships between the infraorders of the order of Diptera. D.melanogaster belongs to the superfamily Drosophilidea of the infraorder Myiomorpha. Sequences homologous to the 0.88Kb Bst I fragment, but not to the 4.5Kb Bst I/Hind III fragment of the Dml03 insertion (see Figure 3) have been found in genomic DNA from Sarcophaga bullata and Musca domestica. It is likely that this is due to the presence of rDNA insertions in these species (Barnet and Rae, 1979). These flies are members of the superfamilies Sarcophagidea and Muscidea respectively, which are part of the infraorder Myiomorpha (Rohlandorf, 1974). No homology could be seen between the insertion of Dml03 and DNA from Sciara coprophila, the crane fly or the robber fly (Barnet and Rae, 1979). These species are members of the infraorders Bibionomorpha, Tipulomorpha and Asilomorpha respectively.

The approach by Barnet and Rae (1979) of hybridising D.melanogaster insertion probes to genomic DNA from other species, using the technique of Southern (1975) has its limitations. Insertions in the rDNA of other species may have little or no homology with those of D.melanogaster, and they may be present in only a few copies, and so might not be detected. The rDNA of various Diptera has been cloned and analysed. S. copiophila has 45 copies of rDNA per haploid genomes 90% of which have a repeat length of 8.4Kb (Renkawitz et al., 1979). These have been cloned and the 26s gene shown to be contiguous (Renkawitz et al., 1979). Further analysis of S. coprophila has shown that a small proportion of the 26s genes are interrupted (R. Renkowitz-Pohl, cited in Renkawitz et al., 1979). Ribosomal DNA of Calliphora erythrocephala has been cloned and some of 26s genes were found to be split by a 2.5Kb insertion. This contains sequences homologous to the right hand end of the Dml03 insertion (Beckingham and White, manuscript submitted). C.erythrocephala like D. melanogaster is a member of

Figure 4 The phylogenic relationships between the infraorders of the order Diptera

(Redrawn from Rohlendorf, 1974)

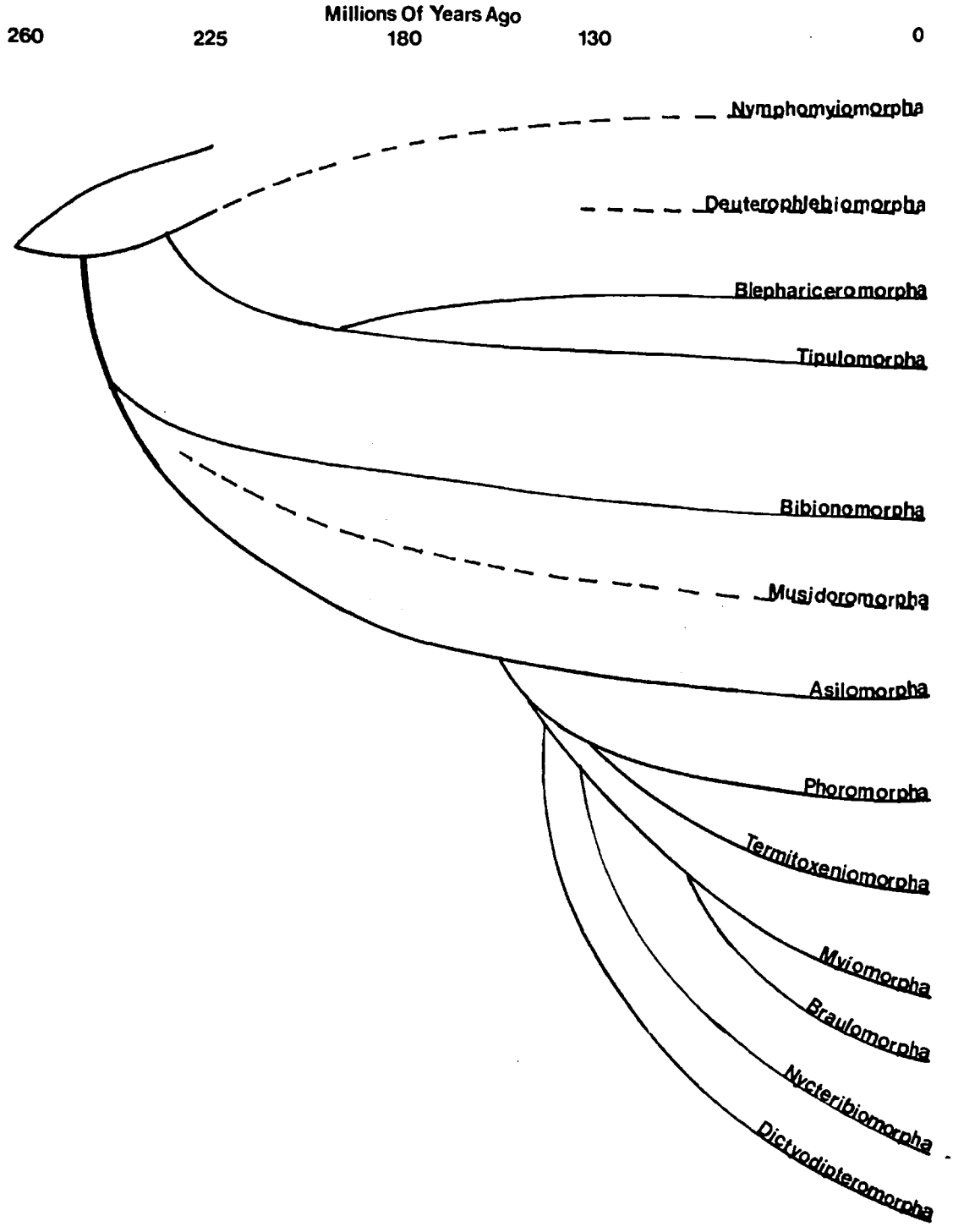

Millions Of Years Ago
260
225
180
130
0
Nymphomyiomorpha
Deuterophlebiomorpha
Blephariceromorpha
Tipulomorpha
Bibionomorpha
Musidoromorpha
Asilomorpha
Phoromorpha
Termitoxeniomorpha
Myiomorpha
Braulomorpha
Nycteribiomorpha
Dictyodipteromorpha

Figure 4.

the infraorder Myiomorpha but of a different superorder, Sarcophagidea. Ribosomal DNA has been cloned from Chironomus tentans, a member of the infraorder Tipulomorpha. It is similar to S.coprophila, it has only 43 to 72 copies of rDNA per haploid genome (Hollenburg, 1976), and a small repeat size of 8.4Kb. When these were cloned, the 26s gene was found not to contain an insertion (Degelmann et al., 1979). These cloned fragments do not represent all the the rDNA units in C. tentans as radioactive rRNA hybridises to larger fragments in genome DNA (Dengelmann et al., 1979) which might be due to a small number of rDNA units with insertions.

It is difficult to generalise on the basis of this data on the evolutionary aspects of rDNA insertions in Diptera. It does indicate that in lower Diptera (those not in the infraorder Myiomorpha) such as S.coprophila, and probably C.tentans, the crane fly and robber flies that very few insertions are found. Any insertions that they do have are probably not homologous to those of D.melanogaster. Both the Type I and II insertions of D.melanogaster, and the insertions of D.virilis and C.erythocephala occur in the same position, near the 3' end of the 26s gene. With the exception of Physarum the introns (insertions) in the large rRNA (26s) genes of non Diptera also occur near the 3' end. A similar situation occurs in vertebrate globin genes, where the introns are found in the same position, although the various genes diverged over 500 million years ago (see Konkel et al. 1979) for references). This does not imply that the rDNA insertions in Diptera and non Diptera are related, but it might indicate some significance about this region of the large rRNA gene. Among the higher Diptera the insertions are related, as sequences homologous to the right hand end of the Dm103 insertion are found in all those examined. This means that they have been preserved since the superfamilies of the Myiomorpha diverged 130 to 180 million years ago (Rohlendorf, 1974; see Figure 4).

It is likely that the interrupted rDNA units in D. melanogaster are not used to produce rRNA (Long and Dawid, 1979b, see also next section). This is also suggested by their lack of replication in polytene tissue (Endow and Glover, 1979), and by the decreased proportion of interrupted rDNA units in flies with low amounts of rDNA, such as species of the D.melanogaster subgroup and S.coprophila (Tartof, 1979; Renkawitz et al., 1979). So at low levels of rDNA, there may be selection pressure to remove split rRNA genes, or not to replicate them during polytenisation. This conflicts with their continual presence in the rDNA because cells presumably have mechanisms of removing non-functional units. One of the proposed mechanisms of maintaining homogeneity in tandemly repeated sequences such as rDNA is unequal exchange both between sister chromatids and chromosomes (Smith, 1973). This hypothesis has been used to account for the sequence organisation of X.laevis rDNA, which can be divided into two regions, coding and non-coding sequences. The coding sequences are homogeneous in the tandem array, and are closely related in different Xenopus species (Forsheit et al., 1974). Mutations which adversely affect the function of rRNA are a selective disadvantage and are removed from the population. The reverse occurs if they confer an advantage on the animal. They are then spread throughout the rDNA by unequal exchange, thus maintaining the homogeneity of the coding sequences in the tandem arrays. In contrast, the non-coding NTS consists of repetetive DNA, it is heterogeneous in size in rDNA units of the same nucleolus, and differs in sequence in different species (Wellauer et al., 1976; Forsheit et al., 1974). This can be accounted for by its ability to accept sequence changes without impairing its function (Smith, 1973). Non-functional interrupted rDNA units might be maintained because they are transcriptionally inactive (Long and Dawid, 1974). If they do not produce RNA which adversely

affects the function of ribosomes they will not be selected against, and then will be maintained as 'silent' genes (Fedoroff, 1979). This is attractive, because not all of the rDNA is needed to give a wild type phenotype (Ritossa, 1976), and not all the rDNA is being transcribed when visualised by the Miller spreading technique (McKnight and Miller, 1976). If the interrupted units were non functional 'silent' genes they should have undergone sequences changes in the same way as the NTS. However, the coding and spacer sequences appear to be identical in contiguous and interrupted repeats of D.melanogaster (White and Hogness, 1977; Wellauer and Dawid, 1977; Pellegrini et al, 1977; David et al., 1978; Long and Dawid, 1979a), though this has not been tested by DNA sequencing. Furthermore the rRNA coding sequences also appear to be almost identical in higher Diptera, although their insertions have diverged (Barnet and Rae, 1979; Beckingham and White, manuscript submitted). This is similar to the split genes of vertebrates, such as globin, where the intron (insertion) diverges more rapidly than the coding sequences (see Konkel et al., 1979 for references). It suggests that, at some stage in the fly's life cycle, the insertion or the interrupted 26s gene has some function.

8. Expression of Split Genes

As well as occurring in Drosophila rDNA, introns (insertions) are found in a variety of other genes. In most cases these are genes that occur only once per haploid genome, and so the cell needs a mechanism to produce mature RNA from them (see Crick, 1979; Abelson, 1974). The bulk of the evidence today shows that a primary transcript is produced which contains both intron and exon sequences. The introns are then 'spliced' out and the exons ligated together to form the mature transcript. Evidence for the primary transcript, has been found for a variety of interrupted genes, such as chick ovalbumin (Roop et al., 1978), mouse β

globin (Ross, 1976; Curtis and Weissmann, 1976; Tilghman et al., 1978) and yeast tRNA (Knapp et al., 1978; O'Farrel et al., 1978). For the latter two the transcript has been shown to be a precursor to mature RNA.

In Drosophila the situation is more complex due to the presence of uninterrupted rDNA units. This means that it is not necessary for the production of mature rRNA that rDNA units containing insertions are transcribed. Long and Dawid (1979b) have demonstrated by DNA:RNA re-association kinetics, and hybridisation of insertion probes to electrophoretically fractionated RNA immobilised on DBM paper that potential precursors from rDNA units with Type I insertions are present at less than one copy per nucleus. Only a heterogeneous array of molecules homologous to the insertion were found, and these are found at very low levels in the cell. This indicates that rDNA units with Type I insertions are not used to produced mature rRNA.

An alternative approach has been to measure the size of actively transcribing units using the Miller spreading technique. For D.hydei the transcription units of spermatocytes are homogeneous, indicating that only units lacking an insertion are transcribed (Glatzer, 1979). Similar experiments using embryos and nurse cells of D.melanogaster gives different results. Here a more heterogeneous size distribution of transcription units is found (Chooi, 1979). As less units were measured in these experiments it is possible that the increased heterogenerity is due to stretching of the chromation. It was suggested that this is not the case because the number of fibres per transcription unit correlates with the length of the unit, and when an allowance is made for stretching, a significant proportion of the units are still longer than the standard unit (defined as an uninterrupted rDNA unit). The distribution of the long transcription units between the X and Y chromosomes was determined using nurse cells of X/X and $X^{sc^4-sc^8}/X^{sc^4-sc^8}/Y$ genotypes respectively. This showed that 21% of those

on the X chromosome, and 16% on the Y chromosome were longer than standard units. This compares with the observation that 65% of the units on the X chromosome and 16% on the Y chromosome contain insertions (Wellauer et al., 1978). If the longer transcription units represent rDNA units with transcribed insertions, it conflicts with the data of Long and Dawid (1979b) who found no precursors in ovaries (which contain the nurse cells) or embryos. It is possible that the longer transcription units represent Type II insertions, but there are more of these units on the X chromosome than there are Type II insertions, and they are randomly interspersed with standard units (Chooi, 1974), which also probably does not occur for this class of insertion (Wellauer and Dawid, 1977). Two pieces of data favour the interpretation that the longer transcription units represent Type II insertions. Firstly similar numbers (16%) are found on the Y chromosome, and secondly the longer units are smaller than the size of full length Type I insertions. It is also possible that the mechanism for producing mature transcripts from interrupted rDNA in Drosophila is very different from other interrupted genes.

It seems likely that cells possess several enzyme systems for splicing, which probably have different modes of action. Some evidence for this comes from the DNA sequence of intron:exon boundaries. Similar sequences are found at the boundaries in mRNA producing genes (see Abelson, 1979, for a review) in widely different species. For example a similar sequence is found in mammals and in the fibroin gene of silkmoth, Bombyx mori (Tsujimoto and Suzuki, 1979). In contrast the introns which occur in some yeast tRNA's show no homology with those in mRNA precursors (Goodman et al., 1977; Valenzuela et al., 1978). The ends of the introns in the rDNA of T.pigmentosa (M.A. Wild, cited by Jordan, 1974) and the chloroplasts of Chlamydomonas (Allet and Rochaix, 1979) have no homology with those present in genes for tRNA or mRNA. These three classes of genes are

transcribed by different RNA polymerases. Those coding for mRNA are transcribed by Polymerase II, rRNA by Polymerase I, and 5s RNA and tRNA by Polymerase III (Chambon, 1975). The DNA sequencing data would suggest that each RNA polymerase has its own set of splicing enzymes. These enzymes behave in a similar way in different species, for example when cloned yeast tRNA genes or SV40 are injected into X.laevis oocytes both are correctly transcribed and spliced (De Robertis and Olsen, 1979; De Robertis and Mertz, 1977).

This suggests that the splicing of rRNA in different species will occur in a similar fashion. Ribosomal RNA precursors containing intron and exon sequences have been detected in the mitochondria of yeast (Van Ommen et al., 1979) and a temperature sensitive mutant of N.crassa (Mannela et al., 1979). They have also been found in Tetrahymena, where the splicing process is unusually fast. In one species, T.thermophila the half life of the precursor is less than 1 minute. In another, T.pigmentosa the precursor could not be detected and the splicing might occur before the termination of transcription (Din et al., 1979). This is in contrast to the β globin transcript, where the half life of the 15s precursor is less than five minutes (Curtis et al., 1977). This also suggests that rRNA has its own splicing enzymes.

Significant amounts of unspliced transcripts have not been detected in Drosophila. The reason for this could be that they have a very short half life, or are spliced during transcription. The data of Long and Dawid (1979b) suggest that this is unlikely, and so 26s genes in Drosophila with insertions are probably non-functional. However it cannot be excluded that they are transcribed in certain tissues, or that they are used to produce another form of RNA.

9. Summary

A large proportion of the 26s rDNA genes of D.melanogaster are split into two coding regions by an insertion. There are two non-homologous classes of insertion - Type I and Type II. Type I insertions are found only in the X chromosome, where they interrupt 50% of the rDNA units. It is likely that these units are not used to produce mature rRNA in the majority of tissues. Although the rDNA units with Type I insertions are non-functional, the coding sequences have not undergone sequence divergence from uninterrupted units. Significant divergence is also not seen in other higher Diptera with rDNA insertions partially homologous to those of D.melanogaster. This suggests that interrupted units do (or once had) some function. Sequences homologous to the major class of insertion have also been found linked to non-rDNA sequences. The characterisation of these sequences may help in the understanding of the function of rDNA insertions. The isolation and characterisation of a cloned segment of these sequences is described in the following chapters.

C H A P T E R II

MATERIALS AND METHODS

1. DNA Preparation

Bacteria carrying cloned segments of Drosophila DNA were propagated under Catagory II containment conditions as specified by GMAG. Plasmid DNA was prepared from E.coli as previously described (Wensink et al., 1974). In some cases the cleared lysate was deproteinised by phenol extraction, followed by ethanol precipitation of the nucleic acids. The resulting pellet was resuspended in TE_2 containing 1.085g of cesium chloride and 0.1ml of a 10mg/ml solution of ethidium bromide per ml of TE_2. It was then centrifuged to equilibrium and treated as described (Wensink et al., 1974). Deproteinisation of the cleared lysate was also carried out by heating to 65°C for 20 minutes, the precipitated protein was pelleted by spinning in the Sorval SS34 rotor at 8,000 rpm for 10 minutes. The resulting supernatent was treated as described (Wensink et al., 1974).

A rapid method for the preparation of partially pure plasmid DNA from cleared lysates was also used. Cleared lysates from 100ml cultures were prepared as described (Wensink et al., 1974). An unpublished procedure of A. Ramback was then used to partially purify the plasmid DNA.
After the heat treatment described above to deproteinise the cleared lysate, the supernatent was made 10% polyethylene glycol, 0.5M NaCl and kept on ice for 1 hour. The DNA precipitate was then collected by centrifugation in the Sorval SS34 rotor at 8,000 rpm for 20 minutes at 4°C. The DNA pellet was resuspended in TE_2 and was suitable for restriction endonuclease digestion.

Nuclear DNA from D.melanogaster Oregon R embroyos was the gift of D. Glover. Phage λ strain C1857 S 7 DNA was the gift of C. Clayton and M. Scott, and SV40 DNA was the gift of B. Chia, M. Lovett, and P.W.J. Rigby.

2. Enzymes

The restriction endonucleases EcoRI and Kpn I were the gift of A. Atkinson, Bst I was the gift of K. Clarke, Sst I was the gift of L. Woods, M. Lovett and B. Chia. Sal I was the gift of B. Chia and G. Latos. Hind III was the gift of L. Woods, Xma I the gift of G. Latos, and Hae III and Hinf I was the gift of M. Neuberger. Sma I the isoschizomer of Xma I was purchased from New England Biolabs Limited. References for these restriction endonucleases are given in the review by Roberts (1978).

E.coli DNA polymerase I was the gift of A. Atkinson and the T4 DNA Polynucleotide Ligase was the gift of R. Hallewell and also of K. Murray. They were prepared as described by Jovin (1969), Weiss et al. (1968), and Murray et al. (1979) respectively.

3. Restriction Endonuclease Digestion Conditions

The following digestion buffers were used for the indicated enzymes

EcoRI and Bst I; 0.1M Tris pH 7.6, 10mM $MgCl_2$.

Hae III, Hinf I and Hind III; 0.1M Tris pH 7.6, 10mM $MgCl_2$, 50mM NaCl, 1mM DTT.

Xma I and Kpn I, 6mM Tris pH 7.6, 6mM $MgCl_2$, 6mM β Mercaptoethanol.

Sst I; 6mM Tris pH 7.6, 6mM $MgCl_2$, 6mM β Mercaptoethanol, 100 μg/ml gelatin.

Sal I; 6mM Tris pH 7.6, 6mM $MgCl_2$, 6mM β Mercaptoethanol, 100mM NaCl.

Sma I; 6mM Tris pH 8.0, 6mM $MgCl_2$, 6mM β Mercaptoethanol, 20mM KCl, 100μg/ml BSA.

All the digestions were carried out at 37°C. Where the DNA sample was to be digested with two or more enzymes with differing digestion conditions, the enzyme with the lower ionic strength buffer was used first. After the appropriate amount of time the ionic strength was increased and the second enzyme added and the digestion continued. The digestions were stopped either by heating to 80°C for 10 minutes, or by the addition of 1/5th volume of gel loading buffer (20% ficoll, 25mM EDTA pH 7.5, 0.5% wt/vol Orange G tracking dye) and electrophoresis.

4. In Vitro Radiolabelling of DNA

DNA was radiolabelled with ^{32}P *in vitro* using the nick translation reaction of DNA polymerase I using previously described conditions (Rigby et al., 1977). Depending on the required specific activity of the DNA, either one or two ($^{32}P\alpha$) deoxytriphosphates (Amersham; ≥ 350 Ci/mmole) were included in the reaction at a specific activity of 200Ci/mmole. The reaction was terminated by heating to 65°C for 10 minutes and unincorporated ($^{32}P\alpha$) deoxytriphosphates removed by passage over a sephadex G 50M column equilibrated in TNE_2. The specific activity of the DNA was normally between 10^7 and 10^8 cpm/µg, depending on the specific activities of the radioactive substrates in the nick translation reaction.

5. Gel Electrophoresis

For both agarose and acrylamide gels, the gel and running buffer was 89mM Tris pH 8.0, 89mM Boric acid and 2.5mM EDTA (TB buffer), the agarose gels normally contained 1µg/ml ethidium bromide. Agarose gels of between 0.5% and 1% in concentration were electrophoresed at 1.5V/cm for 13 hours, if the agarose concentration was greater than 1%, electrophoresis was for 6 hours at 2.5V/cm.

Linnear 2.5% to 7.5% wt/vol gradient polyacrylamide gels were made with the assistance of M. Lovett from a stock of 30% wt/vol acrylamide:0.8% bis acrylamide. The 7.5% acrylamide solution was made 0.15% TEMED, 0.06% ammonium persulphate and 20% glycerol, while the 2.5% acrylamid solution was made 0.15% TEMED and 0.1% ammonium persulphate before pouring. The restriction digestions for gradient polyacrylamide gel electrophoresis were ethanol precipitated, washed in 70% ethanol, and resuspended in 10µl of TE_2 . An equal volume of 60% sucrose, 1.5% ficoll, 25mM EDTA pH 7.5 and 0.2% wt/vol bromophenol blue was added and the sample loaded on the gel, which was then electrophoresed at 1V/cm for 12 hours.

Gels which did not ethidium bromide were stained in a 1µg/ml solution of the dye in electrophoresis buffer for 1 hour, and then destained in 5mM EDTA pH 7.5 for 30 minutes. The DNA in the gel was visualised under UV irradiation and photographed using an orange filter.

6. Transfer of DNA from Agarose Gels to Nitrocellulose and Hybridisation with ^{32}P DNA

DNA was transferred from agarose gels to nitrocellulose essentially as described by Southern (1975), except with the following modifications. The DNA in the gel was denatured in 0.5M NaOH, 1.5M NaCl for approximately 4 to 6 hours, and then prior to transfer the gel was neutralised in 1.5M Tris pH 7.4, 3M NaCl for 1 hour. Before commencing the transfer of DNA on to it the nitrocellulose was rinsed several times in distilled water and once in 2x SSC. The transfer was carried out in 20 x SSC for 48 hours, after which the nitrocellulose was rinsed extensively in 2 x SSC and baked at 80°C for at least 2 hours.

A modification of the procedure described by Jeffreys and Flavell, (1977) was used for the hybridisation of nick translated ^{32}P DNA to

restriction fragments transferred onto nitrocellulose filters. All the following procedures were carried out at 65°C. The nitrocellulose filters were first preincubated for 3 hours in 3 x SSC 10 x PM (PM is Denhardts (1966) solution) 0.1% SDS, and then for a further hour after the addition of denatured sonicated salmon sperm DNA to 100µg/ml. The hybridisation was in the above solution for approximately 16 hours with 10 to 20 ng/ml of denatured nick translated ^{32}P DNA except that 5 x SSC was used. The nitrocellulose sheets were washed 4 x 30 minutes in 2 x SSC, 5 x PM 0.1% SDS, 0.1% tetrasodium pyrophosphate and 100µg/ml denatured sonicated salmon sperm DNA, then 6 x 30 minutes in 0.2x SSC, 0.1% SDS, and finally 2 x 20 minutes at room temperature in 3mM Tris base. After drying at 65°C they were autoradiographed at -70°C, using preflashed Fuji RX medical X-ray film (Laskey and Miles, 1976).

7. Binding of DNA to DBM Paper Disks and Hybridisation with ^{32}P DNA

Prior to being covalently bound to DBM paper the DNA was partially degraded by boiling in 0.1M NaOH for 20 minutes, then HCl and Tris pH 8.0 was added to 0.1M and 50mM respectively, the DNA was then ethanol precipitated and washed once with 70% ethanol. Between 4 and 8µg of the DNA was then covalently bound to 9mm disks of DBM paper as described (Goldberg et al., 1979).

The procedure for hybridisation of nick translated ^{32}P DNA to the DNA fixed on the DBM paper as is described in Section 6 with the following exceptions. The preincubation and hybridisations were carried out at 55°C, and the washes at 50°C, these did not contain the 0.1% tetrasodium pyrophosophate, also the 3mM Tris base washes at room temperature were omitted. The final wash was in 15mM NaCl, 5mM Tris pH 8.0 and 2mM EDTA (melt buffer) at 50°C. The disks were stored in this buffer at 4°C after the final wash.

8. Subcloning of Restriction Fragments

The bacterial vector chosen for the subcloning of restrcition fragments was pBR322, which contains genes for ampicillin and tetracycline resistance (Bolivar et al., 1977). In general pBR322 and the plasmid from which the fragment was to be subcloned were digested with the appropriate restriction endonucleases. When the digestion was complete the enzymes were inactivated by heating to 80°C for 10 minutes. Then 0.1µg of cleaved pBR322 and 5µg restricted plasmid were mixed and heated to 65°C for 5 minutes and quickly cooled on ice and made 100 mM Tris pH 7.6, 10mM $MgCl_2$, 1mM ATP, 15mM DTT and 1mM spermine, with the DNA concentration at between 50 and 100µg/ml. T4 DNA ligase was added and the mixture incubated for 1 hour at 15°C, after which it was diluted with the above buffer to a DNA concentration of 5 to 10 µg/ml and more T4 ligase added, the incubation was continued for a further 15 to 20 hours. The DNA was then ethanol precipitated and used to transform E.coli strain HB101 as described (Mandel and Higa, 1970) to ampicillin resistance.

Where the subcloned restriction fragment was inserted into a restriction site in the tetracycline resistance gene of pBR322, bacteria carrying the recombinant plasmids were ampicillin resistant tetracycline sensitive. These were isolated by picking the ampicillin resistant transformants into L. plates containing 100µg/ml ampicillin, 25µg/ml tetracycline, those that did not grow were tetracycline sensitive and contained recombinant plasmids. If the subcloned restriction fragment was inserted into the Eco RI site of pBR322, recombinant plasmids were identified by the colony hybridisation procedure.

9. Colony Hybridisation

Two procedures for colony hybridisation were used, in both cases colonies were picked onto a grid. In one a modification of the Grunstein and Hogness (1975) procedure was used. The colonies were replica plated onto a nitrocellulose disk placed on an L. plate, the colonies were then grown up overnight on the nitrocellulose disk. This was then removed and placed on Whatman 3mm paper soaked in 0.5M NaOH, 1.5M NaCl for 2 minutes, then neutralised in a similar way for 2 minutes in 1M Tris pH 7.4, 3M NaCl. The colonies were finally fixed to the nitrocellulose by washing under suction in 95% ethanol, after rinsing in 2 x SSC the disks were baked at 80^{o}C for at least 2 hours. In the other procedure a disk of Whatman 541 paper was placed over the grid of colonies. It was then removed and placed on a fresh L. plate and the colonies grown up for 2 to 4 hours, after which the disk was removed and treated as the nitrocellulose disks but omitting the 95% ethanol washes. The procedure used for hybridisation of nick translated ^{32}P DNA to colonies bound on either nitrocellulose or Whatman 541 paper disks is the same as described in Section 6.

10. Plaque Hybridisation

A pool of recombinant phage λ containing D.melanogaster Oregon R DNA was provided by D.S. Hogness and constructed by L. Prestidge, D. Kemp and E. Meyerowitz. The pool was constructed by inserting approximately 20Kb segments of randomly sheared Drosophila DNA between the left and right arms of the vector λ Sep 6 by dG:dC tailing. This λ vector has identical arms to λ 647 (Murray et al. 1977) but is red^{-} and contains a spontaneous clear mutation. The recombinant λ phage were packaged in an *in vitro* packaging extract generating 5 x 10^{4} phage.

These were banded in an cesium chloride equilibrium gradient and the peak (2 x 10^4 phage) amplified in E.coli strain KH802. The amplified phage pool (λ bDm) was banded again in cesium chloride and divided into high and low density phage pools, which correspond approximately to phage containing large and small inserts of Drosophila DNA respectively.

A total of 2.4 x 10^4 phage from the high density pool were used to infect a lawn of E.coli strain KH802 on 6 L. plates (4 x 10^3 plaques/56.7 cm^2 plate). Four replicas of the plaques on each plate were taken on nitrocellulose disks by the procedure of Benton and Davies (1977). They were denatured in 0.5MNaOH, 1.5M NaCl for 30 seconds and neutralised for the same time in 1M Tris pH 7.4, 3M NaCl. After rinsing in 2 x SSC they were baked at 80°C for 2 hours. The procedure for hybridisation of nick translated ^{32}P DNA to the plaques is the same as in Section 6, except that 0.1% tetrasodium pyrophosphate was included in the preincubation and hybridisation solutions, and also in all the 65°C washes.

11. Cesium Chloride, Actinomycin D Equilibrium Density Gradients

A total of 520µg of D.melanogaster DNA was cleaved with EcoRI, phenol extracted and ethanol precipitated. The DNA was resuspended in 25ml of TE_2 containing cesium chloride at a density of 1.66 g/cc and 30µg of Actinomycin D per 100µg of DNA. After centrifugation in the Beckman 60 ti rotor at 31,500 rpm at 20°C for 72 hours, 1ml fractions were collected. The main peak was located by spotting 10µl aliquots of the fractions on agarose gels containing 1µg/ml ethidium bromide. Fractions containing the main peak were extracted with cesium chloride saturated isopropanol to remove the Actinomycin D, concentrated by dialysis against polyethylene glycol and then dialysed against TE_2. The DNA concentration of each fraction was determined by its OD_{260}

reading.

12. Purification of Restriction Fragments

The main procedure used is as follows. After fractionation by agarose gel electrophoresis, restriction fragments were located by ethidium bromide staining and irradiation with long wave UV. The segment of the gel containing the desired restriction fragment was cut out and broken up by syringing through a 23 gauge needle into 2mls of 10x TNE_2 buffer and 2mls of TE_2 saturated phenol. This was carried out several times after which the solution was incubated with shaking for 1 hour at 65°C. The phases were separated by centrifugation for 10 minutes at 8,000 rpm at 4°C in the Sorval HB4 rotor. The aqueous phase was removed and replaced with 1ml 10 x TNE_2 buffer and the above procedure repeated. The DNA in the pooled aqueous phases was ethanol precipitated, and after washing once with 70% ethanol was resuspended in distilled water and was then suitable for labelling in vitro by nick translation.

Restriction fragments for nick translation were also fractionated by sucrose gradient centrifugation. The restriction endonuclease cleaved DNA was layered on a 27ml 10 to 30% wt/vol sucrose gradient in TNE_2 buffer and centrifugated at 25,000 rpm for 12 hours at 20°C in the Beckmann SW27 rotor, after which 1 ml fractions were collected. An aliquot of each fraction was then analysed by agarose gel electrophoresis, and fractions containing the desired fragment pooled and ethanol precipitated. After washing in 70% ethanol it was resuspended in distilled water and used for nick translation.

Restriction fragments for restriction mapping by digestion with a second enzyme were purified by using LGT agarose. Restriction endonuclease cleaved DNA was electrophoretically fractionated in an

low gelling temperature (LGT) agarose gel (LGT agarose from Miles) using TB buffer. Restriction fragments were located by ethidium bromide staining and long wave length UV irradiation and the gel slice containing them cut out. This was then melted at 65°C, cooled to 37°C and $MgCl_2$ added to 20mM and an excess of the second restriction endonuclease added. The digestion was carried out for 1 hour at 37°C then fractionated on a second normal agarose gel.

13. Phenol Extractions and Ethanol Precipitations

In general the phenol extractions described in this Chapter were carried out by the addition of 2 volumes of TE_2 saturated phenol to the aqueous phase. After mixing the phases were separated by centrifugation, and the aqueous layer collected. Ethanol precipitations of DNA were carried out by the addition of 1/5th volume of 2M sodium acetate pH 7.0 and 2.5 volumes of absolute ethanol to the aqueous solution. The solution was then kept either at -20°C for at least 12 hours, or in a dry ice ethanol bath for 20 minutes. The DNA precipitate was then collected by centrifugation. Normally after phenol extraction DNA was ethanol precipitated, where necessary residual phenol was removed by washing the DNA pellet in 70% ethanol, it was then dessicated to dryness and resuspended in TE_2.

14. Length Standards for Gel Electrophoresis

Restriction digests were fractionated on agarose gels ranging in concentration from 0.4 to 2.5%. The latter allowed fragments of $\leq$ 200 bp to be resolved. As size markers for small fragments *Hind* III, *Hae* III, or *Hinf* I digests of SV40 were used. The sizes of the fragments were determined from the SV40 sequences (Fiers et al., 1978).

For large fragments, an EcoRI/Bst I digest of phage λ was used (Haggerty and Schleif, 1976). In this case the size of the λ genome was taken as 46.5 Kb (Davidson and Szybalski, 1971). The sizes of the unknown restriction fragments was determined by comparing their mobility with that of the length standards. This was carried out using a graph drawn of the lenth of the standard restriction fragments (in Kb) plotted against their mobility.

15. Electron Microscopy

This was carried out using the aqueous spreading technique of Davis et al. (1971). SV40 DNA (5.226Kb, Fiers et al., (1978)) was included as a length standard in the spreading mixture. They were examined in a Philips EM301 microscope. The length measurements were made by projecting 35mm micrographs onto a Hewlett Packard digitiser in line with a 9821 computer.

CHAPTER III

ISOLATION AND CHARACTERISATION OF cDm 219

1. Introduction

The hybridisation of Type I insertion probes to gradient purified non-rDNA Type I sequences (Dawid and Botchan, 1978; Dawid and Wellauer, 1978) indicate that they are too complex to be analysed by this approach. Therefore these sequences were isolated by cloning in bacterial plasmids. This chapter describes the isolation and characterisation of a cloned segment of D.melanogaster DNA which contains five tandemly arranged units homologous to the insertion of Dm103.

During the course of these experiments a variety of recombinant plasmids derived from Dm103 were used. This 17Kb EcoRI fragment was originally isolated cloned in the vector pSC101 where it was named pDm103 (Glover et al., 1975). It was then transferred into the vector ColEI, where it was called cDm103 (Grunstein and Hogness, 1975). The vector ColEI is not homologous to pSC101, and so pDm103 can be hybridised to segments of Drosophila DNA cloned in ColEI, as no hybridisation occurs due to the bacterial vectors. Dm103 has also been cloned in the vector pBR322 (pcDm103, see Appendix I) which has slight homology with both pSC101 and ColEI. A diagram of Dm103 and sub-clones derived from it is shown in Figure 5.

2. Isolation of Cloned Non-rDNA Type I Insertion Sequences

As previously described there are two classes of sequence in the D.melanogaster genome homologous to the Type I insertion of Dm103.

Figure 5 Subclones of Dm103

The subclones are indicated below the restriction map of Dm103. The construction of pDm103A, cKDm103/A1, cKDm103/B and cKDm103/D is described by Glover and Hogness (1977). With the exception of pDm103/A the vector is the ColE1-Kan plasmid (PML2), for pDm103/A the vector is pSC101. The construction of pC2, pC25 and pC4 is described in Appendix I, the vector is pBR322.

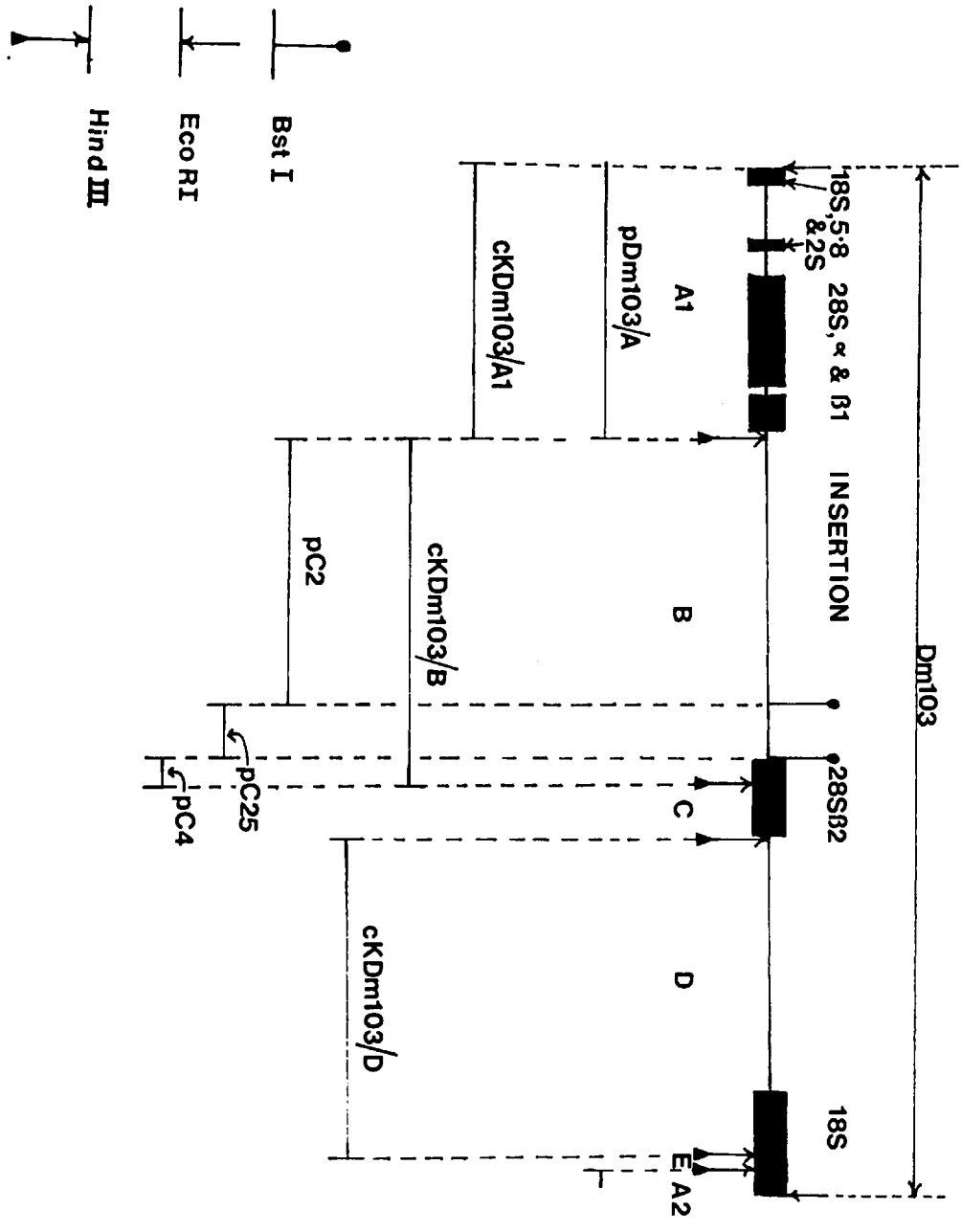

Hind III
Eco RI
Bst I
18S,5·8 &2S
28S,α & β1
INSERT ION
28Sβ2
18S
A1
A2
pDm103/A
cKDm103/A1
pC2
cKDm103/B
pC25
pC4
cKDm103/D
A
B
C
D
E
Dm103

Figure 5.

There are those which occur in the rDNA as Type I insertions, and those which are found outside of it linked to different sequences. This provides the basis for the isolation of cloned segments of Drosophila DNA containing non-rDNA Type I insertion sequences. Bacterial colonies containing cloned D.melanogaster DNA which hybridise to insertion sequences can be differentiated into the two classes by whether they hybridise to an rDNA probe. Those that do not will contain Type I insertion sequences from outside the rDNA.

A cloned segment of non-rDNA Type I sequences was isolated from a pool of segments of D.melanogaster DNA cloned in the vector ColEI. The recombinant plasmids were of two kinds- those where the D.melanogaster DNA was randomly sheared and inserted into the vector by the dA:dT tailing method (Wensink et al., 1974), or those where the DNA had been partially digested with EcoRI and ligated into the vectors EcoRI site using DNA ligase (Glover et al., 1975). These plasmids were constructed at Stanford University by G. Rubin and D. Glover respectively. They were screened for those containing sequences complementary to cRNA transcribed from pDm103. As the two vectors pSC101 (for pDm103) and ColEI are not homologous, only colonies containing plasmids with sequences homologous to Dm103 will hybridise. Because this contains a 17Kb rDNA unit with a Type I inserion, colonies which hybridise to it will contain plasmids with either rDNA sequences or non-rDNA Type I sequences. A total of 144 colonies were isolated which hybridised to Dm103. Of these 135 contained randomly sheared D.melanogaster DNA, and nine had fragments generated by partial EcoRI cleavage.

Colonies containing recombinant plasmids were then screened for non-rDNA Type I sequences by the differential hybridisation of rDNA

and Type I insertion probes. As an rDNA probe pDm103/A was used; it contains fragments A1 and A2 of Dm103 (Figure 5, Glover and Hogness, 1977), which contain 18 and 26s rDNA sequences and approximately 200bp of insertion, cloned in the vector pSC101. The only hybridisation which will occur will be due to Dm103/A since the plasmid vectors do not show homology. The insertion probe contained Hind III fragment B of Dm103 (Dm103/B, see Figure 5). This contains almost all of the 5.3Kb insertion and approximatley 200bp of 26s rDNA sequences (Glover and Hogness, 1977). It was purified from cKDm103/B in order to prevent hybridisation due to the ColEI sequences of the PML2 vector. EcoRI cleaves cKDm103/B twice, generating the 6.5Kb ColEI, plus a 13Kb fragment which includes Dm103/B. This fragment was separated from the 6.5Kb ColEI on a sucrose gradient (material and methods), and recovered for use as the insertion probe. The two probes were radiolabelled in vitro with ^{32}P by nick translation (Rigby et al., 1977) and used for colony hybridisation (Grunstein and Hogness, 1975).

Several colonies hybridised to the insertion probe, but not to the rDNA probe. But when plasmid DNA was prepared from these colonies all but one, cDm219, hybridised to rDNA. Figure 6 shows that cDm219 hybridises to an rDNA unit with a Type I insertion, Dm103, (track 1) but not to an uninterrupted unit, Dm238 (track 2, D. Glover, personal communication). This indicates that cDm219 contains non-rDNA Type I insertion sequences.

The plasmid cDm219 is one of those constructed by inserting D.melanogaster DNA into the bacterial vector by the dA:dT tailing method. The length of the inserted segment of Drosophila DNA was determined by measured the contour length of X-irradiated cDm219 in the electron microscope. This gave a length of 32.12 Kb (SD = $\pm$ 0.72 Kb, N =11) which subtracting the length of the ColEI vector (6.54Kb, Glover and

Figure 6 Hybridisation of cDm219 to pDm238/29 and pcDm103

EcoRI cleaved pcDm103 (track 1) and pDm238/29 (track 2) was electroretically fractionated on a 0.6% agarose gel and transferred to nitrocellulose. Lanes marked A are the ethidium bromide stain of the gel, and those marked B are the autoradiograph of ^{32}P labelled cDm219 hybridised to the DNA transferred on to the nitrocellulose. The faster migrating fragment in both tracks is the pBR322 vector.

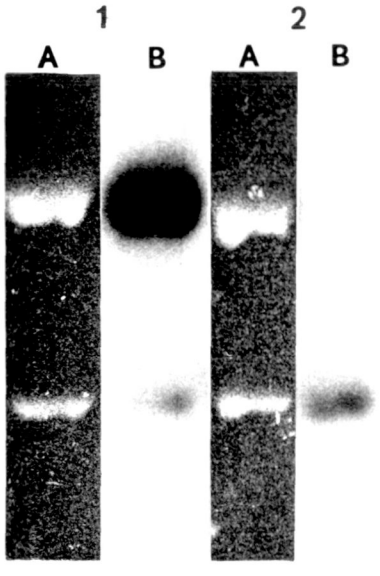


Figure 6.

Hogness, 1977) gives a length of 25.58Kb for the inserted segment of Drosophila DNA (Dm219).

3. Construction of the Restriction Endonuclease Cleavage Map of cDm219

If the Type I insertion sequences of Dm 219 have not significantly diverged from those of Dm103, they can be located by comparison of its restriction map with that of the Dm103 insertion. This section describes the construction of a detailed restriction map of cDm219.

The initial approach to producing a restriction map of cDm219 was to order restriction fragments by single, double and in some cases triple and partial digestions of cDm219 with a variety of restriction endonucleases. However, the sum of the sizes of the restriction fragments generated by several enzymes did not equal the length of cDm219. Furthermore, densitometric tracings of restriction digests of cDm219 fractionated on agarose gels, and stained with ethidium bromide, indicated that some fragments were present in greater than molar amounts. This suggested that cDm219 was internally repetitious.

This made the production of an unambiguous restriction map of cDm219 difficult. However, it was possible to construct a map for the enzymes Bst I and EcoRI and a partially complete map for Xma I. The construction of the maps for these endonucleases is given in Appendix 2, which included a diagram of the Bst I and EcoRI cleavage sites of cDm219 (Figure 28). Xma I was used to orientate the Bst I/EcoRI double digestion products, but it was not possible at this stage to do the same for the 2.6 and 1.8Kb EcoRI fragments, and the two 4Kb Bst I fragments. The partial Xma I map also allowed the ColEI vector of cDm219 to be positioned on the cleavage map, since it has a single Xma I site 1.35Kb away from its single EcoRI site.

In order to reduce the complexity of the restriction endonuclease cleavage pattern of cDm219, the Bst I, EcoRI and the Bst I/EcoRI double digestion products were subcloned into a bacterial plasmid vector. The vector chosen for this was pBR322 (Bolivar et al., 1977). This is 4.3Kb, and has genes coding for tetracyline and ampicillin resistance, it has single cleavage sites for EcoRI BstI Sal I, Hind III and Pst I. A map of pBR322 is shown in Figure 7. It can be seen that with the exception of EcoRI all the above cleavage sites are in an antibiotic resistance gene. When a restriction fragment is inserted into one of these sites the gene is normally inactivated. This provides a quick preliminary screen for recombinant plasmids. For example insertion of a Bst I fragment into the Bst I site of the tetracycline resistance gene of pBR322 makes bacteria carrying the recombinant plasmid sensitive to tetracycline. These can be isolated by screening for bacteria which are ampicillin resistant and tetracycline sensitive.

This method was used as a preliminary screen for recombinant plasmids containing the Bst I and Bst I/EcoRI fragments of cDm219. Plasmids carrying the EcoRI fragments were identified by the colony hybridisation procedure (Grunstein and Hogness, 1975). The inserted segment of cDm219 was then identified by restriction endonuclease cleavage (Appendix 1).

Digestion of the subcloned fragments of cDm219 with restriction endonucleases gave rise to a far less complex cleavage pattern than that of the parental clone. In these digestions none of the fragments appeared in greater than molar yields. This meant that by a combination of single, double and partial digestions of the subclones their restriction maps could be deduced. By comparison of similar digestions of cDm219, the 2.6 and 1.8Kb EcoRI fragments and the two 4Kb BstI fragments could be ordered. This is shown in Figure 28 of Appendix 2, together with the names of the subclones containing these fragments. The

Figure 7 Restriction map of pBR322

The sizes of various restriction fragments (in Kb) or pBR322 is taken from its DNA sequence (Sutcliffe, 1978), and the positions of the ampicillin and tetracycline resistance genes (shown inside the restriction map) is from Boilvar et al. (1977).

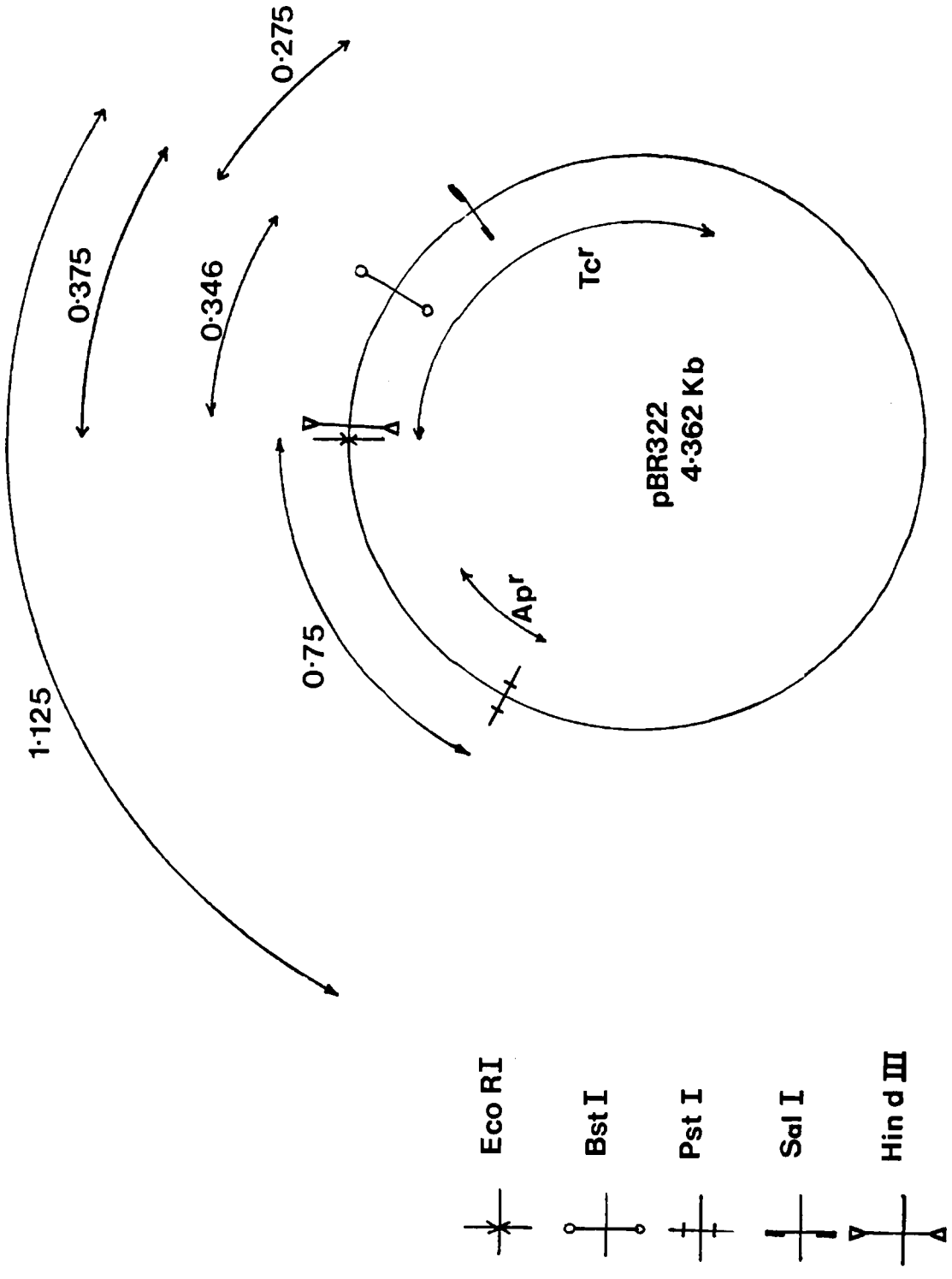

0·275
0·375
0·346
1·125
0·75
Tc^r
Ap^r
pBR322
4·362 Kb
Eco RI
Bst I
Pst I
Sal I
Hin d III

Figure 7.

complete restriction map of cDm219 is shown in Figure 8, its construction is described in Appendix 2. Only the four Hind III sites of cDm219 which correspond to the single site of the Dm103 insertion were mapped. There are additional Hind III sites present in unit 2 of Dm219 (see Figure 8). A more detailed map of the subclone pB74 was produced by mapping its Pst I cleavage sites. This is described in Appendix 2, and is shown in Figure 12.

Also included in Figure 8 is a map of the cleavage sites of cDm103 for the same enzymes as were mapped on cDm219 (Glover and Hogness, 1977; Hind III and EcoRI, B. Jordan and D. Glover; Xma I, and D. Glover; Sal I and Bst I; personal communications), while the Sst I and Kpn I sites were mapped as described in Appendix 2. For some of these mapping experiments subcloned fragments of Dm103 were used. These were; cKDm103/B (Glover and Hogness, 1977), together with the 4.5Kb Hind III/Bst I fragment (pC2) and the 0.88Kb Bst I fragment (pC25) from the insertion, both were cloned in pBR322 (Appendix 1), and are as shown in Figure 5.

4. The Restriction Map of cDm219

Figure 8 shows the restriction maps of cDm219 and cDm103, comparison of the two indicates that a large portion of Dm219 contains Type I insertion sequences. The left hand end of the Dm103 insertion contains a characteristic set of four Sst I sites; representatives of which are also found in repeating arrays with Dm219. At the right hand end of the Dm103 insertion is a set of Bst I and Kpn I sites, these are found incomplete in four regions of Dm219. This allows Dm219 to be divided into five tandemly arranged units of Type I sequences. Three of the units are complete (units 2,3, and 4, of Figure 8) and two are interrupted by the ColEI vector (units 1 and 5, of Figure 8). It can be seen that the units are separated from each other by the presence

Figure 8 Restriction map of cDm103 and cDm219

The map of cDm219 shows the five tandemly arranged units homologous to the insertion of Dm103. The sequence substitution, zeta, of unit 2 is indicated by the wavy line. Identical deletions are found in units 3 and 4. This is shown for unit 3 by the alignment of its restriction map with that of the Dm103 insertion, δ represents the deletion, and γ and ζ are the sequences to either side of it. Some subclones of Dm103 are shown below its restrction map, and all of those of Dm219 are indicated above its map.

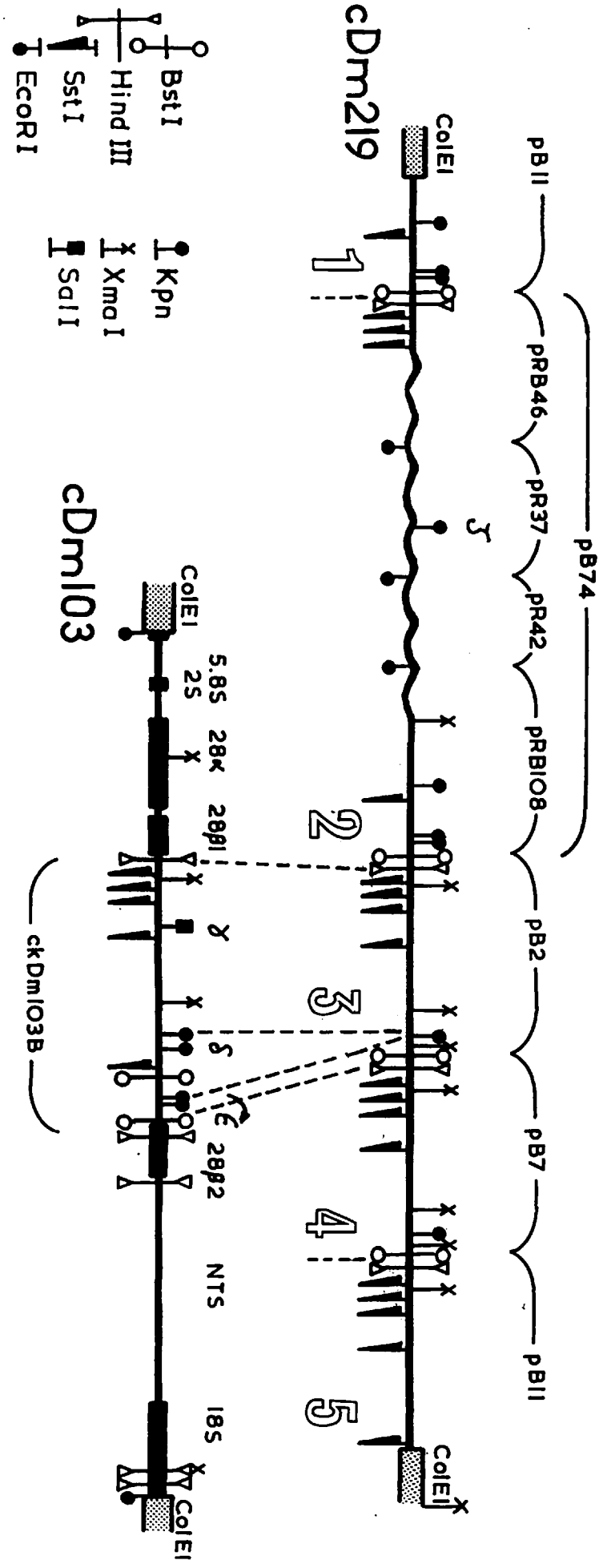

BstI
HindIII
SstI
EcoRI
KpnI
XmaI
SalI
cDm219
ColEI
pB11
pRB46
pR37
pR42
pRBIO8
pB74
pB2
pB7
pB11
ColEI
1
2
3
4
5
cDm103
ColEI
5.8S
25
28α
28β1
α
δ
ε
28β2
NTS
18S
ColEI
ckDm103B

Figure 8.

of a single Bst I site at their left hand ends. This means that each unit is represented by a subcloned Bst I fragment except for the incomplete units 1 and 5, which are found in the clone pB11.

Many of the restriction sites of the Dm103 Type I insertion occur in common arrangements within several of the Dm219 units. Both units 3 and 4 have the 2.6 Kb Xma I fragment found in the left half of the Dm103 insertion (Figure 9,A). The 1.05 and 0.07Kb Kpn I fragments of Dm103 are found in units 1 and 2 (Figure 9,B) while the 0.37Kb Bst I/Kpn I fragment is found in units 3 and 4, as well as units 1 and 2 of Dm219 (Figure 9,C). There are, however, differences in the Sst I cleavage sites of Dm219 which are not apparent in Figure 8. Figure 10 shows Sst I digestions of subclones of cDm219 which contain either entire units or their left hand ends fractionated in a high resolution gradient polyacrylamide gel. A Sst I digest of pC2 is also shown, this contains the left hand half of the Dm103 insertion (see Figure 5). It can be seen that all units with intact left hand ends, except for unit 2, have the 0.68Kb Sst I fragment of Dm103 (Sst I fragment F of Dm219). In contrast, of the equivalent fragments in Dm219 (Sst I fragments G to K) to the 0.32 and 0.35 Kb Sst I fragments of Dm103 only one; fragment I of units 2 and 3 comigrates with a fragment in Dm103 (the 0.32 Kb Sst I fragment). The rest vary in size, ranging from 0.355 to 0.29Kb. Only units 4 and 5 of Dm219 have identical Sst I cleavage sites, both giving fragments G and J. Although these small Sst I fragments are heterogeneous in Dm219, they do map in the equivalent positions of the tandem arrays as the 0.35 and 0.32Kb fragments in the Dm103 insertion, this suggests a basic underlying homology in this region of Dm103 and Dm219. This is supported by restriction digests of three other independently isolated cloned rDNA units with full length Type I insertions; cDm206, cDm435 and cDm801.

Figure 9 Restriction endonuclease digests of cDm219 and its subclones

(A) Xma I cleaved pcDml03 (track 1), pB2 (track 2), pB7 (track 3), and cDm219 (track 4) were fractionated on a 0.6% agarose gel. The arrow indicates the 2.6Kb fragment from the insertion of Dml03, which comigrates with fragment D of cDm219.

(B) Tracks 1, 4 and 6 are Kpn I digestions of cDml03, pB74 and cDm219 respectively, tracks 2, 3 and 5 are Kpn I/Bst I digestions of cDml03, pB74 and cDm219 respectively.

(C) Tracks 1 and 6 are Kpn I digestions of cDml03 and pB11 respectively, the remaining tracks are cDml03 (track 2),pB2 (track 3), pB7 (track 4) and pB11 (track 5) all cleaved with Kpn I/Bst I. The digests in (B) and (C) were fractionated on 2.5% agarose gels, the arrows indicate a, the 1.05Kb and c, the 0.07Kb Kpn I fragments and b, the 0.37Kb Kpn I/Bst I fragments of cDml03 which comigrate with fragments of cDm219.

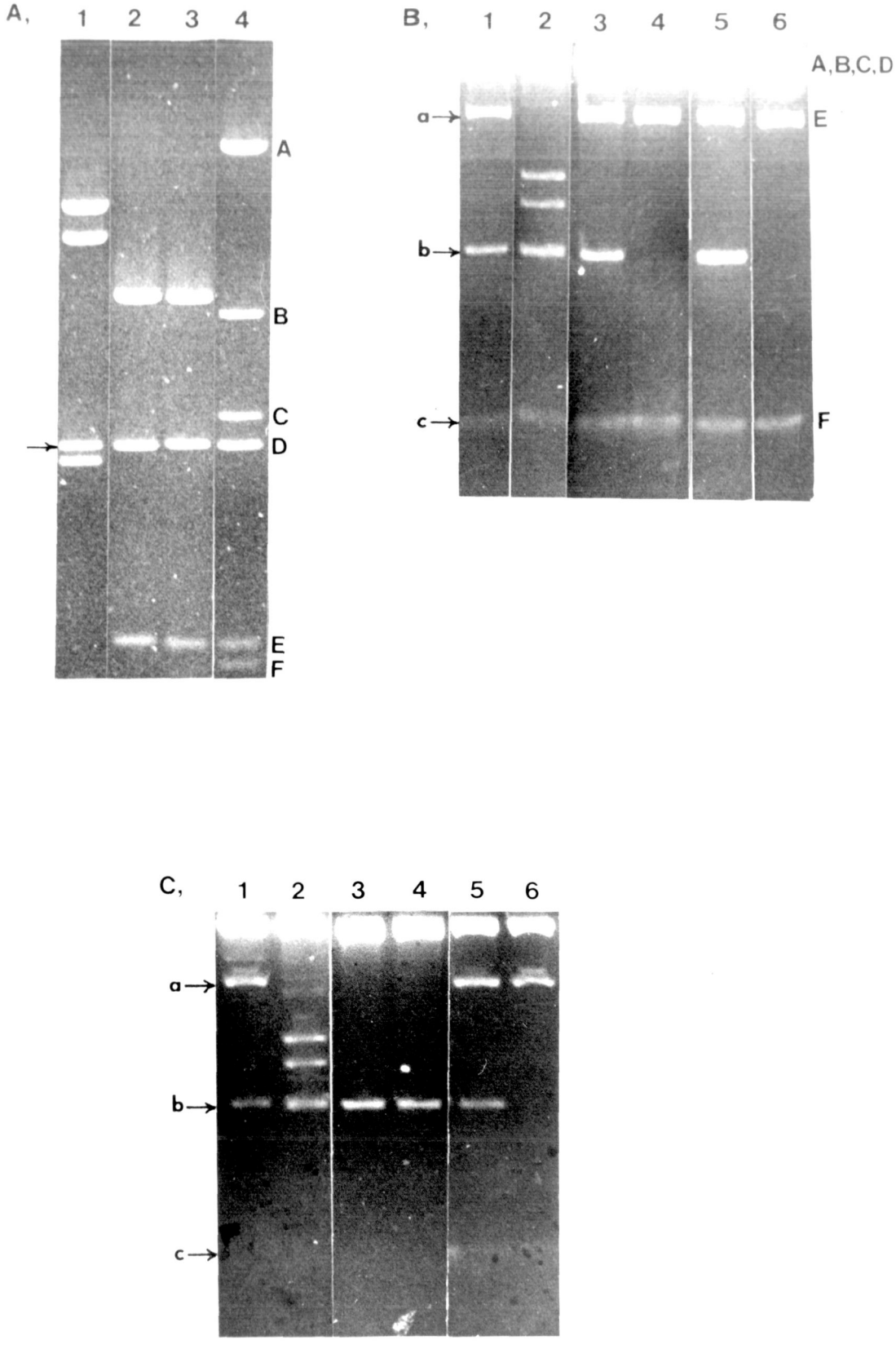


Figure 9.

Figure 10 Sst I digestion of cDm219, its subclones and of rDNA Type I insertions

Sst I digestions of cDm219 (track 1), cDm206 (track 2), pC2 (track 3), cDm219 (track 4), pRB46 (track 5), pB11 (track 6), pB7 (track 7) and pB2 (track 8) fractionated on a 2.5-7.5% gradient polyacrylamide gel (see materials and methods). The letters alongside track 1 refer to the Sst I fragments of cDm219.

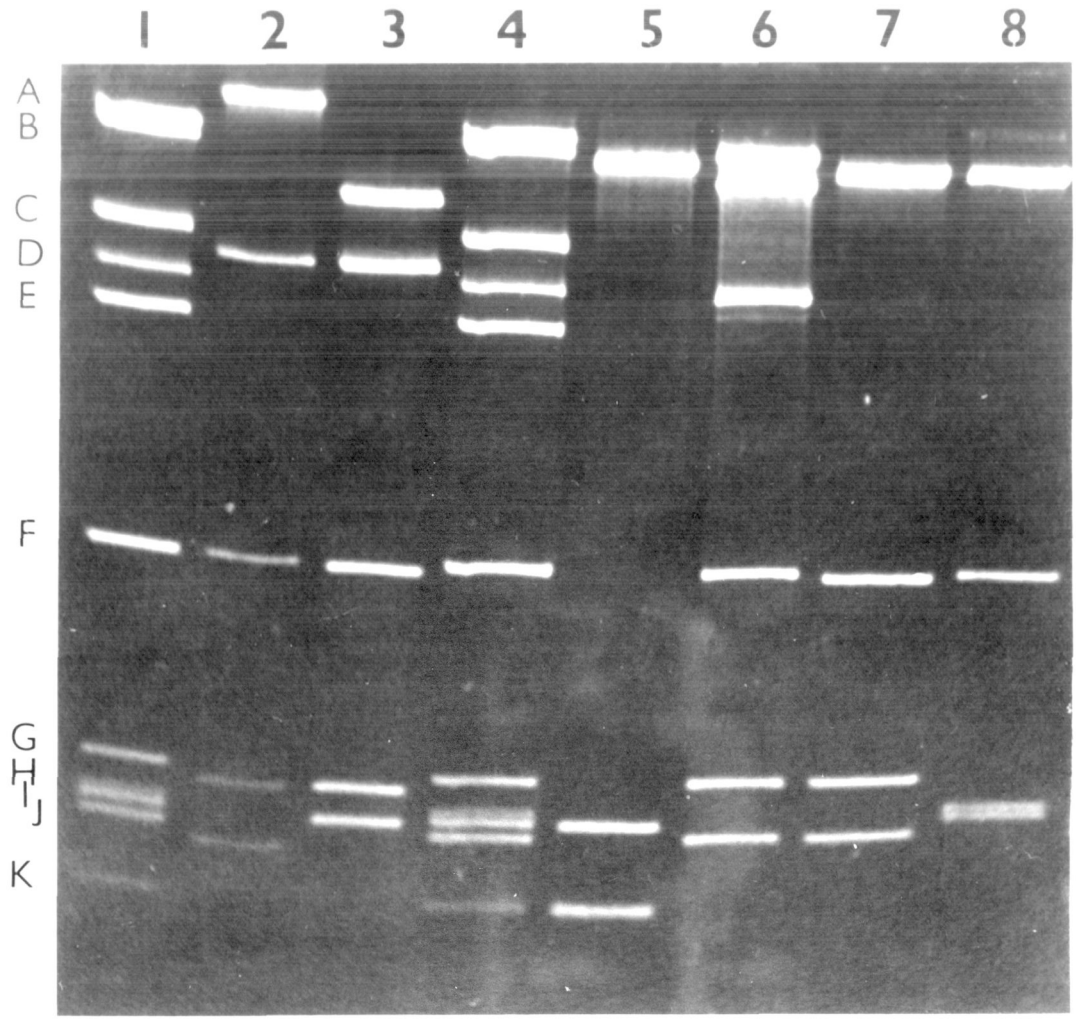


Figure 10.

These were digested with Sst I, Kpn I, Bst I and Sal I. The only difference between these three Type I insertions and that of Dm103 was found in Dm206, where the 0.32Kb Sst I fragment of Dm103 was found to be approximately 20 bp smaller (see tracks 2 and 3 of Figure 10). This indicates that slight variation in these Sst I sites can occur in both rDNA and non-rDNA Type I sequences, although with higher frequency in the latter. Several other differences occur in the restriction maps of Dm103 and Dm219. For example the 0.68Kb Sst I fragment of Dm103 contains a Sal I site, this does not occur in the equivalent fragments of Dm219. Other variations can be seen in Figure 8, only some of them are due to the major sequence variations described in the next section.

5. Major Sequence Differences Between the Units of Dm219 and the Insertion of Dm103

The differences in the restriction maps of the Dm219 units and the insertion of Dm103 described in the previous section, can be explained by single base pair changes. However, the restriction maps of figure 8 also reveals that large sequence alterations can occur between the two Type I sequences.

At the extreme right hand end of the Dm103 insertion there are two Bst I sites, 0.88Kb apart. In the units of Dm219 the left hand Bst I site is missing. In Dm103 this occurs in the 1.05Kb Kpn I fragment, which is also present in units 1 and 2 of Dm219, which suggests the absent Bst I site is due to minor sequence changes. Units 3 and 4 of Dm219 however, lack all but one of the Kpn I sites, and so the 1.05Kb fragment is missing. Digestion of cDm219 with either Hind III, Kpn I or Bst I results in a 4Kb doublet fragment, corresponding to units 3 and 4. As the insertion of Dm103 is 5.4Kb, and units 3 and 4 are

4Kb, it suggests that they have 1.4Kb deletions, which would account for the missing Bst I and Kpn I restriction sites. The region deleted in these units was found by aligning their restriction maps with that of the Dm103 insertion (see Figure 11). The restriction sites between the 2.6Kb Xma I and 0.37Kb Kpn I/Bst I fragments of Dm103 are missing in units 3 and 4, indicating that the deletion occurs between these fragments. The position of the deletion shown in Figure 11 is based on the assumption that before the deletion occurred, units 3 and 4 had similar restriction maps to units 1 and 2. On this basis the region of the Dm103 insertion deleted in units 3 and 4 starts approximately 0.37Kb from the right hand end and extends 1.4Kb to the left, removing the 0.07 and 1.05Kb Kpn I fragments. This suggests that a second Kpn I site should be found in these units (see Figure 11), but if the assumption stated above is correct it was not there in the first place. Confirmation that a deletion has occurred was obtained by the heteroduplex experiments of J. Shepard and D. Glover who localised a 1.1Kb deletion 0.5Kb from the right hand Bst I site of these units.

Unit 2 of Dm219 is unusual because it contains three cleavage sites for EcoRI which does not normally cleave rDNA Type I insertions. This unit is also 11.4Kb in length, which is 6Kb larger than the Dm103 insertion. This suggests that it contains non-insertion sequences, and is confirmed by the experiments described in the following section.

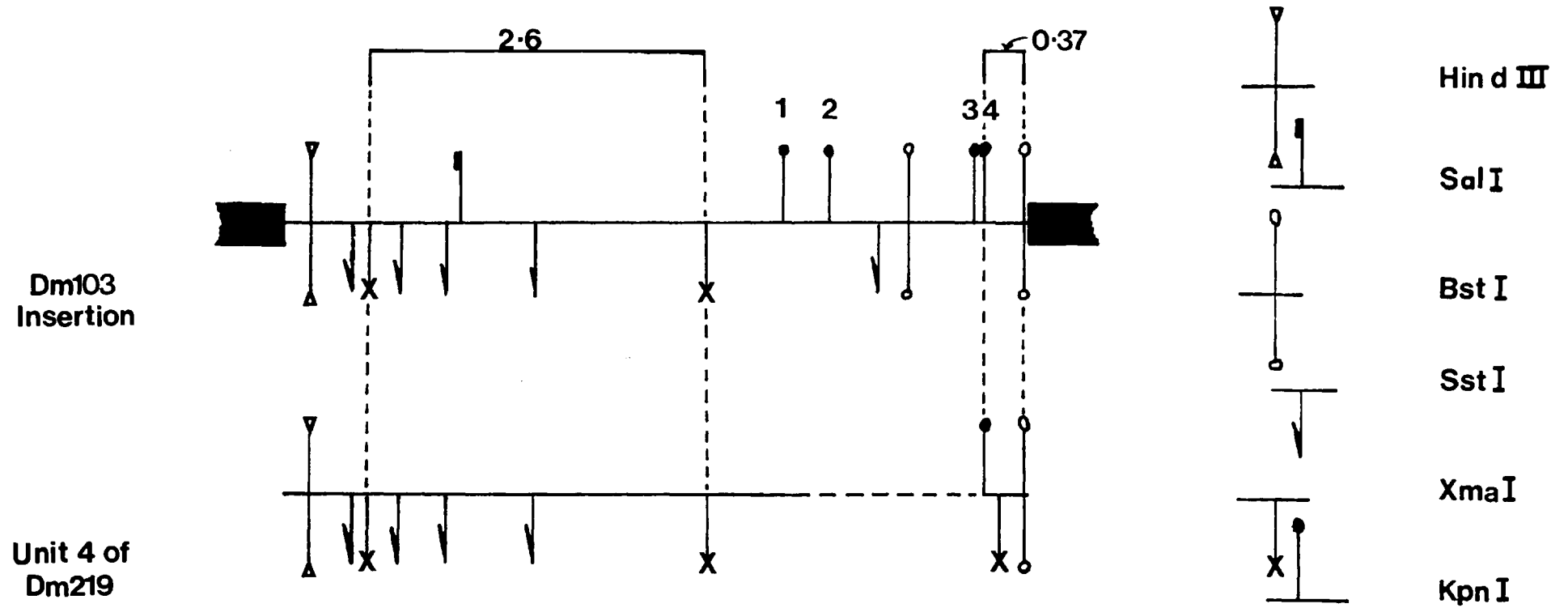
6. The Hybridisation of Type I Insertion Sequences to cDm219

In order to confirm the freatures described above, particularly the presence of non-insertion sequences in unit 2, cDm219 was hybridised to in vitro radiolabelled rDNA Type I insertion probes (Rigby et al., 1977). This was carried out by the procedure of Southern (1975). Restriction endonuclease cleaved cDm219, or its subclones was electro-

Figure 11 Alignment of the restriction maps of unit 4 of Dm219 with the Dm103 insertion

The restriction map of unit 4 of Dm219 has been aligned with that of the Dm103 insertion to show how its 1.4Kb deletion removes the 0.07Kb and 1.05Kb Kpn I fragments. A similar alignment can be made with unit 3 of Dm219.

Figure 11.


2·6
0·37
1
2
3
4
Dm103
Insertion
Unit 4 of
Dm219
Hin d III
Sal I
Bst I
Sst I
Xma I
Kpn I

phoretically fractionated on agarose gels, stained with ethidium bromide for photography, and then the fragments transferred to a nitrocellulose filter. Either radioactive cKDm103/B or cDm103 could be used as insertion probes, as cDm219 does not contain rDNA sequences. The only hybridisation which can occur, will be either due to the ColEI sequences in the probes and the vectors of cDm219 and its subclones, or to the presence of insertion sequences in the Dm219 restriction fragments.

The Non-Insertion Sequences of Unit 2

Figure 12 shows the restriction map of unit 2. In this figure the 2.6 and 1.8Kb EcoRI fragments are referred to by the names of the subclones which contain them (pR37 and pR42 respectively). When Dm103/B is hybridised to an EcoRI/Bst I digest of cDm219 all of the fragments hybridise with the exception of the 2.6 and 1.8Kb EcoRI fragments (Figure 13, track 1). This confirms that the three EcoRI sites of unit 2 are due to a non-insertion sequence, which was named the zeta element. This size of the zeta element was determined by hybridising cDm103 to a variety of other digests of cDm219, or its subclones. This revealed that neither the 2.0Kb Pst I/EcoRI (Figure 13, track 10) or the 1.6Kb Xma I/EcoRI (Figure 13, track 7) fragments to the left and right of the EcoRI sites respectively contain insertion sequences. In contrast the 2.27Kb Sst I/EcoRI fragment to the left, and the 1.9Kb Pst I/EcoRI fragment to the right hybridise to cDm103 (Figure 13, track 4 and 9 respectively). This data is illustrated in Figure 12, where it can be seen that the limits of the zeta element are within the 0.27Kb separating the SstI and Pst I sites at its left end, and within the 0.3Kb between the Xma I and Pst I sites at its right hand end. This makes the zeta element between 8.0 and 8.57 Kb, which when subtracted from the length of unit 2 (11.4Kb), shows that the unit only

Figure 12 Restriction map of unit 2 of Dm219

The fragments indicated above the restriction map do not hybridise to the Dm103 insertion, while those below do show hybridisation. This was used to position the zeta substitution which is shown by the thick line. The fragments and the subclones indicated above this were used as probes for hybridisation to genomic DNA.

Figure 12.

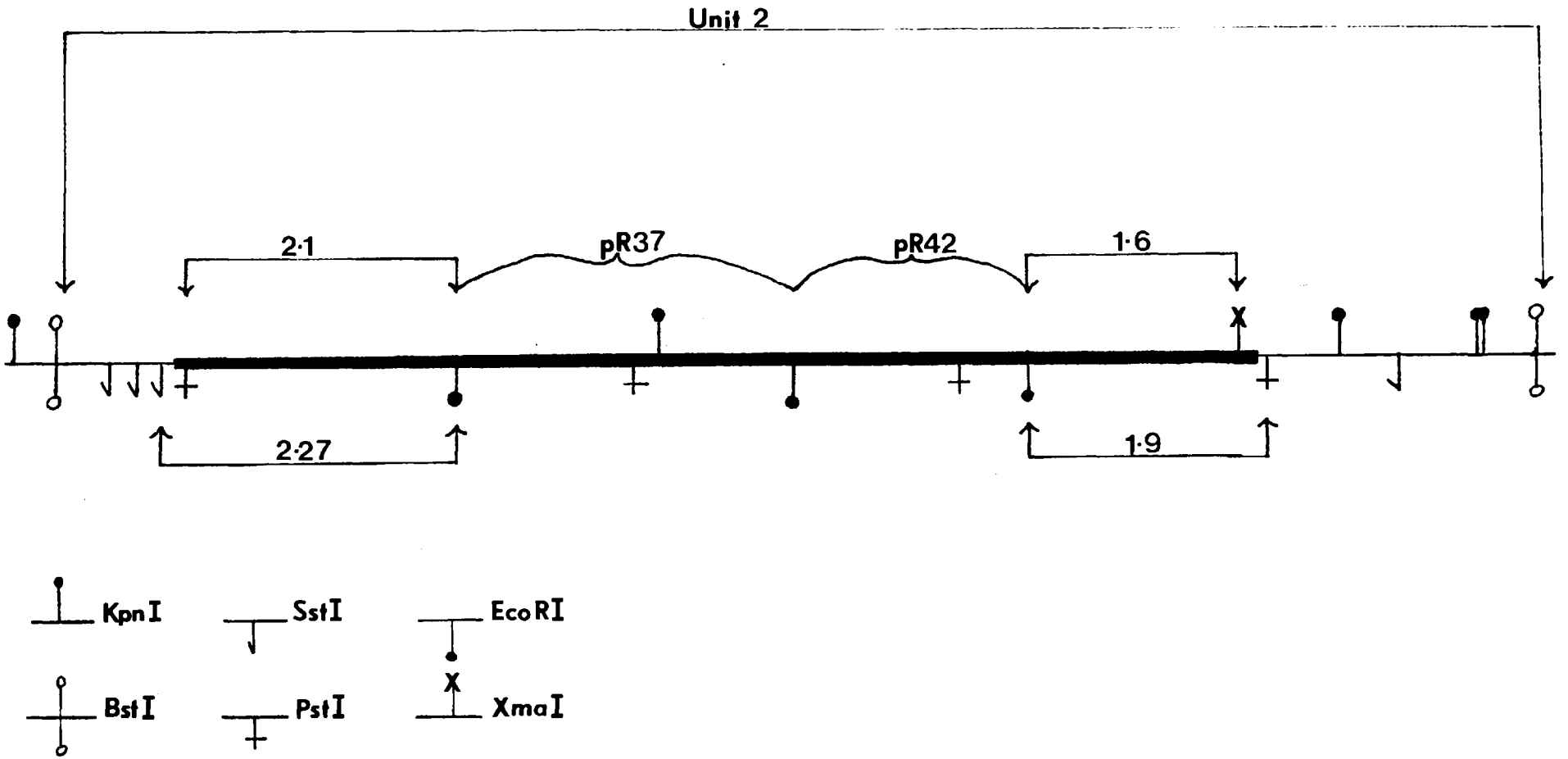


Figure 13 Hybridisation of Type I insertion probes to restriction fragments of Dm219

Hybridisation of ^{32}P labelled cKDm103/B to Bst I/EcoRI cleaved cDm219 (track 1), ^{32}P labelled cDm103 to Sst I cleaved cDm219 (track 2), Bst I/Sst I cleaved pB11 (track 3) and pB2 (track 6), Bst I/EcoRI/Sst I digested pRB46 (track 4) and pRB108 (track 5). Tracks 8, 9, and 10 show the hybridisation of ^{32}P labelled cDm103 to Pst I/EcoRI digests of pB74, pRB108 and pRB46 respectively. The digests of tracks 2 to 6 were fractionated on a 1% agarose gel, the remainder on 0.6% gels. Lanes marked A are the ethidium bromide stained gel, those marked B show the result of transferring the fragments to nitrocellulose and hybridisation with ^{32}P labelled DNA. The sizes (in Kb) of the fragments are indicated alongside tracks 1A, 7A and the two relevant fragments of tracks 8 to 10. The letters alongside track 2 denote the Sst I fragments of cDm219. The position of the pBR322 vector in tracks 3 to 6 is indicated, as is the weakly hybridising 2.27 Kb Sst I/EcoRI fragment in track 4.

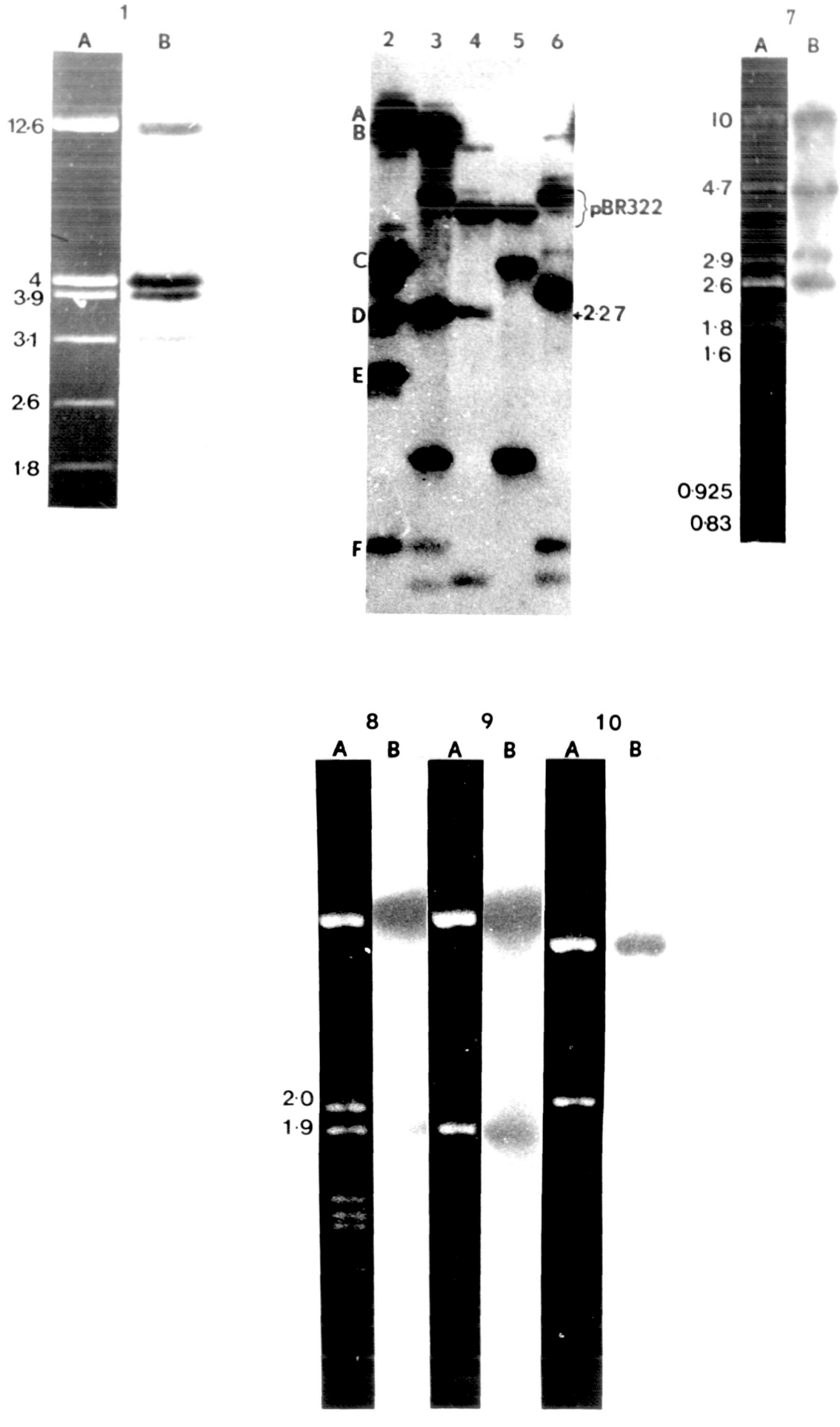

1
A
B
12·6
4
3·9
3·1
2·6
1·8
2
3
4
5
6
A
B
C
D
E
F
pBR322
2·27
7
A
B
10
4·7
2·9
2·6
1·8
1·6
0·925
0·83
8
A
B
9
A
B
10
A
B
2·0
1·9

Figure 13.

contains between 3.4 and 2.8Kb of insertion sequence. This indicates that between 2.0 and 2.57Kb of the Type I sequences are deleted from unit 2. A comparison of the restriction maps of unit 2 and the Dm103 insertion suggest that the deletion extends from just to the right of the equivalent of the third Sst I site of Dm103. This would make the position of the deletion and the zeta element of unit 2 coincident indicating that a substitution of approximately 8.0Kb for 2.0Kb of insertion has occurred. The equivalent position of the unit 2 substitution in Dm103 is 1.1 to 3.6Kb from the left end of the insertion. Similar results were obtained by the heteroduplex experiments of D. Glover, who showed that the zeta element is 8.05Kb, and that 0.95Kb of Type I sequences are deleted in unit 2. In these experiments the substitution was found to lie between 1.2 and 2.37Kb from the left end of the Dm103 insertion.

In some experiments described in the following chapter it was noticed that the 1.6Kb Xma I/Eco RI fragment from the right end of the zeta element hybridised to fragments from the left end of unit 2. A similar result is seen when pRB108 from the right of unit 2 is used as a probe. It hybridises to the 2.27Kb SstI/EcoRI fragment of pRB46 (from the left side of unit 2, see Figure 8 and Figure 14, track 2). The converse also applies; the 2.0KbPstI/EcoRI fragment gel purified from pRB46 hybridises to the 2.8Kb Sst I/EcoRI of pRB108 (Figure 14, track 4).

The probes used in these experiments were either subclones of cDm219, or purified from them, and so the weak hybridisation cannot be due to contamination. The weak homology between the left and right ends of unit 2 most likely resides in the zeta element, as no hybridisation was seen with the 2.0Kb PstI/EcoRI fragment to Dm103.

With the exception noted above the zeta element does not appear to be internally repetitive at the level recognised by the

Figure 14 A repeating sequence at or near the ends of the zeta element

This was detected by cross hybridisation between restriction fragments from either end of the Zeta element. Tracks 1, 2, and 3 show the hybridisation of the 2.0Kb Pst I/EcoRI labelled with ^{32}P to Sst I/EcoRI digestions of pB74, pRB108 and pR46 respectively. The 2.27Kb Sst I/EcoRI fragment which shows homology with the probe is indicated, the high molecular weight fragment in track 3 is a partial digestion product. Tracks 4 and 5 show the hybridisation of ^{32}P labelled pRB108 to Sst I/EcoRI digests of pRB46 and pB74 respectively. All the digests were fractionated on 0.6% gels, lanes marked A are the ethidium bromide stained gel and B after transfer to nitrocellulose and hybridisation with ^{32}P DNA.

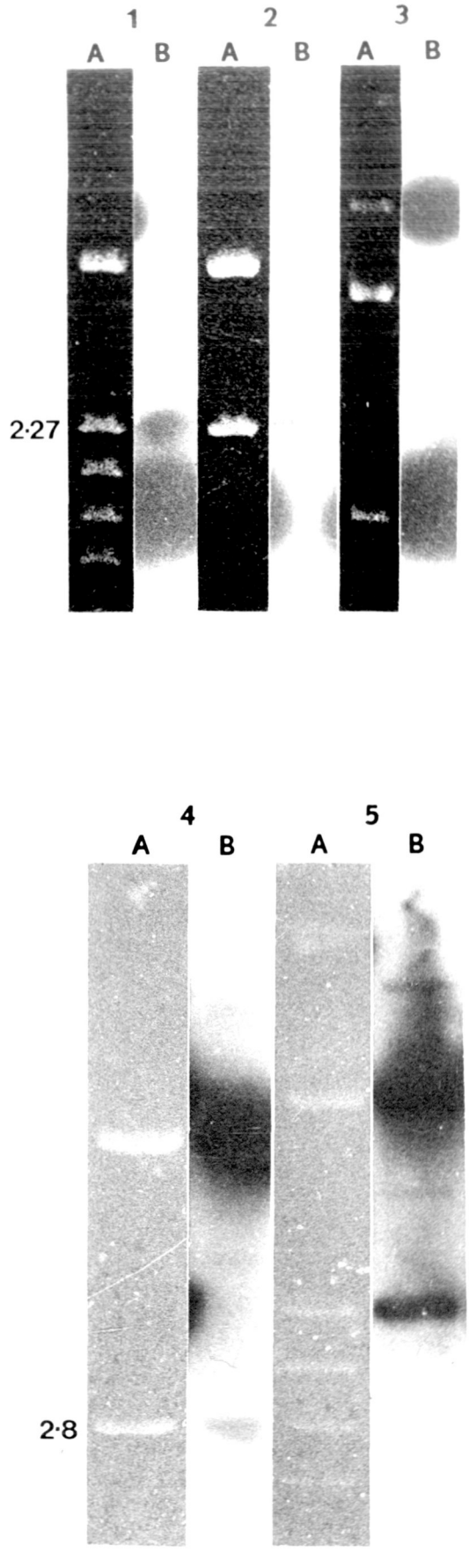


Figure 14.

restriction endonucleases EcoRI, Kpn I, Pst I or Xma I. In addition the two cloned EcoRI fragments of the zeta element do not cross hybridise or show homology with any other region of unit 2.

The Insertion Sequences of cDm219

It can be seen (Figure 13) that restriction fragments from cDm219 or its subclones hybridise to cDm103 or cKDm103/B with an intensity expected of their length and their molar yield. Weak hybridisation occurs if the Dm219 fragment contains part of the zeta element, or if it is due to the partial homology of ColEI with the pBR322 vector of the subclones. This leads to the conclusion that, with the exception of the zeta element, Dm219 is composed entirely of insertion sequences.

None of these experiments, or the heteroduplex experiments of J. Sheppard and D. Glover, addressed the question of the junctions between the tandemly arranged units of Dm219. This is because cKDm103/B which was used to form heteroduplexes with the subcloned Bst I fragments of cDm219 lacks approximately 200 bp of the left end of the insertion. Experiments were therefore carried out to see if cDm219 contained this region, and also the right hand end of the insertion.

Figure 8 shows that the junctions between the units of Dm219 are contained within a 0.29Kb Bst I/Hind III fragment. It can be found in the subclones pRB46, pB2, pB7 and pB11, which represent units 2,3,4 and 5 of Dm219 respectively. In addition pB2, pB7 and pB11 contain a 0.346Kb BstI/HindIII fragment cleaved from the tetracycline resistance gene of the pBR322 vector. This fragment is not found in pRB46 as it is contained in the 0.398Kb Bst I/EcoRI fragment of pBR322 which is substituted by the Dm219 fragment during the cloning procedure. The position of the Bst I and Hind III sites in these subclones is shown in Figure 31 of Appendix 2.

The probes for the left and right ends of the Dm103 insertion

were cKDl03/Al and pC4 respectively (Figure 5). cKDml03/Al includes approximately 200bp of insertion sequences extending from the left of the insertions single Hind III site to the β2 26s rDNA, and almost all of pSCl0l, its construction and structure is described by Glover and Hogness (1977). pSCl0l codes for tetracycline resistance, and so shows homology with the 0.346Kb Bst I/Hind III fragment of pBR322, pC4 contains a 0.32Kb Bst I/Hind III fragment of Dml03 inserted in pBR322 deleting the 0.346Kb Bst I/Hind III fragment of the vector. This 0.32Kb fragment of Dml03 contains at least 200bp of β2 26s rDNA and the sequences extending to the insertions right hand Bst I site.

Figure 15 shows the hybridisation of these two probes to Bst I/ Hind III digests of pBR46, pB2, pB7 and pBll. It can be seen that cKDml03/Al hybridises to the 0.29Kb fragment, but at an intensity which is only about a third of that to the 0.346Kb fragment from pBR322 (Figure 15, lane C of tracks 2 to 5). This suggests that the junction fragments of Dm219 are only partially homologous to the left end of the Dml03 insertion. However, these fragments should also contain sequences from the right end of the insertion and so should also hybridise to pC4. Figure 15 (lane A, of tracks 2 to 5) shows that this does not occur, although pC4 does hybridise back to the 0.32Kb Bst I/Hind III fragment (Figure 15, track 1). A negative result was also obtained when cKDml03/B was used as a probe (data not shown). The 0.29Kb Bst I/Hind III should contain the 200bp of insertion sequences found in Dml03/Al. The level of hybridisation seen supports this conclusion. But as the fragment is only partially homologous to Dml03/Al, the lack of hybridisation to pC4 or Dml03/B is surprising. It is possible that hybrids between the 0.29Kb fragment and pC4 are not stable in the hybridisation and washing conditions used in this experiment (see Materials and Methods). Alternatively the

Figure 15 The junctions between the units of Dm219

The junctions between the units of Dm219 are contained in a 0.29Kb Bst I/Hind III fragment. Bst I/Hind III digests of cKDm103/B (track 1), pB2 (track 2) pB7 (track 3), pB11 (track 4) and pRB46 (track 5) were electrophoresed through a 1.6% agarose gel which was stained with ethidium bromide (lanes marked i or B), the fragments were transferred to nitrocellulose and hybridised with either ^{32}P labelled pC4 (lanes marked ii or A) or cKDm103/A1 (lanes marked C). The 0.32, 0.346 and 0.29Kb Bst I/Hind III fragments from cKDm103/B (cloned in pC4), pBR322 and the subcloned fragments of Dm219 respectively are indicated.

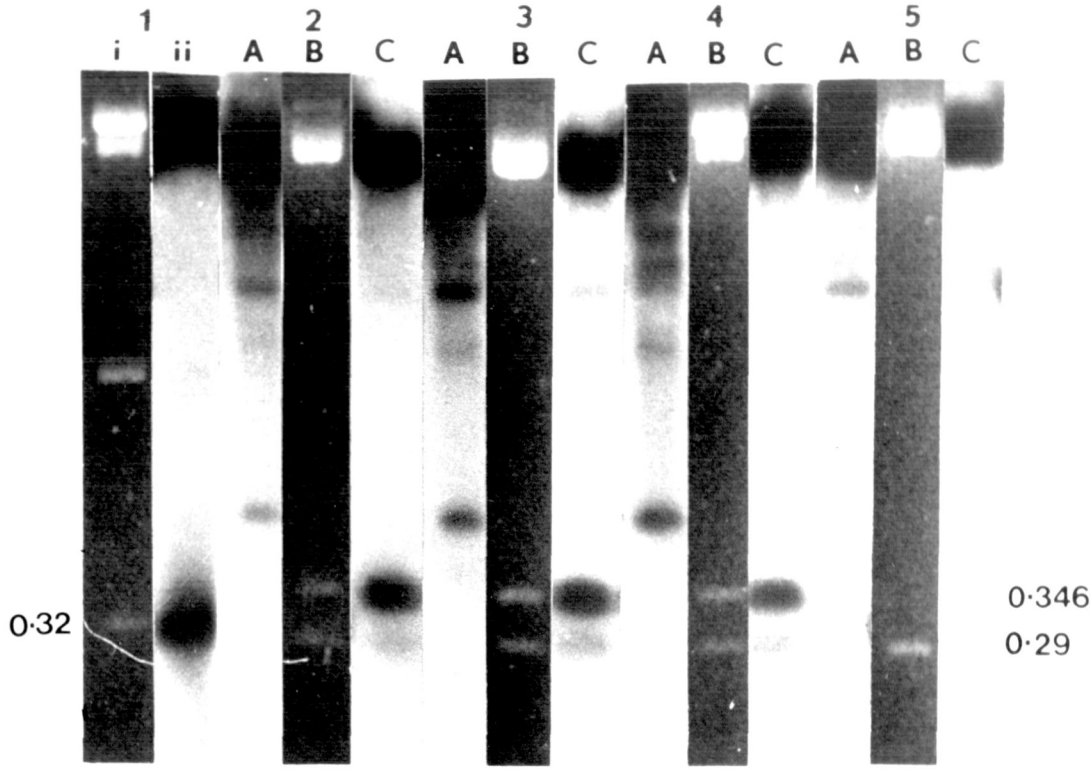


Figure 15.

fragment might not contain sequences homologous to pC4, which would mean that it contains non-insertion sequences. The nature of these junction sequences is best determined by DNA sequencing.

7. Thermal Melting Experiments between Individual Units of Dm219 and the Insertion of Dm103

Comparison of the restriction maps of the tandemly arranged units of Dm219 with the Dm103 insertion indicate that some sequence divergence has taken place. This is too small (less than 100 bp) to be detected in the heteroduplex experiments of J. Sheppard and D. Glover. A more accurate determination of the degree of sequence divergence was obtained by comparing the temperature at which Dm103 homoduplexes melted with that of hybrids between it and units of Dm219. The difference between the two temperatures is a measure of the percentage of mis-matched base pairs in the hybrids.

Plasmids containing either the Dm103 insertion, or various units from Dm219 were covalently fixed to DBM paper (Goldberg et al, in press) and hybridised with a radioactive probe (see Table I). The hybrids and homoduplexes were then melted by raising the temperature through $5^{o}C$ intervals, and the amount of radioactivity still hybridised to the DNA bound on the DBM paper determined. Melting curves (a plot of the amount of radioactivity on the DBM paper versus temperature) resulting from such an experiment are shown in Figure 16, and the results of two experiments are given in Table I. Hybrids between units of Dm219 and the Dm103 insertion have a melting temperature $1.5^{o}C$ to $0.1^{o}C$ lower than Dm103 insertion homoduplexes. It is not clear if these values, particularly the latter, are significant. However it does indicate that the insertion sequences of Dm103 and Dm219 are closely related,

TABLE 1 The Tm of cKDm103/B homoduplexes and hybrids between cKDm103/B and the units of Dm219

Plasmids used to form duplexes with cKDm103/B	Experiment 1		Experiment 2	
	Tm	ΔTm	Tm	ΔTm
cKDm103/B	60.25	0.0	66.25	0.0
pB2	58.75	1.5	65.25	1.0
pB7	58.75	1.5	65.15	0.9
pB74	60.15	0.1	66.35	0.1

In experiment 1 cKDm103/B was bound to disks of DBM paper, while in experiment 2 cKDm103/B, pB2, pB7 and PB74 were each bound to separate disks. The DNA bound to the disk was then hybridised with ^{32}P labelled DNA at a concentration of 2ng/ml, and a specific activity of 9×10^7 cpm/μg. In experiment 1 the ^{32}P DNA was cKDm103/B, pB2, PB7 and pB74, while in experiment 2 it was cKDm103/B. The DBM paper disks were washed as described in materials and methods and then placed together in a beaker containing melt buffer. The melting temperature of the hybrids was then determined by raising the temperature through 5^oC intervals. At each temperature the disks were removed and counted by Cerenkov radiation in a scientillation counter. In experiments 1 and 2 the cpm's remaining on the disks at 75^oC and 80^oC respectively were treated as background and subtracted from the values determined at lower temperatures. The results were then plotted as the percentage of the cpms hybridised to the DNA bound on the disk at 50^oC against temperature. This is shown in Figure 16 for experiment 2. These melting curves were used to determine the Tm (the temperature at which half

the cpm s have been lost from the filter) of the duplexes. The difference between the Tm of cKDm103/B homoduplexes and hybrids between it and pB2, pB7 and pB74 is given by the ΔTm value.

Figure 16 Melting curves between hybrids of the units of Dm219 and the Dm103 insertion

The plasmids pB74, pB2, pB7 and cKDm103/B representing units 2, 3 and 4 of Dm219 and the insertion of Dm103 respectively were fixed to DBM paper and then hybridised with ^{32}P labelled cKDm103/B (materials and methods). The melting curves were constructed as described in the legend to Table I.

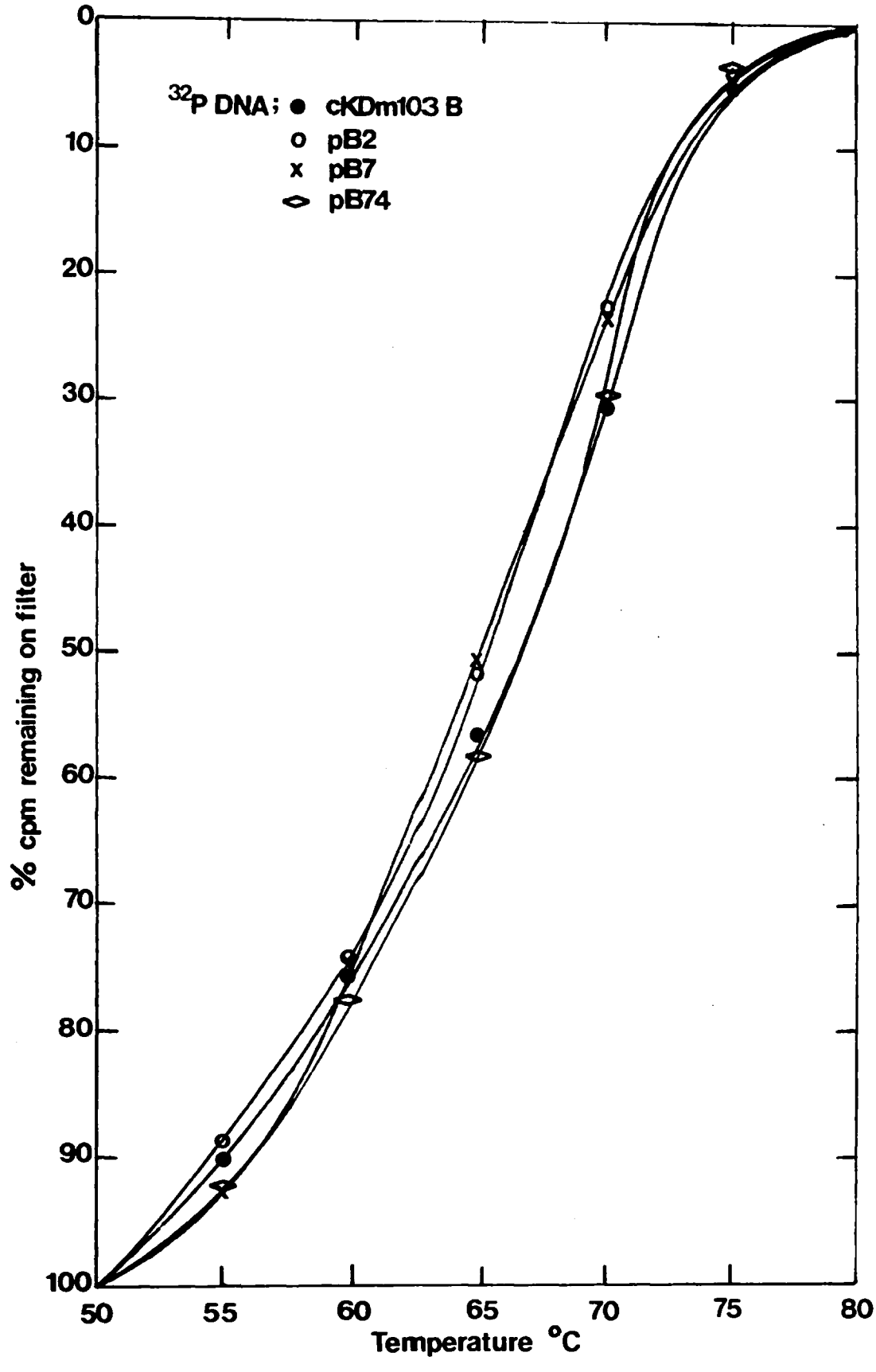

^{32}P DNA; ● cKDm103 B
o pB2
x pB7
◇ pB74
% cpm remaining on filter
0
10
20
30
40
50
60
70
80
90
100
50
55
60
65
70
75
80
Temperature °C

Figure 16.

and have undergone very little sequence divergence (0.1 to 1.5%, assuming 1^{o} equals 1% mismatched base pairs, Britten et al., 1974). Together with the restriction mapping experiments it suggests that non-rDNA Type I sequences have been more free to evolve than Type I rDNA insertions.

CHAPTER IV

THE ARRANGEMENT OF Dm219 SEQUENCES WITHIN THE Drosophila melanogaster GENOME

1. Introduction

The cloned segment of D.melanogaster DNA, Dm219 consists of five tandemly arranged Type I insertion sequences. Two of the units have identical deletions, and another a substitution; the zeta element. Before generalities can be drawn from this sequence arrangement, it is important to know whether it is reflected in the D.melanogaster genome. The technique developed by Southern (1975) can allow the organization of a cloned segment of DNA to be determined within total genomic DNA. This procedure when coupled with fractionating genomic DNA by density was used to show that the sequence arrangement in Dm219 reflects a significant proportion of Type I sequences in the Drosophila genome.

The hybridisation probes used to determine the organization of Dm219 sequences in the D. melanogaster genome were as follows. The probe for rDNA sequences was pDm238/29, this is an 11.6Kb EcoRI fragment containing an uninterrupted rDNA unit cloned in pBR322. The insertion probe was pB7, a subclone of Dm219, which contains a 4Kb Bst I fragment (see Figure 8). The zeta element probes are indicated above the cleavage map of Unit 2 (Figure 12). Two of the probes are the cloned 2.6 and 1.8Kb EcoRI fragments of Dm219 (pR37 and pR42 respectively). Also the 2.0Kb Pst I/EcoRI and 1.6Kb XmaI/EcoRI fragments were purified by gel electrophoresis and used as zeta element probes.

2. Tandem Arrays of Type I Sequences within the D.melanogaster Genome

EcoRI does not cleave Type I insertion sequences, so if tandem arrays such as those of Dm219 exist in the D.melanogaster genome they should

be found in large EcoRI fragments. The problem being to distinguish such fragments from the large EcoRI fragments which hybridise to Type I insertion probes, but are derived from the rDNA. This can be seen in Figure 17 where pB7 has been hybridised to an electrophoretically fractionated EcoRI digest of genomic DNA, which was denatured and transferred to nitrocellulose (Southern, 1975). The insertion probe hybridises to a complex array of fragments including the 17Kb EcoRI fragments representing rDNA units with full length Type I insertions. A different pattern of hybridisation is seen when a rDNA probe (pDm238/29) is used. This is also of high complexity and includes the 17Kb EcoRI fragments and the 11.6Kb fragments derived from contiguous rDNA units.

Use was then made of the observation of Dawid and Botchan (1978) that it is possible to fractionate non-rDNA Type I sequences away from the rDNA on cesium chloride equilibrium gradients containing actinomycin D. This provides the basis for a partial purification of the large EcoRI fragments containing tandem arrays of Type I sequences. The Type I insertions have a G:C content of 55%, and the rRNA coding sequences have a G:C content of 39% (Dawid and Wellauer, 1977). This means that EcoRI fragments with tandem arrays of Type I sequences will have a higher G:C content than the rDNA and bind more actinomycin D so decreasing their density. Thus when EcoRI cleaved genomic DNA is fractionated on cesium chloride - actinomycin D equilibrium gradients the large EcoRI fragments with tandem arrays of Type I sequences will be found in the low density region of the gradient. These can be detected by electrophoresing an aliquot of each gradient fraction on an agarose gel, followed by transfer to a nitrocellulose filter and hybridisation with the insertion probe pB7.

Figure 18A shows the result of such an experiment with EcoRI cleaved D.melanogaster DNA. In fractions from the high density region of the gradient the hybridisation pattern is fairly simple. EcoRI fragments corresponding to rDNA units with full length Type I insertions can be seen,

Figure 17 Hybridisation of Type I and rDNA sequences to EcoRI cleaved D.melanogaster DNA

EcoRI cleaved D.melanogaster DNA was electrophoresed through a 0.6% agarose gel and transferred to nitrocellulose (Southern, 1975). It was then hybridised with ^{32}P labelled pDm 238/29 (track 1, track 2 is a shorter exposure) or pB7 (track 3). The position of the 11.6, 17 (contiguous rDNA units and those with full length Type I insertions respectively) 7.8 and 5.9 Kb EcoRI fragments is indicated. The latter two fragments are produced from rDNA units with Type II insertions.

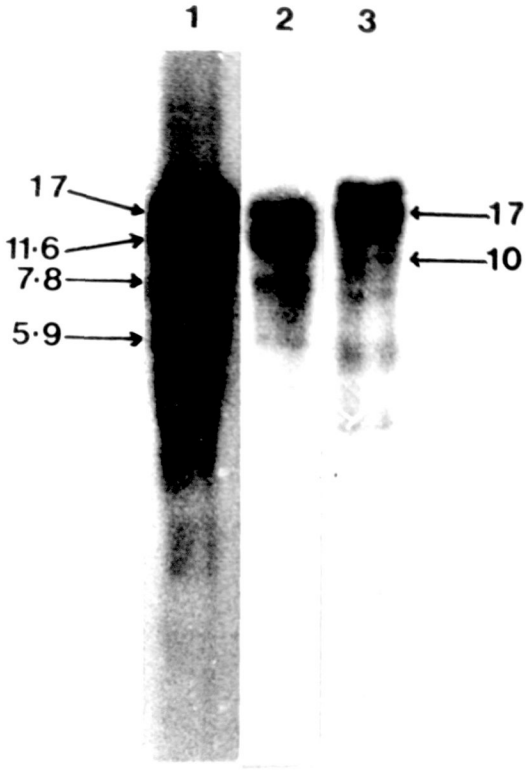


Figure 17.

there are also fainter bands between it and a broad band at 12Kb. These probably correspond to rDNA units with smaller Type I insertions (Dawid and Wellauer, 1978; Wellauer and Dawid, 1978). Towards the centre of the gradient, as the density becomes lower, these fragments decrease in intensity, and are replaced by a more complex hybridisation pattern. This becomes essentially a smear in fractions from the low density region of the gradient (Figure 18(A), tracks 6 to 8), with EcoRI fragments ranging from approximately 25Kb to 11Kb. Over the smear can be seen several strong bands. In tracks 5 and 6 (Figure 18(A)) they are of 10.9 and 9.6Kb. These are probably not derived from the rDNA as they have a higher density than rDNA units with or without Type I insertions, and there are no strongly labelled rDNA EcoRI fragments of this size (Figure 17). They probably represent another form of organisation of non-rDNA Type I sequences, unlike that of Dm219. As expected fractions from the low density region of the gradient contain large EcoRI fragments of at least 25Kb (Figure 18A, tracks 7 and 8).

The structure of Dm219 predicts that if these large EcoRI fragments contain tandem arrays of Type I sequences they should be cleaved to their unit length by restriction endonucleases which cleave them once. To test this an aliquot of each gradient fraction was digested with BstI, electrophoretically fractionated and hybridised to pB7 (Figure 18B). The 17Kb EcoRI rDNA fragments are cleaved twice by Bst I into fragments of 9, 7, and 0.88Kb. Of these pB7 only hybridises to the 9 and 0.88 fragments, which are seen in the high density regions of the gradient (Figure 18B, tracks 1 to 3). Other fragments of 4.1 and 3.5Kb are seen which are most abundant in tracks 5 and 6 (Figure 18B). These are possibly derived from the 10.9 and 9.6Kb EcoRI fragments which occur in the same region of the gradient. Another fragment of 4.7Kb is found, this was expected due to BstI/EcoRI cleavage of rDNA units with 0.5 and 1.0Kb Type I insertions. But its

Figure 18 Fractionation of restriction endonuclease fragments homologous to the Type I insertion on a cesium chloride actinomycin D gradient

EcoRI cleaved D.malanogaster DNA was fractionated on an cesium chloride actinomycin D gradient as described in materials and methods. An aliquot of approximately 0.3µg from each fraction was either electrophoresed on a 0.5% agarose gel for 40 hours at 1.5v/cm (A) or digested with Bst I and fractionated on a 0.6% gel (B). They were then transferred to nitrocellulose and hybridised with ^{32}P labelled pB7. The tracks (1 to 8) are arranged in order of decreasing density (increasing G:C content). The markers in (A) are EcoRI cleaved cDm219 (25 Kb), cDm103 (17 and 6.6Kb) and pBR322 (4.3Kb), while in (B) they are an EcoRI/Bst I digest of cDm219.

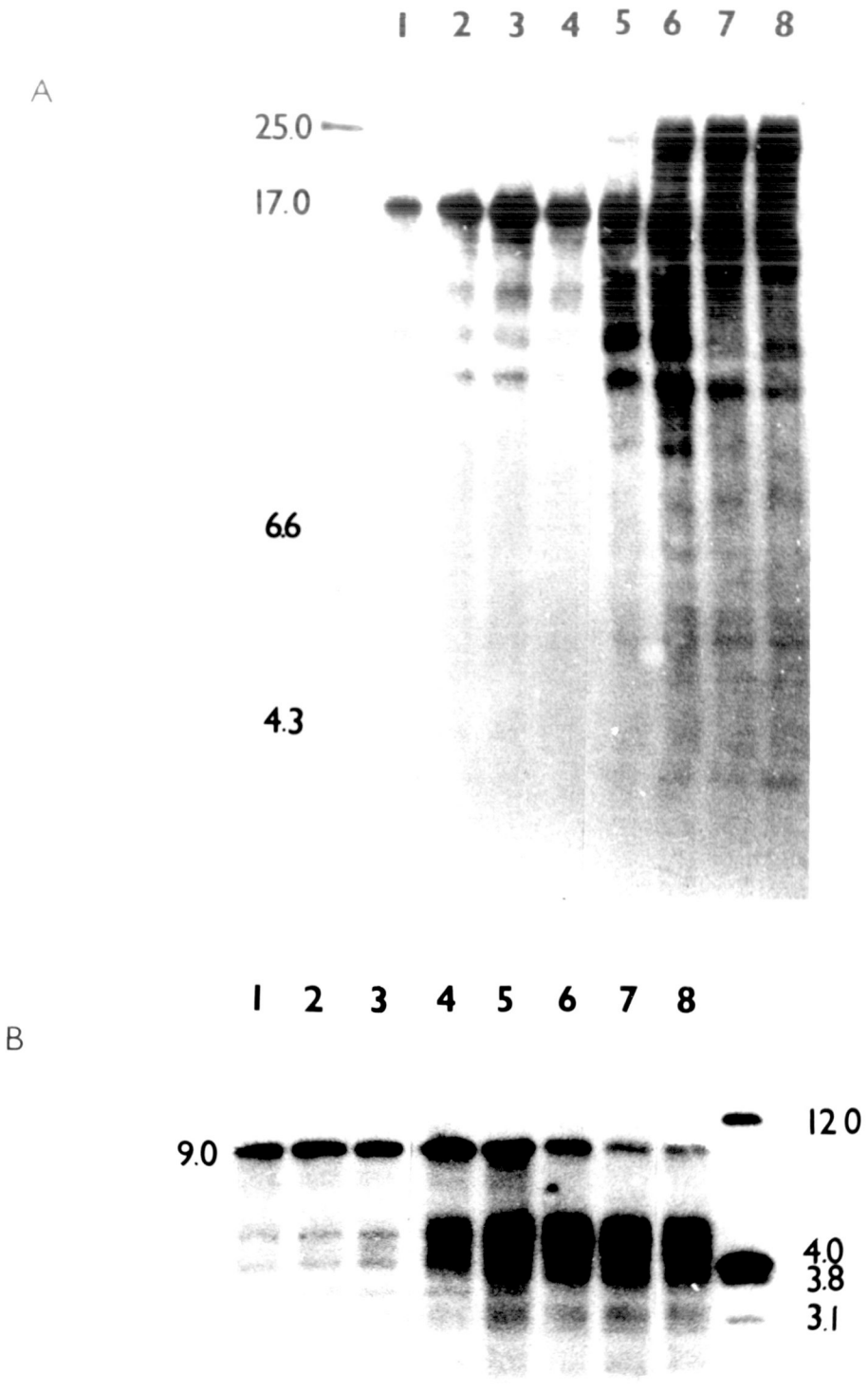


Figure 18.

abundance increases with decreasing density and so is unlikely to be derived from the rDNA. One other intensely hybridising fragment is seen, this is 4.3Kb and is found in the low density region of the gradient, so it is likely to originate from non-rDNA Type I sequences. The most striking result is seen in tracks 7 and 8 (Figure 18(B)) where the complex smear of EcoRI fragments, including the predominate fragment of at least 25Kb, is compacted to an intensely hybridising region of 3.8 to 5.4Kb.

The above results are consistent with the structure of Dm219. Large EcoRI fragments were found with a high G:C content, these appear to contain tandemly arranged units which are heterogeneous in size, their upper size limit being 5.4Kb - the size of the Dm103 insertion. Like Dm219 they appear to lack the 0.88Kb Bst I fragment of Dm103, as this appears to be absent in the low density region of the gradient.

3. The Arrangement of Sequences in the D.melanogaster Genome Homologous to the Zeta Element of Dm219

Introduction

Unit 2 of Dm219 contains a substitution of about 8Kb which was named the Zeta element. Its organization within the D.melanogaster genome was investigated by hybridising zeta element probes to restriction endonuclease cleaved genomic DNA by the technique of Southern (1975).

Hybridisation of pR37 and pR42 to EcoRI Cleaved D.melanogaster DNA

The plasmids pR37 and pR42 contain the 2.6 and 1.8Kb EcoRI fragments respectively, of the zeta element of Dm219 (see Figure 12). When genomic DNA is cleaved with EcoRI and hybridised with pR37, most of the autoradiographic response is from a fragment which comigrates with the 2.6Kb EcoRI fragment of Dm219 (Figure 19, tracks 1 and 2). In contrast, pR42 as well as hybridising to a fragment of similar mobility to the 1.8Kb EcoRI fragment of

Figure 19 Hybridisation of Zeta sequences to restriction endonuclease cleaved Drosophila melanogaster DNA

Total D. melanogaster DNA was cleaved with various restriction endonucleases and then electrophoresed through agarose gels. It was then transferred to nitrocellulose (Southern, 1975) and hybridised with in vitro ^{32}P radiolabelled zeta sequences (Rigby et al., 1977). ^{32}P labelled pR37 was hybridised to EcoRI, Bst I and Kpn I cleaved D.melanogaster DNA (tracks 1, 7 and 9 respectively) and to EcoRI (track 2) and Kpn I (track 10) cleaved cDm219. ^{32}P labelled pR42 was hybridised to EcoRI cleaved D.melanogaster DNA (tracks 3, 5 , 13, 15 and 19) and cDm219 (tracks 4 and 18), and to Bst I, Kpn I, Bst I/EcoRI, Kpn I/EcoRI and Sst I/EcoRI digested D.melanogaster DNA (tracks 8, 11, 14, 16 and 17 respectively), it was also hybridised to Kpn I cleaved cDm219 (track 12). The 1.6Kb Xma I/EcoRI fragment was gel purified, radiolabelled and hybridised to EcoRI cleaved D.melanogaster DNA (track 6). The samples of tracks 7 and 8 were electrophoresed through a 0.5% agarose gel for 40 hours at 1.5V/cm, the remaining samples were electrophoresed on 0.6% gels.

Samples of digested DNA which were electrophoresed on the same agarose gel are arranged in groups (tracks 1 to 4, 5 and 6, 7 and 8, 9 to 12, 13 and 14, 15 and 16, and 17 to 19). EcoRI cleaved D.melanogaster DNA hybridised with ^{32}P pR42 is included where relevant as a reference in the separate groups (tracks 5, 13, 15 and 19). The sizes of various restriction fragments in genomic DNA or cDm219 which hybridise to Zeta sequences are given in Kb alongiside the tracks.

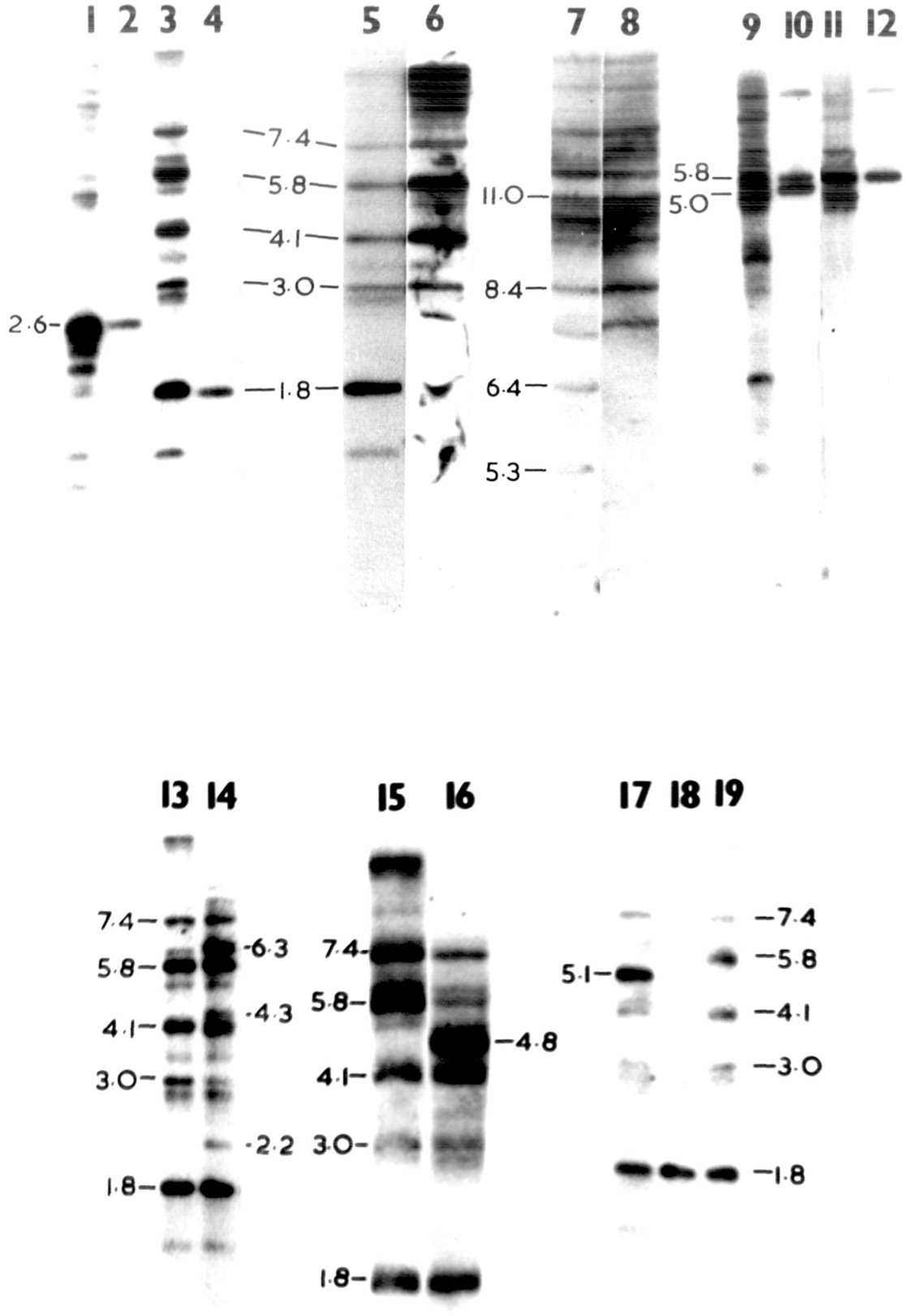

1 2 3 4 5 6 7 8 9 10 11 12
2·6
7·4
5·8
4·1
3·0
1·8
11·0
8·4
6·4
5·3
5·8
5·0
13 14 15 16 17 18 19
7·4
5·8
4·1
3·0
1·8
6·3
4·3
2·2
7·4
5·8
4·1
3·0
1·8
4·8
5·1
7·4
5·8
4·1
3·0
1·8

Figure 19.

Dm219, also hybridises to five other fragments of 3.0, 4.1, 5.8, 7.4 and ∿25Kb with approximately equal intensity.(Figure 19, tracks 3 and 4). These results indicate that the zeta element of Dm219 is part of a family of moderately repetitive sequences within the D.melanogaster genome. The sequences of pR37 and pR42 are found not only in the expected 2.6 and 1.8Kb EcoRI fragments which correspond to those of Dm219, but in the case of pR42 they are also linked to five other sequence arrangements in the genome.

Hybridisation of pR37 and pR42 to BstI Cleaved D.melanogaster DNA

BstI does not cleave the zeta element of Dm219, it also does not cleave the majority of the EcoRI fragments in the genome to which pR42 hybridises. One exception is the ∿25Kb which is digested by BstI to a size of 6.3Kb, in addition several other fragments at a lower intensity are seen (Figure 19,.track 14). This means that pR37 and pR42 will hybridise to identical BstI fragments if they are linked together in the genome. Figure 19, tracks 7 and 8 show the hybridisation of pR37 and pR42 respectively to Bst I cleaved genomic DNA. Both probes hybridise to an identical set of fragments of greater than 7.4Kb, including one which comigrates with the 11.4Kb fragment of Dm219, which contains the zeta element. In addition, pR37 hybridises to fragments of 5.3 and 6.4Kb, which show no hybridisation to pR42. This shows that the majority of sequences homologous to pR37 and pR42 in the genome are linked, but some in pR37 occur independently of those in pR42.

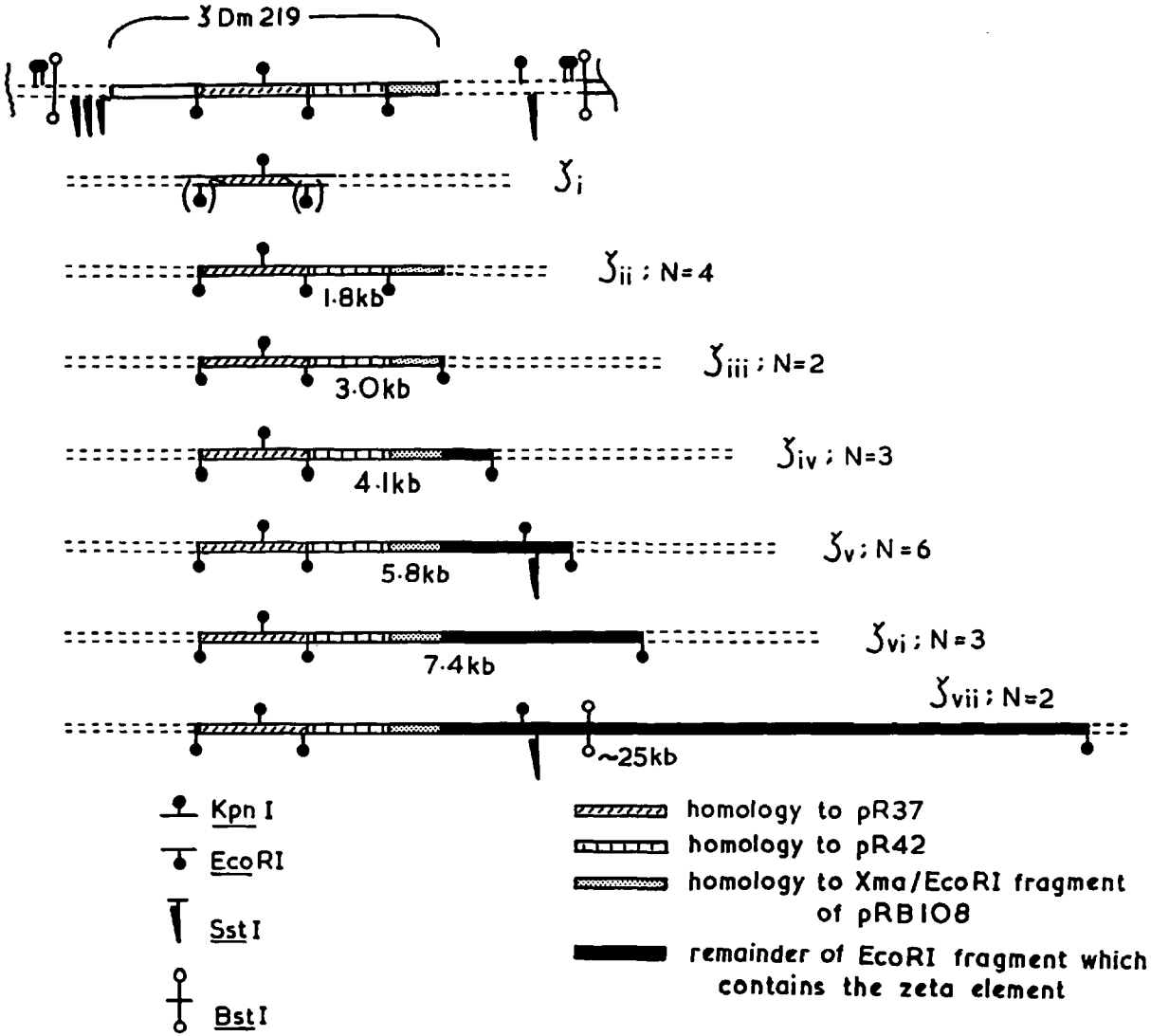
Seven Major Sequence Arrangements of the Dm219 Zeta Element in the D.melanogaster Genome

The above data reveals that there are seven major arrangements of the zeta element within the D.melanogaster genome, these are shown in Figure 20. The smallest of these (ζ_1) consists of the unlinked pR37 sequences, while

Figure 20 The major arrangements of zeta sequences within the D. melanogaster genome

The seven major arrangements were determined from the hybridisation of Dm219 zeta element sequences to restriction endonuclease cleaved D.melanogaster DNA electrophoretically fractionated and transferred to nitrocellulose (Figure 19). For comparison the zeta element of Dm219 and its flanking Type I sequences is shown. The determination of the copy numbers of the various zeta arrangements within the genome is described in the text and in Figure 22.

Figure 20.


ʒ Dm 219
ʒi
ʒii ; N=4
1·8kb
ʒiii ; N=2
3·0kb
ʒiv ; N=3
4·1kb
ʒv ; N=6
5·8kb
ʒvi ; N=3
7·4kb
ʒvii ; N=2
~25kb
Kpn I
EcoRI
Sst I
Bst I
homology to pR37
homology to pR42
homology to Xma/EcoRI fragment of pRB108
remainder of EcoRI fragment which contains the zeta element

the others consist of the 2.6Kb Eco fragment linked to either the 1.8Kb fragment or any of the other five EcoRI fragments to which pR42 will hybridise. The ζ_{11} arrangement of Figure 20 corresponds to the zeta element of Dm219, the restriction map of which is also shown in this figure.

The Flanking Sequences of the Zeta Arrangements

Several experiments were carried out to determine the nature of the sequences flanking the zeta arrangements within the D.melanogaster genome. In particular to see if they were linked to Type I insertion sequences.

Figure 19 (track 6) shows the hybridisation of the 1.6Kb XmaI/EcoRI fragment to EcoRI cleaved genomic DNA. In Dm219 this fragment is located to the immediate right of the 1.8Kb EcoRI fragment cloned in pR42. It hybridises to the same EcoRI fragments in the genome as does pR42 with the exception of the 1.8Kb fragment. Thus not only are the majority of the sequences present in pR37 and pR42 linked in the genome, but they are also linked to those of the XmaI/EcoRI fragment (see Figure 20). This indicates that the EcoRI site separating this fragment from that of pR42 is not conserved in comparison to those producing the 2.6Kb EcoRI fragment.

Experiments were also carried out to identify other restriction fragments in the D.melanogaster genome which contain both the zeta element and Type I insertion sequences. The 11.4Kb BstI fragment of Dm219 has already been identified. KpnI cleaves the zeta element of Dm219 once in the 2.6Kb EcoRI fragment (see Figure 12) producing fragments of 5.8 and 5.0Kb, both of which contain Type I sequences. The sequences in pR37 will hybridise to both of these fragments, while pR42 will only hybridise to the latter. When they are hybridised to KpnI cleaved genomic DNA a large number of fragments are revealed. pR37 hybridises to a smear of fragments of

between 5.0 and 6.0Kb (Figure 19, track 9) which makes it difficult to identify the 5.8 and 5.0Kb fragments of Dm219. A smaller number of fragments hybridise to pR42, most of which are also between 5.0 and 6.0Kb, the most intensely hybridising fragment comigrates with the 5.8Kb KpnI fragment of Dm219 (Figure 19, tracks 11 and 12). This is surprising as the 11.4Kb Bst I fragment of Dm219 is a minor fragment amid a large number of others (Figure 19, tracks 7 and 8). The most likely explanation is that some of the other zeta arrangements in the genome contain an internal 5.8Kb Kpn I fragment. This predicts that some of the other EcoRI fragments which hybridise to pR42 will be cleaved by Kpn I. Figure 19 (track 16) shows the result of hybridising pR42 to an Kpn I/EcoRI digest of Drosophila DNA. Both the ∿ 25 and 5.8Kb EcoRI fragments are cleaved by Kpn I, and a new fragment of 4.8Kb appears. This when combined with the pR37 sequences as shown in Figure 20 equals the 5.8Kb Kpn fragment. A similar experiment was done with Sst I. Both the ∿ 25 and 5.8Kb EcoRI fragments are cleaved by Sst I, resulting in a new fragment of 5.1Kb (Figure 19, track 17).

These two results are shown in Figure 20, together with that of the Bst I/EcoRI digest. It can be seen that the Kpn I and Sst I sites in ζ_V and ζ_{VI} together with the Bst I site of the latter align with sites in Dm219. This suggests that these two arrangements contain Type I sequences, but the complexity of the hybridisation patterns seen in Figure 19 makes this conclusion ambiguous.

The Zeta Sequences to the left of pR37 in the D.melanogaster Genome

The zeta arrangements shown in Figure 20 were derived from hybridisation experiments using pR37 and sequences to its right of the Dm219 zeta element as probes. In order to characterise the sequences to its left in the D. melanogaster genome the 2.1Kb Pst I/Eco RI fragment was used as a probe.

Figure 21 Hybridisation of zeta sequences to the left of pR37 to restriction endonuclease cleaved D.melanogaster DNA

Restriction endonuclease cleaved DNA was electrophoresed through a 0.6% agarose gel and transferred to a nitrocellulose filter. They were then hybridised with either the ^{32}P labelled 2.1Kb Pst I/EcoRI fragment (tracks 3, 4 and 5), ^{32}P pR37 (tracks 1, 2, 6, 7 and 8) or ^{32}P pR42 (track 9). The tracks contain D.melanogaster DNA cleaved with EcoRI (tracks 2 and 3), Pst I (tracks 5, 6 and 9), or Pst I/EcoRI (tracks 4 and 7). pB74 cleaved with Pst I or Pst I/EcoRI was electrophoresed in tracks 1 and 8 respectively, the sizes of some of its restriction fragments are given alongside the tracks.

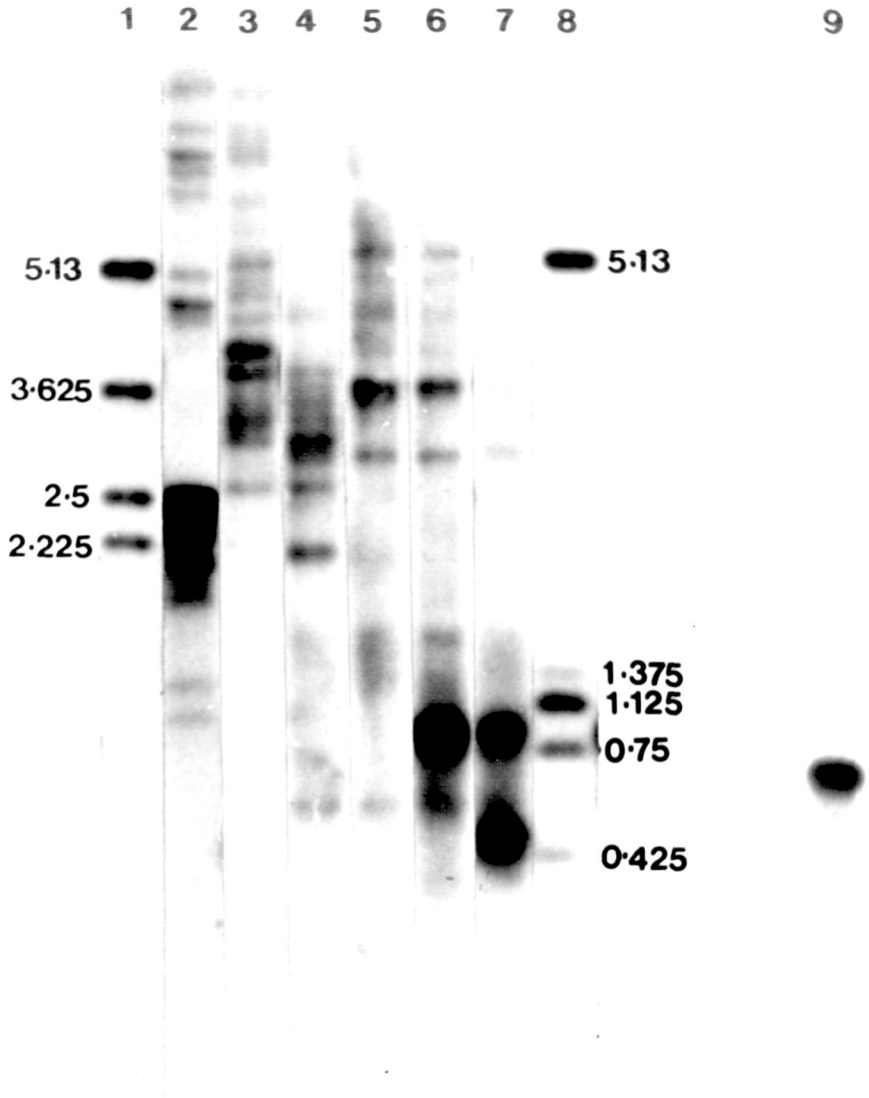


Figure 21.

This fragment together with the Pst I sites of the zeta element are shown in Figure 12. When either this fragment or pR37 is hybridised to Pst I cleaved genomic DNA the 3.6Kb Pst I fragment which is homologous to both of these probes is a minor species among a large number of fragments (Figure 21, tracks 5 and 6). The most intense hybridisation of pR37 is to a fragment of ∿0.90Kb, a slightly smaller fragment is seen when pR42 is used as a probe (Figure 21, track 9). This suggests that it contains sequences homologous to both pR37 and pR42, although its size is different to the 2.5Kb Pst I fragment which overlaps the 2.6 and 1.8Kb EcoRI fragments of Dm219. As expected the 3.6Kb Pst I fragment in genomic DNA, like that of Dm219, is cleaved by EcoRI, but only approximately 60% of the ∿ 0.90Kb Pst I fragments are digested, resulting in a fragment of about ∿0.5Kb (Figure 21, track 7). This is unexpected if it is the equivalent of the Dm219 2.5Kb Pst I fragment which is cleaved by EcoRI, also this site appears to be conserved in genomic DNA (see Figure 20). When the 2.1Kb Pst I/EcoRI fragment is hybridised to a PstI/EcoRI digest of genomic DNA, a fragment which comigrates with it is seen, but there are also about 30 other fragments of equal intensity in the 3.3Kb region of the gel (Figure 21, track 4). This fragment hybridises to a large number of EcoRI fragments in genomic DNA, which range in size from 25 to 2.8Kb, the most intense being 4.8Kb.

These experiments reveal two main features of the zeta arrangements in the D. melanogaster genome. Firstly hybridisation of zeta element probes to Pst I cleaved genomic DNA reveals considerably more internal heterogeneity than was previously shown with Kpn I, EcoRI and Bst I. Secondly sequences homologous to the left hand end of the Dm219 zeta element in the genome are far more heterogeneous than those on the right. This is seen by comparing the hybridisation pattern of the 2.1Kb Pst I/EcoRI fragment to that of the 1.6Kb XmaI/EcoRI fragment to EcoRI cleaved genomic DNA (Figure 19, track 6; Figure 21, track 3).

4. Quantitation of Zeta Sequences in the D.melanogaster Genome

The previous section described the arrangement of sequences in the D.melanogaster genome homologous to the zeta element of Dm219. It indicates that the arrangement found in Dm219 is a minor part of the total number. Some of the other zeta arrangements also appear to be linked to Type I sequences.

The number of copies of each zeta arrangement in the D.melanogaster genome was determined by quantiative gel transfer hybridisations (Lis et al., 1978). A standard curve of the intensity of hybridisation of the ^{32}P probe to known amounts of either cDm219 or pB74 was constructed. By comparing the intensity of hybridisation of restriction fragments in a known amount of genomic DNA with the standard curve, the copy number of the restriction fragment could be determined. Efficient transfer of the larger restriction fragments in the genome relative to those of cDm219 used to construct the standard curve was achieved by short wave ultra violet irradiation of the gel prior to transfer. This nicked the DNA within the gel so reducing its length when denatured for transfer to the nitrocellulose filter. The efficiency of transfer of fragments in different parts of the gel was monitored by including Hind III cleaved phage λ in several gel tracks. It was detected after transfer to nitrocellulose by including in vitro labelled $^{32}P\lambda$ (Rigby et al., 1977) in the hybridisation solution.

The copy number of the EcoRI fragments in genomic DNA homologous to pR42 and pR37 was calculated from the standard curve of the hybridisation of these probes to EcoRI cleaved cDm219 (Figure 22A). The copy numbers of the major EcoRI fragments which hybridise to pR42 and represent the zeta arrangements of Figure 20 are 4,2, 3, 6, 3 and 2 for the 1.8(ζ_{11}), 3.0 (ζ_{111}), 4.1 (ζ_{1V}), 5.8(ζ_{V}) 7.4(ζ_{V_1}) and ∿ 25Kb(ζ_{V11}) EcoRI fragments respectively. This

Figure 22 Quantitative gel transfer hybridisations of zeta sequences to D.melanogaster DNA

Quantitative gel transfer hybridisations were carried out as described by Lis et al. (1978). After electrophoresis the gel was irradiated with short wave ultraviolet light for 10 minutes. The DNA was then transferred to nitrocellulose filters and hybridised at a concentration of 20ng/ml with ^{32}P pR42 (specific activity 8.8 x 10^7 cpm/µg) or ^{32}P pR37 (specific activity 6.4×10^7 cpm/µg). In both cases ^{32}P labelled λ DNA (specific activity 7 x 10^7 cpm/µg) was included at a concentration of 0.23 ng/ml. Several autoradiographs were taken, and ones which gave a linear dose response with increasing amounts of plasmid DNA when traced with a Joyce-Loebal microdensitometer were used to estimate the amount of zeta sequences within the D. melanogaster genome.

(A) Track 1 contains 1.03µg of EcoRI cleaved D.melanogaster DNA, tracks 2, 3, 5, 6, 7, 9 and 10 contain 0.019, 0.034, 0.078, 0.16 and 0.31 ng of EcoRI cleaved cDm219 respectively, tracks 4, 9 and 11 contain 0.5µg of Hind III cleaved λ. After electophoresis through a 0.6% agarose gel and hybridisation with ^{32}P labelled pR42 and λ the autoradiograph which is shown, was used to construct the illustrated linear close response for tracks 2,3, 5 and 6. These contain relative to the D.melanogaster the equivalent of 0.97, 1.94, 3.84 and 7.83 copies of the 1.8Kb EcoRI fragment per genome. A densiometric tracing of track 1 is shown, the ordinate is calibrated from the linear dose response curve.

(B) Hybridisation of ^{32}P pR37 to 1.03µg of EcoRI cleaved D.melanogaster DNA after electrophoresis through a 0.6% agarose gel and transfer to a nitrocellulose filter. A densiometric tracing of the autoradiograph is shown. The ordinate was calibrated from a linear dose response curve of the hybridisation to EcoRI cleaved cDm219 which contained

the equivalent of 0.97 to 31.33 copies of the 2.6Kb fragment per genome. In this exposure the hybridisation of pR37 to the 2.6Kb EcoRI fragment in the D.melanogaster genome is outside the linear dose response of the tracing. A shorter exposure where a linear dose response was obtained from 8 to 64 copies of the 2.6Kb fragment was used to estimate that it is present 18 times per genome.

(C) ^{32}P pR42 was hybridised to 2.06µg of Bst I cleaved D.melanogaster DNA which has been electrophoresed through a 0.5% agarose gel at 1.5v/cm for 40 hours. A densiometric tracing of the autoradiograph is shown, its ordinate was constructed from the linear close response curve of the hybridisation of pR42 to Bst I cleaved pB74 which contained the equivalent of between 1.03 and 16.62 copies of the 11.4Kb Bst I fragment per genome.

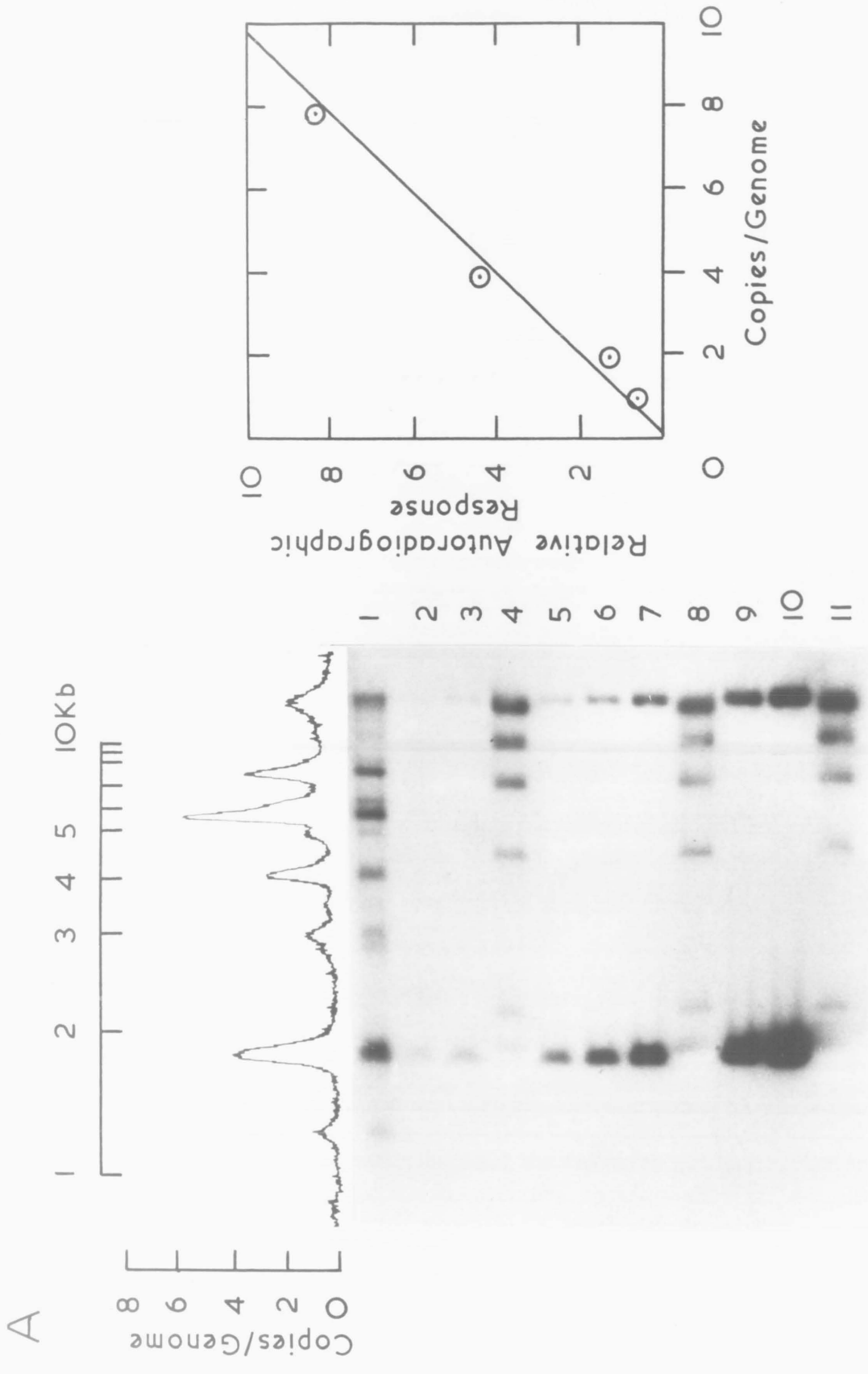

A
Copies/Genome
0 2 4 6 8
1 2 3 4 5 10Kb
1 2 3 4 5 6 7 8 9 10 11
Relative Autoradiographic Response
0 2 4 6 8 10
Copies/Genome
0 2 4 6 8 10

Figure 22.

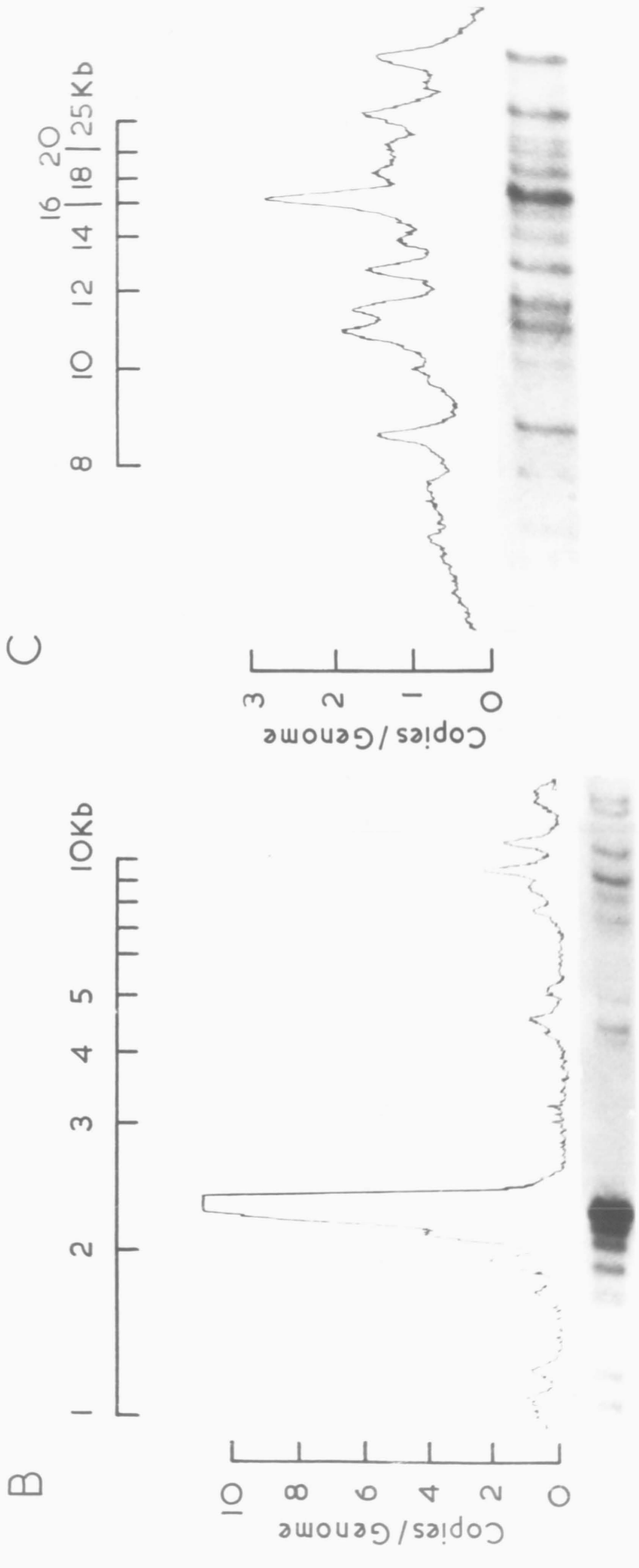

C
8
10
12
14
16
18
20
25Kb
Copies / Genome
0
1
2
3
B
1
2
3
4
5
10Kb
Copies / Genome
0
2
4
6
8
10

Figure 22.

adds up to 20 copies of pR42 sequences in these arrangements. The total number was estimated to be 32 by summing the area under the densitometric tracing of pR42 hybridised to EcoRI cleaved genomic DNA.

pR37 hybridizes mainly to an EcoRI fragment in genomic DNA which comigrates with the 2.6Kb EcoRI fragment of Dm219, this was found to be present 18 times per haploid genome (Figure 22B). This is close to the value of 20 for the copy number of the major EcoRI fragments which hybridise to pR42, and so it is consistent with the 2.6Kb EcoRI fragments being linked to these in the genome as is shown in Figure 20. When the minor fragments which hybridise to pR37 are included, the total number of sequences is 42 per haploid genome. The discrepancy between the total number of sequences homologous to pR37 and pR42 in the genome (42 compared with 35) might be due to the presence of additional pR37 sequences unlinked to those of pR42, which would result in a higher copy number.

Both pR37 and pR42 hybridise to identical Bst I fragments which are larger than 7.4Kb in the genome, indicating that the majority of these two sequences are linked. The copy number of the Bst I fragments homologous to pR42 was determined using a standard curve of the hybridisation of pR42 to Bst I cleaved pB74. From this the copy number of each of the 14 major Bst I fragments which hybridise to pR42 was found to be between one and two per haploid genome (Figure 22C). The 11.4Kb Bst I fragment of Dm219 was found to be present approximately twice per haploid genome. The total number of sequences homologous to pR42 in this experiment was 31, which is close to the value of 35 estimated from its hybridisation to EcoRI cleaved D.melanogaster DNA.

The copy numbers from these experiments suggest that the same zeta arrangements are present on several different Bst I fragments in the genome. This is because there are only one or two copies of each of the major Bst I fragments per genome, while there are between two and six copies of each of

the zeta arrangements shown in Figure 20. This means that some of the members of a particular arrangement have to be present on different Bst I fragments. It then accounts for there being four copies of the ζ_{11} arrangement found in Dm219, but only two copies of the Dm219 11.4Kb Bst I fragment. The additional members of ζ_{11} presumably are found on other Bst I fragments. Alternatively the members of a particular zeta arrangement might differ due to the presence or absence of Bst I sites, (compare the intensitities of EcoRI fragments in track 13 with those in the Bst I/EcoRI digest of track 14, Figure 19).

Some of the restriction fragments in Bst I or EcoRI cleaved genomic DNA which hybridise to pR37 or pR42 are present less than once per haploid genome. There are several possible reasons for this. Firstly, the restriction fragments in the genome might only be partially homologous to the probe. This could be due to sequence divergence creating additional Bst I or EcoRI cleavage sites, or decreasing the stability of hybrids between the fragment and the probe. Both would lead to a weaker hybridisation signal and hence underestimation of the copy number. Alternatively the genomes of individual flies in the population from which the DNA was prepared might vary in their restriction sites, which would also lead to fragments being present less than once per genome.

5. Linkage between Type I and Zeta Sequences in the D.melanogaster Genome

The copy numbers obtained from the quantitative gel transfer hybridisations described in the previous section allow an estimation on the degree of linkage between Type I and zeta sequences. This assumes that the restriction sites present in ζ_V and ζ_{V11} indicate that they contain Type I insertion sequences (see Figure 20). If this is correct, then together with the two copies of Dm219 in the genome, a total of 12 out of

35 (34%) of the zeta sequences (represented by pR42) are linked to Type I insertion sequences.

Because of the complexity of the hybridisation patterns of zeta element probes to genomic DNA, the assumption that ζ_V and ζ_{V11} contain Type I sequences may be incorrect. More definitive evidence of their linkage in the genome was obtained by an alternative approach. This was to measure the frequencey of plaques obtained from a library of randomly sheared D.melanogaster Oregon R DNA cloned in phage λ which hybridise to both zeta element and Type I insertion sequences. The library was obtained from D.S. Hogness, and its construction is described in Chapter 2. It was used to infect a lawn of E.coli. Four replicas of the resulting plaques were then taken onto nitrocellulose disks by the procedure of Benton and Davies (1977). These were then bybridised to radioactive probes containing zeta element (pR37 and pR42), Type I insertion (pB7) or rDNA (pDm238/29) sequences. After autoradiography the number of plaques homologous to each probe was counted (see Table 2). Using this data the copy number of each sequence in the genome could be calculated from:

$$\frac{\dfrac{P}{2.4 \times 10^4} \times \dfrac{L}{20} \times 1.65 \times 10^5}{L}$$

Where 2.4×10^4 is the number of plaques screened, 20 is the average size (in Kb) of the D.melanogaster DNA carried by each plage, and 1.65×10^5 is the size of the D.melanogaster haploid genome in Kb (Rasch et al., 1973). P is the number of plaques which hybridised to each probe, and is given in Table 2. L is the length of the segment of Drosophila DNA in the probe, which was 1.8 for pR42, 2.6 for pR37, 4.0 for pB7 and 12Kb for pDm238/29. From this the following copy numbers per haploid genome was calculated to be 18,

TABLE 2 The number of plaques homologous to Type I, rDNA and zeta Sequences obtained from a library of D.melanogaster DNA cloned in bacteriophage λ

^{32}P labelled probes	pDm238/29		pB7		pR37		pR42	
	Number	%	Number	%	Number	%	Number	%
pDm238/29	149	63.5	69	26.5				
pB7	69	31.7	171	65.7	8	12.3	12	23
pR37			8	3.0	44	67.6	17	32.6
pR42			12	4.8	17	26.1	23	44.4
TOTAL	218	100	260	100	65	100	52	100

A total of 2.4 x 10^4 plques generated from a library of D. melaogaster DNA cloned in bacteriophage λ (materials and methods) were screened by the procedure of Benton and Davies (1977). The nitrocellulose filters were hybridised with ^{32}P labelled DNA which had specific activities of approximately 10^7 cpm./μg at a concentration of 2ng/ml. The table gives the number of plaques which hybridised to one or two probes. Four plaques hybridised to pB7, pR37 and pR42. The percentages are of the total number of plaques which hybridised to a given probe.

22, 89 and 75 for pR42, pR37, pB7 and pDm238/29 respectively.

These values rest on two assumptions, the first is that each sequence is found only once in any given recombinant phage, and secondly that there is no preferential loss of particular sequences from the library. For tandemly repeated sequences of less than 20Kb in length such as rDNA and the Type I sequences of Dm219 the first assumption is incorrect. In the case of rDNA where the size distribution of the units is known the average size can be calculated to be 14Kb. This means that phage with 20Kb inserts of Drosophila rDNA will contain on average 1.43 units. When the original copy number is multiplied by this a value of 105 per haploid genome is found. For non-rDNA Type I sequences, the proportion which are arranged in tandem, and their size distribution is unknown, and so the copy number cannot be adjusted. The second assumption is also unjustified as some tandemly arranged sequences are unstable in recombinant phage λ (B. Klein and K. Murray, personal communication). If this were true for the rDNA, then it would account for the low value of 105 units per haploid genome in comparison to values of 250 as determined by DNA:RNA saturation hybridisation experiments (Tartof, 1973). The second assumption might also affect tandemly arranged Type I sequences, and possibly zeta sequences associated with them. This would explain the different copy numbers obtained by this method compared to those obtained by quantitative gel transfer hybridisations (22 against 42 for pR37, and 18 against 35 for pR42).

The degree of linkage between zeta sequences and Type I sequences was determined from Table 2, 33% of sequences homologous to pR42 are linked to those found in pR37, and 23% to the Type I sequences of pB7. This is lower than those determined by quantitative gel transfer hybridisation experiments where values of 100% and possibly 34% were obtained. For pR37, 26% are linked to pR42 and 12% to pB7, which is also lower than the previously

determined values of 83% and possibly 28% respectively.

The plaques which contained Type I sequences could hybridise to either the rDNA or zeta element probes, or to neither. From Table 2 the number of rDNA units with Type I insertions is 31%. The DNA used to construct the Drosophila library was prepared from a mixture of males (X/Y) and females (X/X). As Type I insertions only occur in the rDNA of the X chromosome, in 49% of the units (Wellauer et al., 1978), the expected percentage of plaques containing rDNA which hybridise to Type I probes can be calculated to be about 36.75%. The majority of the plaques which hybridised to pB7 did not hybridise to the rDNA probe, this value (74%) is higher than expected but supports the conclusion that there is a large number of non-rDNA Type I sequences in the genome.

A comparison of the results of the plaque hybridisation experiment with those obtained by other methods indicate that it is only semi-quantitative. The data is consistent with the structure of Dm219 as zeta sequences can be found linked to Type I sequences in the genome.

CHAPTER V

DISCUSSION

The number of cloned segments of non-rDNA Type I sequences isolated was considerably lower than expected. Quantitation of Type I sequences fractionated on cesium chloride gradients containing various dyes indicate that similar amounts are found linked and unlinked to the rDNA (Dawid and Wellauer, 1977). This means that of the colonies which hybridise to cRNA from pDm103, half of those containing Type I sequences should not hybridise to rDNA probes. Thirty-seven out of the 144 screened hybridised to the insertion probe and so approximately 17 should have contained non-rDNA Type I seqences but only one, cDm219, was found. This probably results from preferential transcription of rDNA rather than the insertion sequences of pDm103 by E.coli RNA polymerase. This would bias the original screen with such cRNA towards the selection of colonies containing rDNA. Support for this comes from the finding of Type I sequences in the plasmid cDm207 (Roiha and Glover, manuscript in preparation) previously thought not to contain Type I sequences, since it did not show hybridisation to cRNA from pDm103 (Glover, 1977).

The plasmid, cDm219, which was isolated from this screen was shown by restriction mapping and hybridisation experiments to consist of five tandemly arranged units homologous to the insertion of Dm103 (Figure 8). Units 3 and 4 are only 4Kb, both having an identically positioned deletion of 1.4Kb. In contrast unit 2 is 11.4Kb as a result of an 8Kb substitution, the zeta element, which replaces 2Kb of Type I sequence.

The arrangement of Type I sequences in Dm219 is characteristic of a significant proportion in the D.melanogaster genome. EcoRI does not cleave Type I sequences, and so tandem arrangements like Dm219 are contained in large EcoRI fragments which band at a lower density in cesium chloride-actinomycin D density gradients than rDNA. These large EcoRI fragments are digested to fragments the size of the repeating unit by restriction endonucleases such as Bst I which have only one cleavage site per unit (Figure 18). The BstI generated fragments form a compact array of between 3.8 and 5.4 Kb, suggesting that the units of the tandem arrays are heterogeneous in size.

Not only do the units of Dm219 differ by the presence or absence of deletions and substitutions, they also show variation in the position of restriction sites. Most of this variation is limited to two Sst I fragments at the left of each unit which vary in size from 0.35 to 0.29Kb. Only in units 4 and 5 are these fragments identical. Units 3 and 4 have identically positioned deletions but different restriction maps. Unequal exchange might act on tandem arrays of Type I sequences resulting in the saltation of similar units. The deletion might have occurred in one unit which was then duplicated, for example by exchange events, and the two units might then have undergone sequence divergence to produce the different restriction maps of units 3 and 4. Unlike the Type I sequences of Dm219 those of the rDNA are relatively homogeneous. When three other full length Type I rDNA insertions were examined, only one, Dm206, had a different restriction endonuclease cleavage pattern to Dm103. This was found to be similar to the variation in the two small Sst I fragments of the units of Dm219. One of these fragments in Dm206 is about 20 bp shorter than the equivalent fragment in Dm103. The variation in this region of the Type I sequences of Dm219 and the rDNA might aldo have arisen by unequal exchange in their respective tandem arrays. The differences between the Type I sequences represented by Dm219 and those of the rDNA suggests that exchange does not occur between the two sequences. This would argue that Type I insertions are not transposable elements of the type found in prokaryotes (reviewed by Shapiro et al., 1977; Kleckner, 1977)

or possibly in Drosophila (Finnegan et al., 1977; Potter et al., 1979; Strobel et al., 1979). These transposable elements show very little sequence variation. It is possible that only Type I sequences located in the rDNA are transposable in which case they would have a very specific intergration process limited to a sequence in the 26s rRNA gene.

The substitution found in unit 2 (the zeta element) was found to contain homologous sequences at or near its ends. This is similar to the transposable elements of prokaryotes and the putative elements of Drosophila (Finnegan et al., 1977). The distribution of zeta sequences within the D.melanogaster genome was examined by the technique of Southern (1975). Seven major sequence arrangements of the zeta element were found, constituting a total of 20 copies per haploid genome (Figure 20). Two of the arrangements have restriction maps suggesting that they are linked to Type I sequences. However, the complexity of the hybridisation pattern of zeta element probes to genomic DNA renders this conclusion ambiguous. The frequencies of recombinant phage containing both Type I and zeta sequences in a library of D.melanogaster DNA cloned in phage λ, indicates that 12 to 23% of the zeta sequences are linked to Type I sequences. This is to be compared with possibly 28 to 34% determined by gel transfer hybridisation experiments. A comparison of both copy number and linkage values determined by plaque hybridisation experiments with those obtained by other methods suggest that they are most likely underestimates. For example, there are 105 copies of rDNA per haploid genome when determined by this method, compared with 250 from RNA:DNA saturation hybridisation experiments. Secondly the gel transfer hybridisation experiments indicate that 100% of the sequences in pR42 are linked to those of pR37 in the genome, compared with 17% determined by plaque hybridisation. The frequent association of zeta sequences with Type I sequences might explain the deletion in unit 2. If two zeta elements are present at different sites within tandemly arranged Type I sequences, recombination

between the two elements would lead to a deletion in one of the units (such as unit 2 of Dm219) and a duplication in the other.

The copy number of the zeta element suggests that it forms part of the middle repetitive DNA of the D.melanogaster genome. This sequence class in Drosophila makes up about 16% of the genome, and consists of families of sequences which are each present about 35 to 100 times dispersed in 6.0 Kb blocks throughout the genome (Manning et al., 1975). The chromosomal location of a particular middle repetitive sequence represented by a cloned segment of DNA varies in different strains of D.melanogaster. This suggests that the middle repetitive sequences have the ability to transpose to different chromosomal locations (Finnegan et al., 1977; Strobel et al., 1979; Potter et al., 1979; Young, 1980). The zeta element of Dm219 might be similar to these putative transposable elements. Firstly it has a sequence repeat at or near its ends and secondly, the different zeta arrangements (Figure 20) might represent different chromosomal locations and hence different flanking sequences. As at least 12 to 23% of the zeta sequences are linked to Type I sequences and most of the zeta arrangements are present more than once per genome, it suggests that if the zeta element is transposable, it only integrates into certain sequences.

The Chromosomal location of Dm219

The insertion of Dm103 hybridises in situ to the nucleolus, the chromocentre and to chromosome 4 of polytene chromosomes (Peacock et al., manuscript in preparation). In the salivary glands of D.melanogaster the centromeres of the polytene chromosomes and their surrounding heterochromatin fuse together forming the chromocentre (Lefevre, Jr., 1976). The hybridisation to the chromocentre of the Dm103 insertion could locate these segments to the centromeric heterochromatin of any of the four chromosomes in the D.melanogaster karotype. A better idea of its location is obtained by in situ

hybridisation to the diploid metaphase chromosomes of neuroblasts. Here almost all of the hybridisation is limited to a block of heterochromatin which contains the nucleolus of the X chromosome. Less than 10% of the silver grains are located on the Y chromosome and none at all on any of the other chromosomes. This suggests that the majority of the non-rDNA Type I sequences are located in the heterochromatin surrounding the nucleolus of the X chromosome (Peacock at al., manuscript in preparation). Only 5% of the silver grains are located over the chromocentre when the Dm103 insertion is hybridised in situ to polytene chromosomes, although similar amounts of Type I sequences occur linked and unlinked to the rDNA (Peacock et al., manuscript in preparation; Dawid and Wellauer, 1977). This might be because the repetitive DNA of the chromocentre is under-replicated during polytenisation (Gall et al., 1971) which would lead to a weaker hybridisation signal from the non-rDNA type I sequences compared to those in the nucleolus. Alternatively most of the non-rDNA type I sequences might occur in the nucleolus linked to the rDNA although not interrupting it. The position of the non-rDNA Type I sequences has been further localised using inversions of the X chromosome. The proximal breakpoint of the W^{m4} inversion is located very close to the distal end of the rDNA (see Figure 2). In this inversion a large part of the X heterochromatin (Xh) distal to the rDNA is now located near the telomere. When rDNA is hybridised in situ to this chromosome in diploid neuroblasts the silver grains are still located over the heterochromatin near the centromere. In contrast, when Type I insertion sequences are used as a probe, two regions of the chromosome hybridise, one region near the centromere and the other near the telomere (R. Appels, personal communication). For hybridisation to occur in situ to diploid chromosomes, the chromosomal sequences must be repetitive. This means then a large amount of the non-rDNA Type I sequences are located in the heterochromatin of the X chromosome distal to the nucleolus (see Figure 2).

Bob Levis and David Glover have carried out in situ hybridisations with subcloned fragments of Dm219 to salivary gland polytene chromosomes. pB7 which only contains Type I insertion sequences hybridises as expected to the nucleolus, the chromocentre and to a band of chromosome 4. When pR37 is used as a probe the silver grains are located over the chromocentre and a lesser number over region 42A of the second chromosome. pR42 hybridises solely to the chromocentre. The common region of hybridisation of these probes representing both the Type I and zeta sequences is to the chromocentre; thus Dm219 originates from this chromosomal region. This suggests that Dm219, like most non-rDNA Type I sequences, is located in the heterochromatin distal to the rDNA of the X chromosome, which is part of the chromocentre of polytene chromosomes.

The Chromocentre

The heterochromatin of the chromosomes which fuses to form the chromocentre makes up a larger part of the genome than is apparent in the polytene chromosomes of salivary glands. A better idea of the proportion of heterochromatin in the genome is obtained when diploid metaphose chromosomes from neuroblasts are examined. Here 22 to 31% of the haploid genome of females is darkly staining heterochromatin (Peacock et al., 1977). The Y chromosome which contains about the same amount of DNA as the X chromosome cannot be seen in salivary gland polytene chromosomes because it is completely heterochromatic in this tissue. The Y chromosome does however contain genes for male fertility (Williamson, 1976), suggesting that the heterochromatin is not necessarily genetically inactive.

About 25% of the D.melanogaster genome consists of highly repetitive satellite sequences. In situ hybridisation locates these sequences almost entirely to the heterochromatin, with very little labelling of the euchromatic

arms of polytene chromosomes (Peacock et al., 1977). The heterochromatin of the X and Y chromosomes also contains the genes for rRNA (Ritossa, 1976). Cloned middle repetitive sequences hybridise in situ to the chromocentre (Wensink et al., 1974; Glover et al., 1975). Some of these of the "copia" type are dispersed throughout the genome and are thought to be transposable elements (Finnegan et al., 1977). Also sequences homologous to the tandem arrays located at 87C1-3 which are transcribed during heat shock are found at the chromocentre (Lis et al., 1978). Both polysomal poly A^+ RNA from normal cells and poly A^+ RNA from heat shocked cells hybridise in situ to the chromocentre (Spradling et al., 1975; Spradling et al., 1976). However in both of these cases it is not clear if the RNA is transcribed from chromocentric DNA or homologous sequences in the euchromatin.

The Heterochromatin Distal to the rDNA of the X Chromosome

The sequence organisation of the heterochromatin is clearly complex and the significance of finding Type I and zeta sequences in it is not obvious. It suggests that the zeta element is not part of the dispersed middle repetitive putative transposing sequences of Drosophila. A significant proportion of the non-rDNA Type I sequences are located in the Xh distal to the rDNA. This region of the heterochromatin has been shown to affect the expression of the rDNA.

Procunier and Tartof (1978) have put forward the hypothesis that the Xh contains a region capable of sensing a deficiency in its homologue and then promoting the disproportionate replication (dosage compensation) of its own rDNA. This region, named the compensatory response (cr) locus, was localised to the heterochromatin distal to the rDNA by using various deficiences of the X chromosome. A chromosome which is deficient for $cr(cr^-)$ is sensed by its homologue (cr^+) which in response, disproportionately replicates its rDNA. For this to occur the cr^+ locus has to be contiguous with its rDNA.

If it is separated from its rDNA by an inversion of the X chromosome such as W^{m4} (see Figure 2) and placed with a cr^- chromosome, no dosage compensation occurs. The cr locus of the X can detect cr^-Y chromosomes and cause dosage compensation. However it is claimed the reverse does not occur; a cr^+Y chromosome does not dosage compensate when placed with a cr^-Y chromosome. (Procunier and Tartof, 1978).

Sequences in this region of the Xh have also been found to interact with a autosomal gene. Female homozygous for the gene abnormal oocyte (abo), which is an autosomal recessive maternal effect mutant of the second chromosome, produce eggs which have a decreased change of survival relative to wild type flies. The eggs can be rescued by increasing doses of either the Y chromosome or of Xh (Sandler, 1970).

By testing of the eggs of homozygous abo females caould still be rescued by X chromosome deficient for various regions of Xh, the region responsible for this effect was located to the Xh distal to the rDNA (Xh^{abo}, Parry and Sandler, 1974). When flies are kept homozygous for abo over several generations they cease to show the abo phenotype and have a several fold increase in their rDNA content (Krider and Levine, 1975). It seems possible that this increased amount of rDNA is causing the loss of the abo phenotype and that this is being mediated by the Xh distal to the rDNA.

The molecular basis of cr and Xh^{abo} is completely unknown. The finding that non-rDNA Type I sequences are found in the same region of heterochromatin as cr and Xh^{abo} is suggestive but not proof of a relationship between the function of the rDNA and these sequences. It would be interesting to see if the tandem arrays (large EcoRI fragments) of Type I sequences such as Dm219 occur in the X chromosome deficiencies used to map cr and Xh^{abo}. However, even if they were missing in the deficiencies which gave cr^- phenotypes and failed to rescue the eggs of abo/abo females, it would not necessarily mean a correlation as the deficiencies remove large segments of DNA. The

analysis of other cloned segments of Type I sequences would help in the characterisation of this region of the X heterochromatin. It would be of particular interest to characterise the other zeta arrangements by molecular cloning, in order to see if they contained repeating sequences at or near their ends, and to obtain unambiguous evidence of their high frequency of linkage to Type I sequences. The analysis of the zeta arrangements in different strains of D.melanogaster by Southerns (1975) technique would make it possible to see if the different arrangements reflect zeta elements transposing into different sequences. The multiple copies of each arrangement might suggest that zeta elements have specific sites of intergration in repetitive sequences.

In conclusion then the easiest approach to understanding the role of Type I sequences (both in and outside the rDNA) and the X heterochromatin is a combination of methods. This would be to characterise more cloned segments of DNA from this region, coupled with genetic manipulation of the Xh and the rDNA,to obtain a possible correlation between the phenotypic effect of the manipulation and changes occurring at the DNA level.

APPENDIX I

ISOLATION AND IDENTIFICATION OF SUBCLONES

The general procedure for the subcloning of restriction fragments into pBR322 is given in materials and methods. It also describes the two procedures used to identify recombinant plasmids and their partial purification from cleared lysates for restriction analysis.

Figure 7 shows a restriction map of pBR322 together with the position of its amplicillin and tetracycline resistance genes. It can be seen that the tetracycline resistance gene contains Bst I and Hind III sites. Insertion of a restriction fragment into these restriction sites inactivates the tetracycline resistance gene and hence provides a preliminary screen for the identification of bacteria containing recombinant plasmids. This involved selection for ampicillin resistant transformants and then screening these for tetracycline sensitivity. Bacteria of this phenotype should contain recombinant plasmids. This was confirmed and the identity of the inserted fragment determined by restriction endonuclease analyses of partially purified DNA from cleared lysates prepared from the tetracycline sensitive colonies.

The above procedure for selecting recombinant plasmids could not be used when fragments were inserted into the pBR322 EcoRI site as this does not occur in the tetracycline resistance gene. Instead the colony hybridisation procedure (materials and methods) was used to identify colonies containing recombinant plasmids. This was checked by restriction analysis of DNA prepared from the positive colonies.

1. Subcloning Restriction Fragments from cDm219

The bacterial vector pBR322 and cDm219 were cleaved with either EcoRI, Bst I, or Bst I/EcoRI and ligated together and used to transform E. coli. The

isolation and identification of subclones containing the various restriction fragments of cDm219 is described below.

The 2.6 and 1.8Kb EcoRI fragments of cDm219

Ampicillin resistant transformants were selected, and screened by the colony hybridisation procedure for recombinant plasmids containing cDm219 EcoRI fragments. The probes used were the 2.6 and 1.8Kb EcoRI fragments of cDm219, gel purified and nick translated with ^{32}P. Two of the positive colonies from this screen pR37 and pR92 contain the 2.6 and 1.8Kb EcoRI fragments of cDm219 respectively (Figure 23, tracks 1 and 3).

The Bst I fragments of cDm219

DNA was prepared from the ampicillin resistant, tetracycline sensitive colonies and digested with either Bst I or Sst I. Digestion of plasmids containing the 4Kb Bst I cDm219 fragments with Bst I gives a single band, consisting of the 4Kb Bst I insert and the 4.3Kb pBR322 vector which comigrated on this gel system (Figure 23 tracks 10 and 11). Cleavage of these plasmids with Sst I vave slightly different restriction patterns, two of the clones pB2 and pB7 with different Sst I patterns (Figure 23 tracks 17 and 18) were taken as representative of the two 4Kb Bst I fragments of cDm219.

Plasmids containing the 11.4 and 12.6Kb Bst I fragments were identified by the presence of large Bst I fragments in the plasmids pB11 and pB74 (Figure 23 tracks 8 and 9). Digestion with Sst I gave slightly different restriction patterns with pB11 and pB74. The latter was shown to contain the 11.4Kb Bst I fragment, as upon EcoRI digestion (Figure 23 track 21) it gave the 2.6 and 1.8Kb EcoRI fragments of cDm219.

The Bst I/EcoRI fragments of cDm219

Partially purified DNA was prepared from cleared lysates (materials and methods) of ampicillin resistant, tetracycline sensitive transformants and

Figure 23 Identification of recombinant plasmids containing the subcloned EcoRI, EcoRI/Bst I and Bst I fragments of cDm219

EcoRI cleaved pR37 (track 1) and pR42 (track 3) electrophoresed on a 0.8% agarose gel alongside Bst I/EcoRI cleaved cDm219 (track 2), the position of the 2.6 and 1.8Kb EcoRI fragments are indicated. EcoRI/Sst I digests of cDm219 (track 4), pRB46 (track 5) and pRB108 (track 6) together with EcoRI cleaved pBR322 (track 7) electrophoresed on a 0.8% agarose gel, the position of the 2.8 and 2.27 Kb Sst I/EcoRI fragments are indicated. Bst I digest of cDm219 (track 12), together with Bst I digests of the subcloned Bst I fragments pB74 (track 8) pB11 (track 9) pB7 (track 10) and pB2 (track 11), a Bst I digest of pBR322 is shown in track 13. The position of the 12.6 and 11.4Kb Bst I fragments of cDm219 and of linear pBR322 is indicated. Sst I digestions of pB74 (track 15) pB11 (track 16) pB7 (track 17), pB2 (track 18) and of cDm219 (track 19) electrophoresed on a 1.5% agarose gel. EcoRI digestion of cDm219 (track 20) and pB74 (track 21) electrophoresed on a 0.6% agarose gel. The samples in tracks 5 and 6, 8 to 11 and 15 to 18 are from cleared lysates. The letter 'p' refers to partial digestion products.

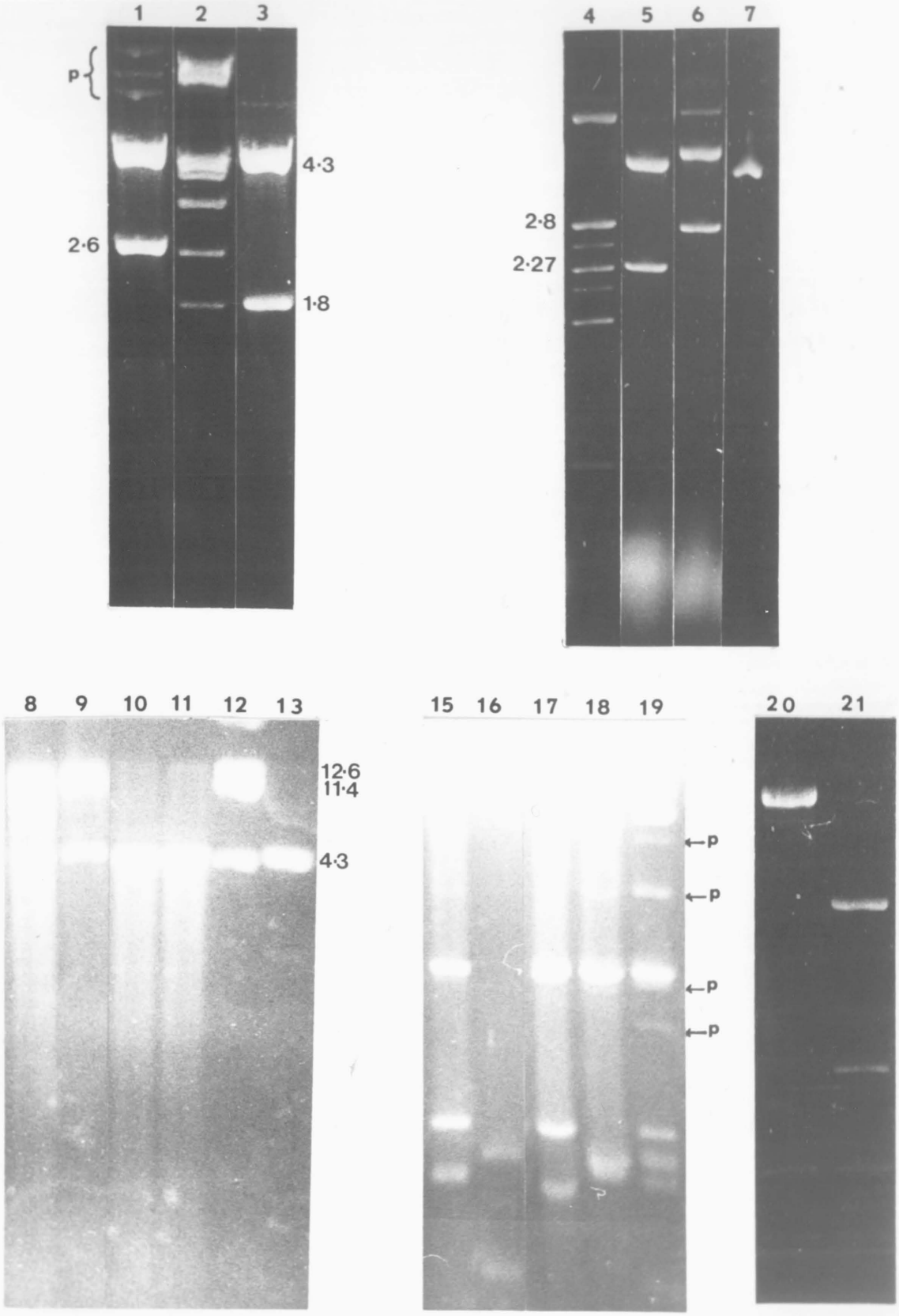

1
2
3
P
4·3
2·6
1·8
4
5
6
7
2·8
2·27
8
9
10
11
12
13
12·6
11·4
4·3
15
16
17
18
19
20
21
P
P
P
P

Figure 23.

the inserted Bst I/Eco RI fragment identified by Sst I/EcoRI digestion. The Sst I site closest to the EcoRI site of the 3.1 and 3.9 EcoRI/Bst I fragments had been mapped by gel purifying these fragments and digesting them with Sst I (data not shown). DNA from the clone pRB46 gave an 2.27Kb EcoRI/Sst I fragment (Figure 23, tracks 5 and 6) characteristic of the 3.1Kb EcoRI/Bst I fragment (Figure 23, track 5) while pRB108 gave an 2.8Kb fragment characteristic of the 3.9Kb EcoRI/Bst I fragment (Figure 23, track 6).

2. Subcloning the 11.4Kb EcoRI fragment of cDm238

The plasmid cDm238 contains a segment of randomly sheared D.melanogaster DNA inserted into the bacterial vector ColE1 by the dA:dT tailing technique. The segment has an 11.4Kb EcoRI fragment containing an entire rDNA unit lacking an insertion (D. Glover, personal communication). EcoRI cleaved cDm238 was ligated to EcoRI cleaved pBR322 and used to transform E.coli to ampicillin resistance. Colonies containing the 11.4Kb EcoRI fragment cloned in pBR322 were isolated by the colony hybridisation procedure. Nick translated ^{32}P pDm103 was used as a probe. The pSC101 vector of this plasmid hybridises weakly to pBR322, but a far stronger hybridisation signal will be seen if the colony contains the desired recombinant plasmid due to the hybridisation of Dm103 to the 11.4Kb EcoRI fragment. DNA was prepared from the colonies which gave strong hybridisation signals and was then cleaved with EcoRI. One plasmid, pDm238/29, gave an 11.4Kb fragment which comigrated with that of cDm238 (Figure 24 track 2) plus the pBR322 vector. pDm238/29 was chosen as a recombinant plasmid containing the 11.4Kb EcoRI fragment of cDm238.

3. Subcloning the 17Kb EcoRI fragment of cDm103

cDm103 contains the 17Kb fragment Dm103 inserted into the vector ColE1. Dm103 was cloned in pBR322 by ligating EcoRI cleaved pBR322 and cDm103 together and using it to transform E.coli to ampicillin resistance. Transformants containing Dm103 inserted in pBR322 were identified by the

Figure 24 Indentification of recombinant plasmids containing the 11.4Kb EcoRI fragment of cDm238 and the 17Kb EcoRI fragment of cDm103

EcoRI digests of cDm238 (track 1) pDm238/29 (track 2) and pBR322 (track 3) electrophoresed on a 0.8% agarose gel. cDm103 cleaved with Bst I/EcoRI (track 4) and EcoRI (track 7), pcDm103/1 cleaved with Bst I/EcoRI (track 5) and EcoRI (track 8) and pBR322 cleaved with Bst I/Eco RI (track 6) and EcoRI (track 9) electrophoresed on a 0.6% agarose gel. The sizes in Kb of relevant restriction fragments are indicated alongside tracks 1, 4 and 17. The samples in tracks 5 and 8 are from cleared lysates.

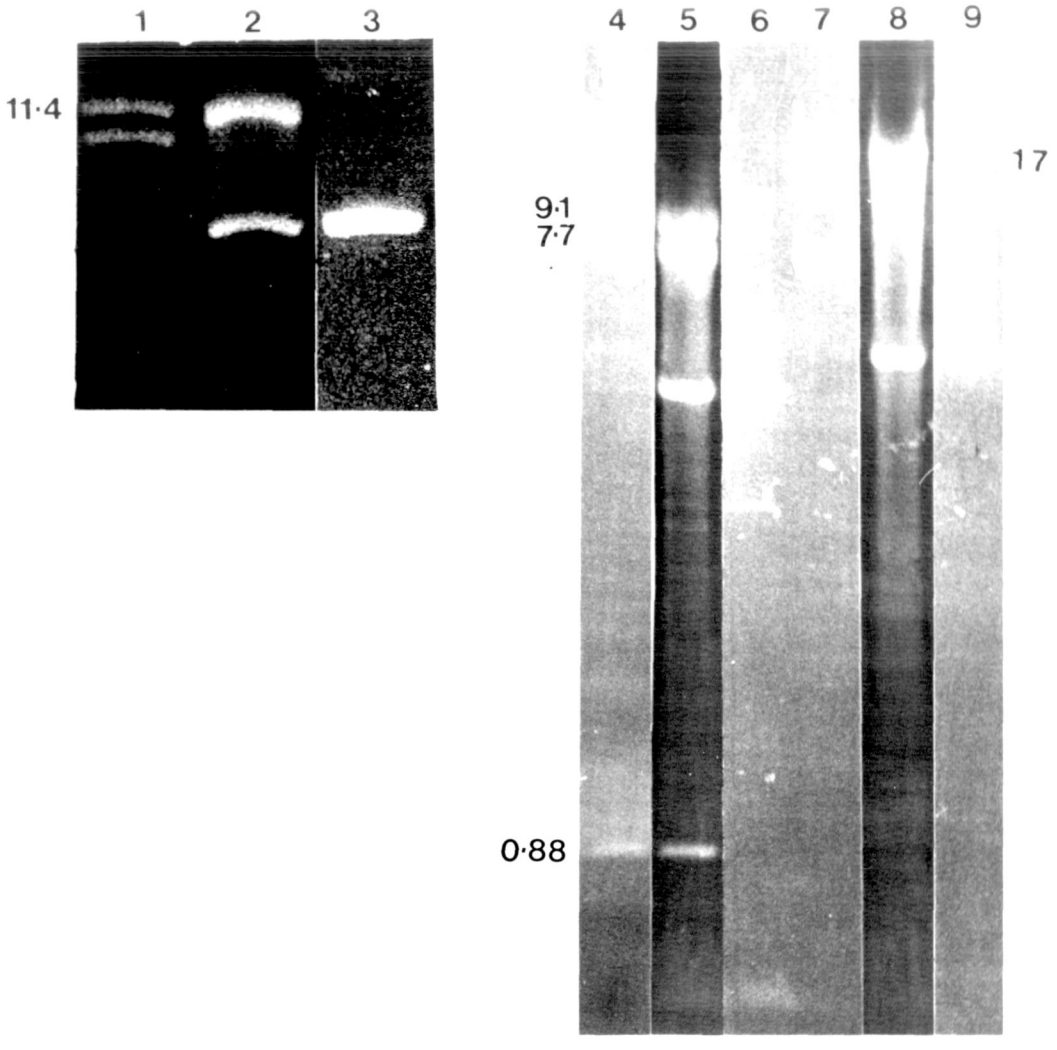


Figure 24.

Figure 25 Identification of recombinant plasmids containing the subcloned Bst I and Bst I/Hind III fragments of cKDm103/B

Bst I/Hind III digests of cKDm103/B (tracks 1 and 6), pC2 (track 2), pC4 (track 5) and pBR322 (tracks 3 and 4) electrophoresed on 0.6% (tracks 1 to 3) and 1.5% (tracks 4 to 6) agarose gels. Bst I cleaved pBR322 (track 7), pC25 (track 8) and cKDm103/B (track 9) electrophoresed in a 1% agarose gel. The sample in track 8 is from a cleared lysate. The positions of the 4.5 and 0.32 Kb Hind III/Bst I and 0.88Kb Bst I fragments are indicated.

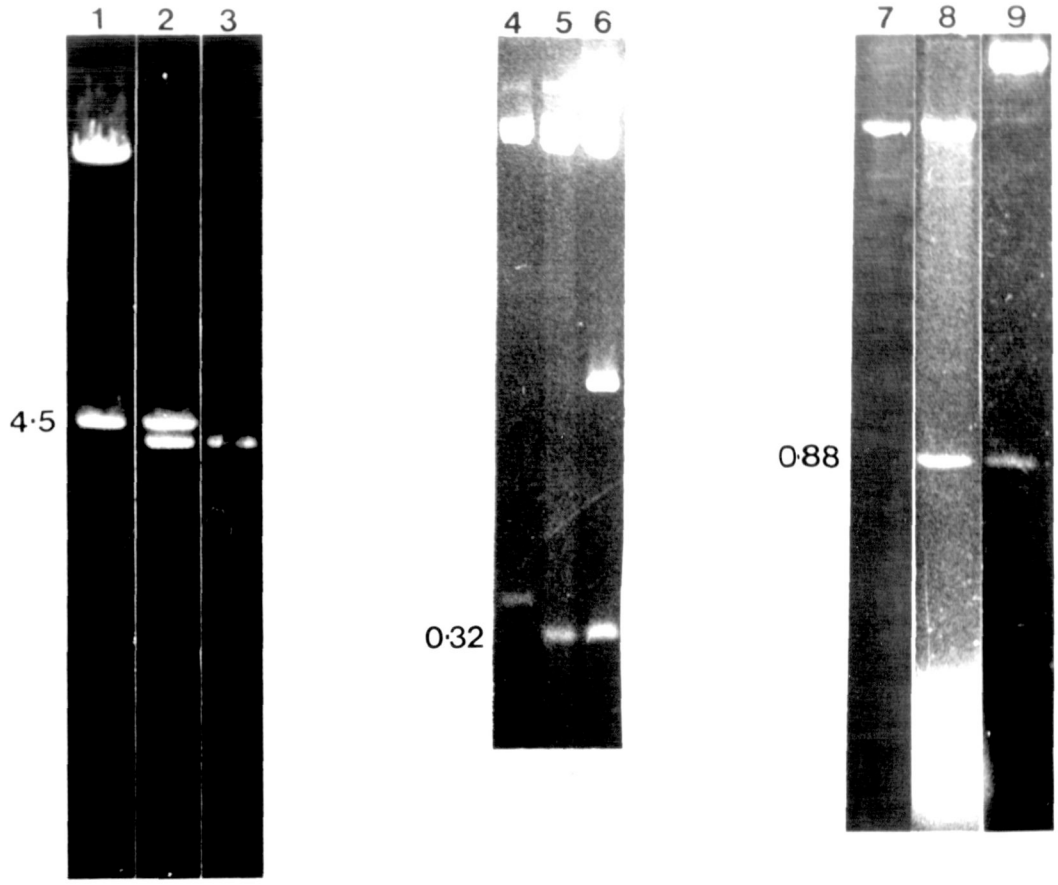


Figure 25.

colony hybridisation procedure.

The 0.88Kb Bst I fragment of Dml03 was used as a probe. It was gel purified and nick translated with ^{32}P. Figure 24 (tracks 4 to 9) shows EcoRI and EcoRI/Bst I digestions of DNA prepared from one of the colonies which hybridised. pcDml03/1 (Figure 24 track 7) was chosen as an example of Dml03 cloned in pBR322. It is normally referred to as pcDml03.

4. Subcloning the HindIII/Bst I fragments of cKDml03/B

Hind III/Bst I digestions of cKDml03/B and pBR322 were ligated together and used to transform E.coli to ampicillin resistance. DNA was prepared from ampicillin resistant, tetracycline sensitive colonies. Figure 25 show Hind III/Bst I digests of cKDml03/B, pBR322 and two of the tetracycline sensitive plasmids. pC2 (Figure 25 track 2) contains the 4.5Kb Bst I/Hind III fragment of cKDml03/B, while pC4 (Figure 25 track 5) contains the 0.32Kb fragment.

5. Subcloning the 0.88Kb Bst I fragment of cKDml03/B

Bst I cleaved pBR322 and cKDml03/B were ligated together and used to transform E.coli to ampicillin resistance. DNA was prepared from tetracycline sensitive colonies and cleaved with Bst I. The plasmid pC25 was chosen as containing the 0.88Kb Bst I fragment of cKDml03/B (Figure 25 track 8).

APPENDIX II

RESTRICTION MAPPING cDm103 AND cDm219

The generation of the restriction maps of cDm103 and cDm219 shown in Figure 8 and of pB74 in Figure 12 is described in this Appendix.

1. The Sst I and Kpn I sites of Dm103

The mapping of the Sst I and Kpn I sites in Dm103 was carried out relative to the previously determined sites for Hind III (Glover and Hogness, 1977), Xma I (B.R. Jordan and D. Glover, personal communication), Sal I and Bst I (D. Glover, personal communication). In the figures of this section these sites and the subclones of Dm103 are only shown where relevant to the mapping experiments.

Sst I sites of Dm103

Both cDm103 and cKDm103/B are cleaved five times by Sst I giving a similar restriction pattern with the exception of the largest fragment. As Sst I does not cleave the bacterial vectors, ColE1 or PML2, of these plasmids then all the Sst I sites are in Hind III fragment B of Dm103. Table 3 gives the sizes of the Sst I fragments of cKDm103/B, and Figure 26 the final Sst I map of Dm103/B.

Bst I cleaves Dm103/B 0.32 and 1.2Kb from its right hand Hind III site, resulting in a 0.88Kb Bst I fragment. An Sst I/Bst I digest of Dm103/B gives the 0.88Kb Bst I fragment plus a new one of 0.22Kb, placing an Sst I site 0.22Kb from the left most Bst I site. Sst I/Hind III digestion of Dm103 gives two new fragments of 1.5 and 0.29Kb. The 1.5Kb Sst I/Hind III fragment has to be due to the Sst I site closest to the right hand end of Dm103/B and this means that the 0.29Kb fragment shows the presence of an Sst I

TABLE 3 The sizes of the restriction fragments produced by Sst I digestion of cKDm103/B and Kpn I digestion of cDm103

Sst I digested cKDm103/B		Kpn I digested cDm103	
Fragment	Size (Kb)	Fragment	Size (Kb)
A	15.63	A	21.53
B	2.65	B	1.05
C	0.68	C	0.37
D	0.35	D	0.07
E	0.33		

Figure 26 The Sst I cleavage map of Dm103

All the Sst I sites of cDm103 are in Hind III fragment B, a restriction map of which is shown in the lower part of the figure, the thick lines indicate the 26s rRNA coding sequences. The new fragments arising from the cleavage of Sst I fragments B, C, E and D with other restriction endonucleases are indicated above the map, together with some of the partial Sst I digestion products. The combinations of restriction endonucleases used are shown beside their cleavage products.

Figure 26.

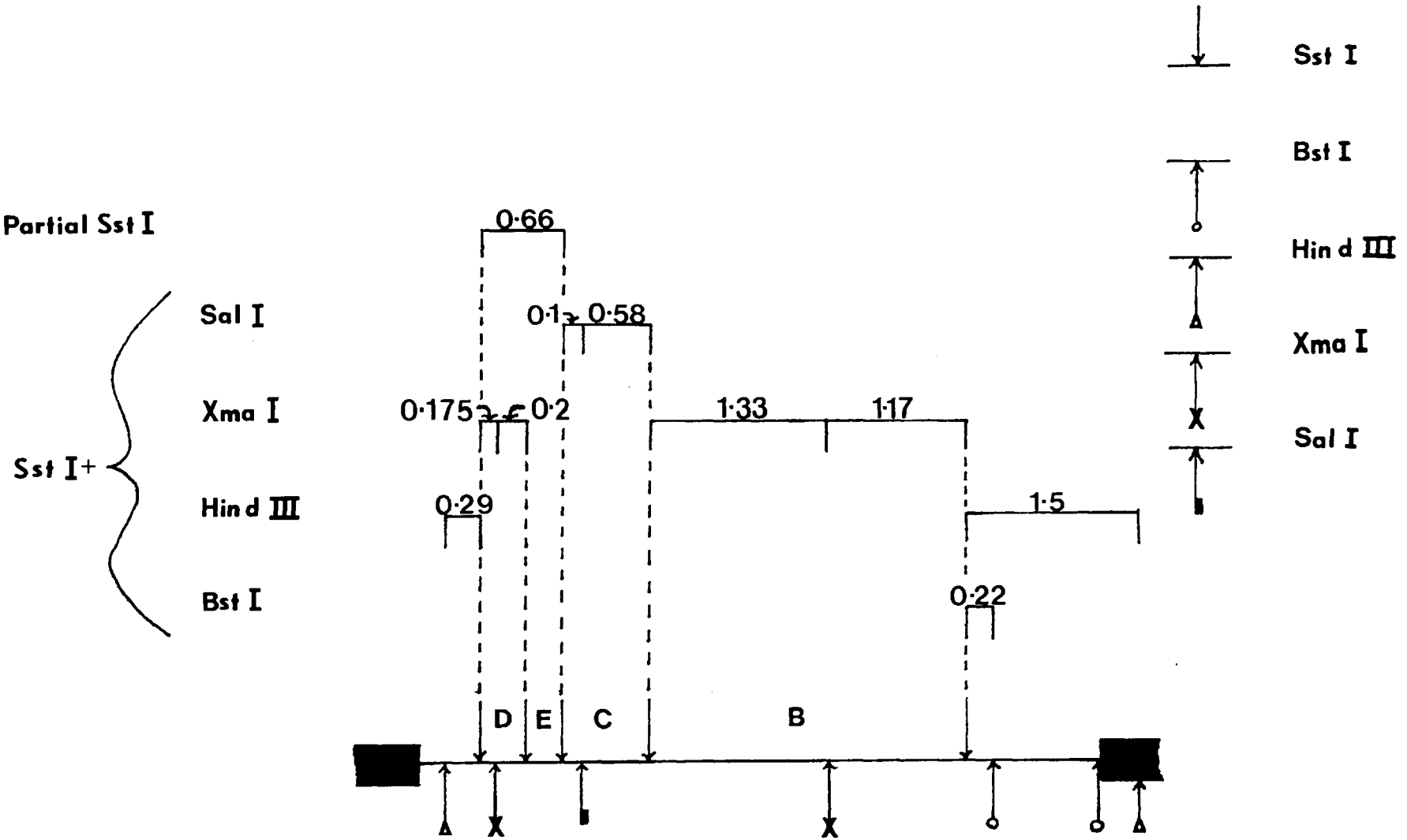
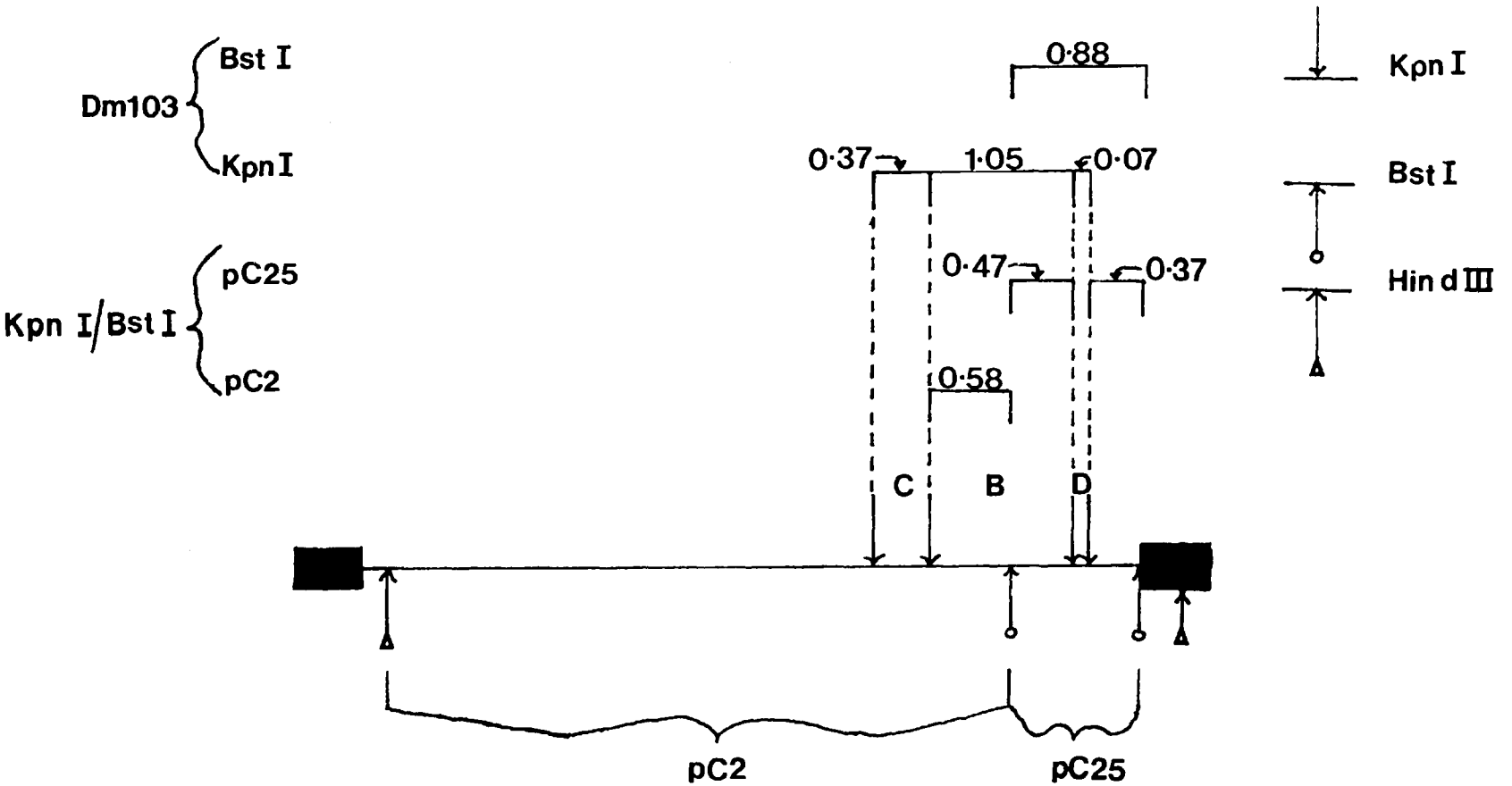

Partial Sst I
Sst I+
Sal I
Xma I
Hin d III
Bst I
0·66
0·1
0·58
0·175
0·2
1·33
1·17
0·29
1·5
0·22
D
E
C
B
Sst I
Bst I
Hin d III
Xma I
Sal I

Figure 27 The Kpn I cleavage map of Dm103

All of the Kpn I sites of cDm103 are in Hind III fragment B, a restriction map of which is shown in the lower part of the figure, the thick lines indicated the sequences coding for 26s rRNA. The sizes and positions of the restriction fragments resulting from Kpn I and Bst I cleavage of cDm103 and Kpn I/Bst I cleavage of pC25 and pC2 are indicated above the map.

Figure 27.


Bst I
Dm103
KpnI
pC25
Kpn I/Bst I
pC2
0·88
0·37
1·05
0·07
0·47
0·37
0·58
C
B
D
pC2
pC25
Kpn I
Bst I
Hin d III

site 0.29Kb from the left end of Dm103/B. Xma I cleaves Dm103/B 0.4 and 3.05Kb from its left hand Hind III site giving a fragment of 2.65Kb. Double digestion of Dm103/B with Xma I/Sst I results in cleavage of Sst I fragments B and D to give fragments of 1.17 and 1.33Kb, and 0.2 and 0.175Kb respectively. This places fragment D adjacent to the 0.29Kb Sst I/Hind III fragment. Fragment E is adjacent to D because a partial Sst I digestion results in a fragment of 0.66Kb, close to the expected sum of D and E (0.68Kb). Sal I cleaves Dm103/B 1.1Kb from its left hand end, an Sst I/Sal I digestion results in fragment C being cleaved into fragments of 0.58 and approximately 0.1Kb placing it adjacent to G. Fragment B must lie between fragment C and the 1.5Kb Sst I/Hind III fragment.

Kpn I sites of Dm103

Kpn I cleaves cDm103 4 times, resulting in the fragments shown in Table 3. A Kpn I/Bst I double digestion of cDm103 gives 3 new fragments of 0.58, 0.47 and 0.37 Kb, placing all the Kpn I sites in the vicinity of the two Bst I sites of the Dm103 insertion. The Kpn I sites were then mapped using the subcloned 4.5Kb Hind III/Bst I fragment (pC2) and 0.88Kb Bst I fragment (pC25) (Figure 27).

Kpn I cleavage of pC2 gives fragment C, further digestion with Bst I gives both this fragment and the 0.58Kb Kpn I/Bst I fragment. This places one Kpn I site 0.58Kb from the left of the left hand Bst I site of Dm103, adjacent to this must be fragment C. pC25 upon Kpn I digestion results in fragment D. Digestion with both Kpn I and Bst I gives Kpn I/Bst I fragments of 0.47 and 0.37 Kb. The sum of the 0.58 and 0.47 Kpn I/Bst I fragments is that of fragment B, resulting in the Kpn I map shown in Figure 27.

2. Restriction mapping cDm219

The basis of the restriction map of cDm219 shown in Figure 8 was the generation of an essentially unambiguous map for the restriction endo-

nucleases Bst I and EcoRI. Incomplete restriction maps could also be produced of cDm219 for several other enzymes. But their completion and the removal of the remaining ambiguities in the Bst I and EcoRI restriction maps was dependent on the use of subcloned restriction fragments of cDm219. The restriction fragments were cloned in pBR322 (material and methods, Appendix I) and were the result of cleaving cDm219 with EcoRI, Bst I or both of these enzymes. A restriction map of pBR322 is shown in Figure 7. As it is not cleaved by Kpn I, Sst I, or Xma I, all the sites for these enzymes in the subclones must occur in cDm219. Unless relevant restriction fragments containing pBR322 sequences have been ignored in describing the restriction mapping experiments.

Most of the results used to produce the final restriction map of cDm219 are given in a tabular form, additional results are also given in the text. As the relationships between restriction fragments shown in the tables were derived both from digestions of cDm219 and its subclones, the names of the subclones are also indicated, alongside the sizes of the EcoRI, Bst I and EcoRI/Bst I digestion products of cDm219. In describing the restriction mapping of cDm219 its three EcoRI sites are taken as a central reference point, and the left and right ends of cloned restriction fragments refer to their orientation in cDm219.

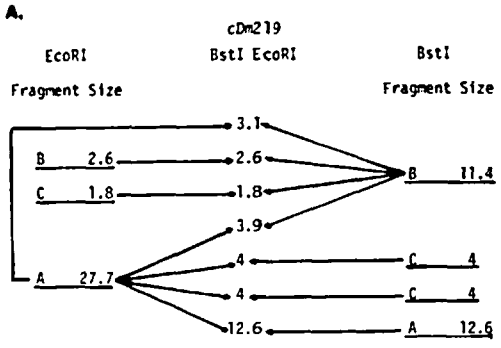
The EcoRI and Bst I sites of cDm219 and the positioning of the ColEI vector

The sites of the EcoRI and Bst I fragments of cDm219, together with the products of digestion with both of these enzymes is given in Figure 28A. Bst I cleavage of cDm219 results in three fragments when the digest is fractionated in an agarose gel. But upon partial Bst I digestion an 8Kb fragment results, indicating that the 4Kb Bst I fragment C is a doublet. As the two 4Kb Bst I fragments are adjacent (giving the 8Kb partial fragment) the linear map for

Figure 28 Mapping the Bst I and EcoRI sites of cDm219

(A) The sizes (in Kb) of the EcoRI, Bst I and EcoRI/Bst I fragments of cDm219. The arrows indicate how the double digestion products originate from the fragments resulting from digestion with a single enzyme.

(B) The Bst I and EcoRI cleavage map of cDm219. The horizontal order of the fragments is the same as the vertical order in (A). The orientation of EcoRI fragments B and C, the 2.9 and 3.1 Kb EcoRI/Bst I fragments, and of the two 4 Kb Bst I fragments was determined as described in the text. The names of the subcloned restrcition fragments of cDm219 are shown below the restriction map.


A.
cDm219
EcoRI
BstI EcoRI
BstI
Fragment Size
Fragment Size
3.1
B 2.6
2.6
B 11.4
C 1.8
1.8
3.9
4
C 4
A 27.7
4
C 4
12.6
A 12.6

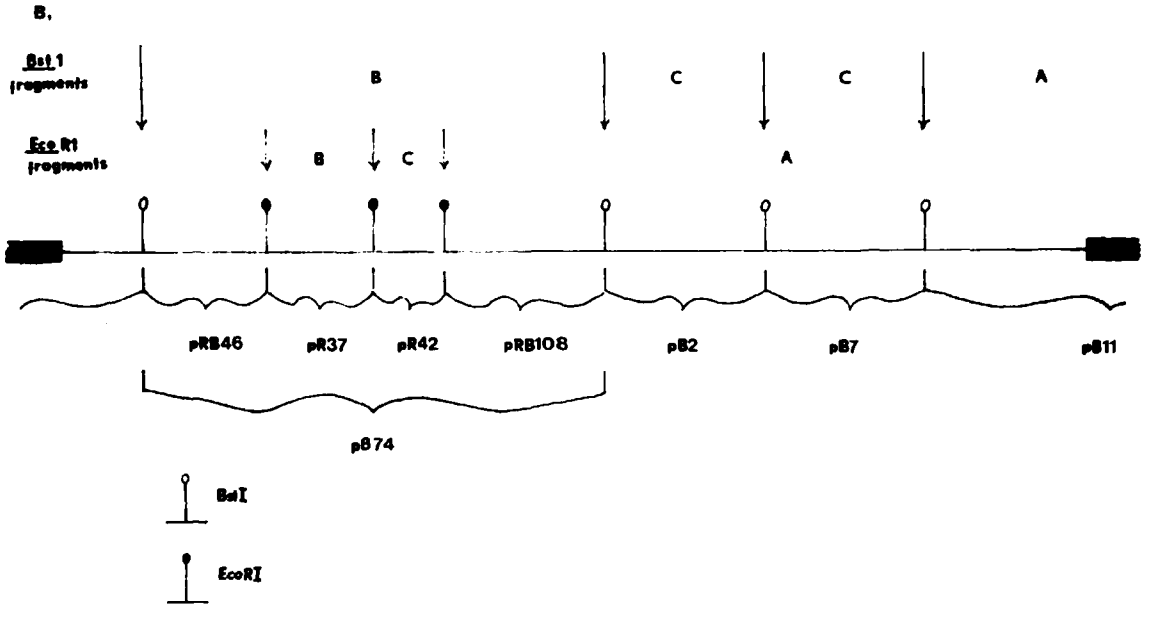

B.
Bst 1 fragments
Eco RI fragments
B
C
C
A
B
C
A
pRB46
pR37
pR42
pRB108
pB2
pB7
pB11
pB74
BstI
EcoRI

Figure 28.

the Bst I fragments in A B C C. A Bst I/EcoRI digestion of cDm219 results in Bst I fragment B being cleaved to 3.9 and 3.1Kb fragments together with EcoRI fragments B and C. The orientation of the 3.9 and 3.1Kb Bst I/EcoRI fragments was determined by further digestion with Xma I.

The sizes of the Xma I fragments of cDm219 is shown in Figure 32A. Bst I cleaves Xma I fragment A into fragments of 8.25 and 7.25Kb, which must lie in the 12.6 and 11.4Kb Bst I fragments. The orientation of these two Bst I/Xma I fragments was determined by cleaving cDm219 with these enzymes together with EcoRI which only cleaves in the 11.4Kb Bst I fragment. This triple digestion of cDm219 results in the 8.25Kb Bst I/Xma I fragment being cleaved by EcoRI into the 2.6 and 1.8Kb EcoRI fragments, the 3.1Kb EcoRI/Bst I fragment and a 1.6Kb EcoRI/Xma I fragment. This establishes the 8.25Kb Xma I/Bst I fragment lies in the 11.4Kb Bst I fragment, and that the 3.1Kb Bst I/EcoRI fragment lies to the left of the EcoRI sites as the 3.9 Kb fragment is cleaved by Xma I which has to cleave to the right of the EcoRI sites. The triple digestion of cDm219 also positions its ColEI vector. ColEI has no Bst I sites and in cDm219 its EcoRI site was destroyed by the dA:dT tails used in its construction. It has a single Xma I site, digestion of the ColEI L plasmid with Xma I and EcoRI results in fragments of 5.2Kb (ColEI L) and 1.35Kb (ColEI S). The only Xma I fragment of cDm219 large enough to contain ColEI L is fragment A. As shown above this fragment is cleaved by Bst I into fragments of 8.25 and 7.25Kb, the 8.25Kb fragment was shown to contain cDm219's three EcoRI sites and so ColEI L must be in the 7.25Kb Xma I/Bst I fragment, which is located in Bst I fragment A. The remaining portion of ColEI (ColEI S) then occurs in the Xma I fragment adjacent to fragment A.

The orientation of the 2.6 and 1.8Kb EcoRI fragments is described in the section on the Sst I and Xma I sites of cDm219. The data from these experiments is included in Figure 28B which shows the EcoRI and Bst I

cleavage sites of cDm219. All of the products of single or double digestions of cDm219 with these enzymes were subcloned in pBR322. The names of the subclones are indicated below the restriction map (Figure 28B).

The Kpn I sites of cDm219

Figure 29A gives the results of cleaving cDm219 with Kpn I alone or in various combination with EcoRI and Bst I.

Figure 30 shows the positions the EcoRI and Bst I sites in fragments B and C. EcoRI cleaves fragments B and C, resulting in fragments of 3.35, 2.6, 1.8, 1.6 and 1.03Kb in length. The 2.6 and 1.8Kb fragments comigrate with the two small EcoRI fragments of cDm219. But Kpn I/EcoRI digestion of the cloned EcoRI fragments of cDm219 shows that the 2.6Kb EcoRI fragment contains a Kpn I site resulting in fragments of 1.6 and 1.03Kb. This indicates that the 3.35 and 2.6Kb Kpn I/EcoRI fragments are due to Kpn I sites to the left and right of cDm219's EcoRI sites. The orientation of these two Kpn I sites was determined with Bst I. This does not cleave fragment B but does cleave fragment C to 4.8 and 0.37 Kb, which means that either the 3.35 or the 2.6Kb Kpn I/EcoRI fragments will be cleaved by Bst I. This as expected from its size was the 3.35Kb fragment which was digested by Bst I to the 3.1Kb Bst I/EcoRI fragment. This places Kpn I fragment C to the left of fragment B. As only the 2.6Kb EcoRI fragment is cleaved by Kpn I it must overlap fragments C and B, and as fragment C is only cleaved once by EcoRI (giving the 3.35 and 1.6Kb fragments) then the 1.8Kb EcoRI fragment must be in fragment B placing it to the right of the 2.6Kb EcoRI fragment. Figure 30 shows the order of fragments B and C and of the 2.6 and 1.8Kb EcoRI fragments.

Bst I cleaves fragments A C and D resulting in the following new fragments of 10.4, 4.8, 3.55 and 0.37 Kb. Fragment C has already been

Figure 29 Mapping the Kpn I sites of cDm219

(A) The results of single, double, or triple digests of cDm219 with the restriction endonucleases Kpn I, EcoRI and Bst I. The arrows indicate the result of cleaving single digestion products with a second or third enzyme.

(B) The Kpn I map of cDm219, the horizontal order of the fragments is the same as the vertical order in (A).

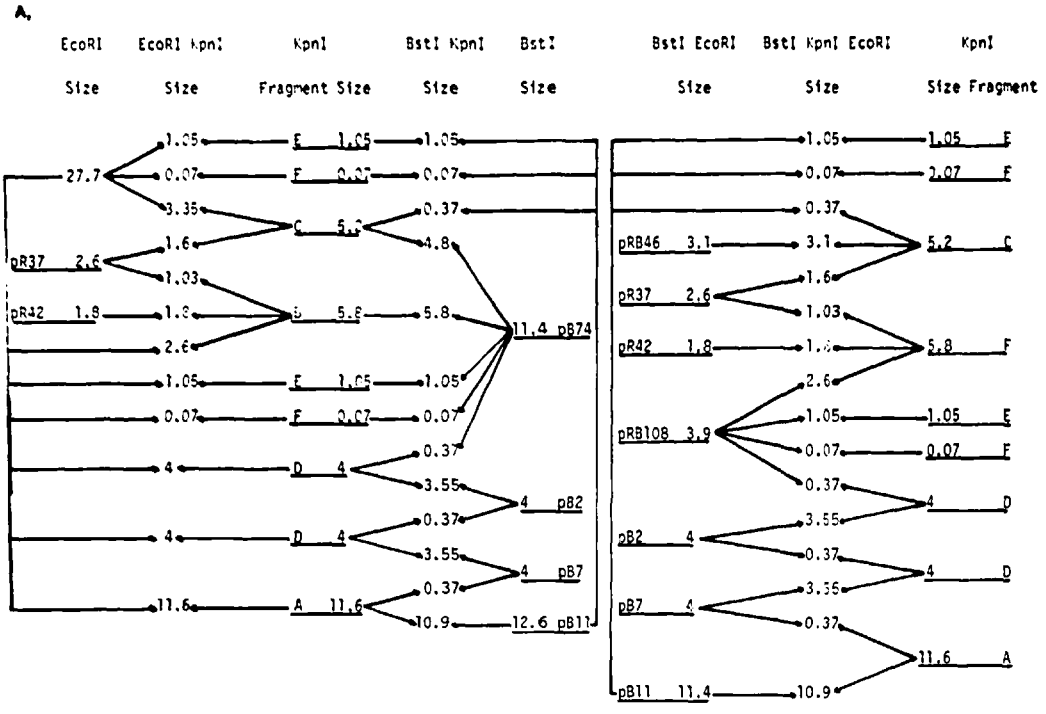
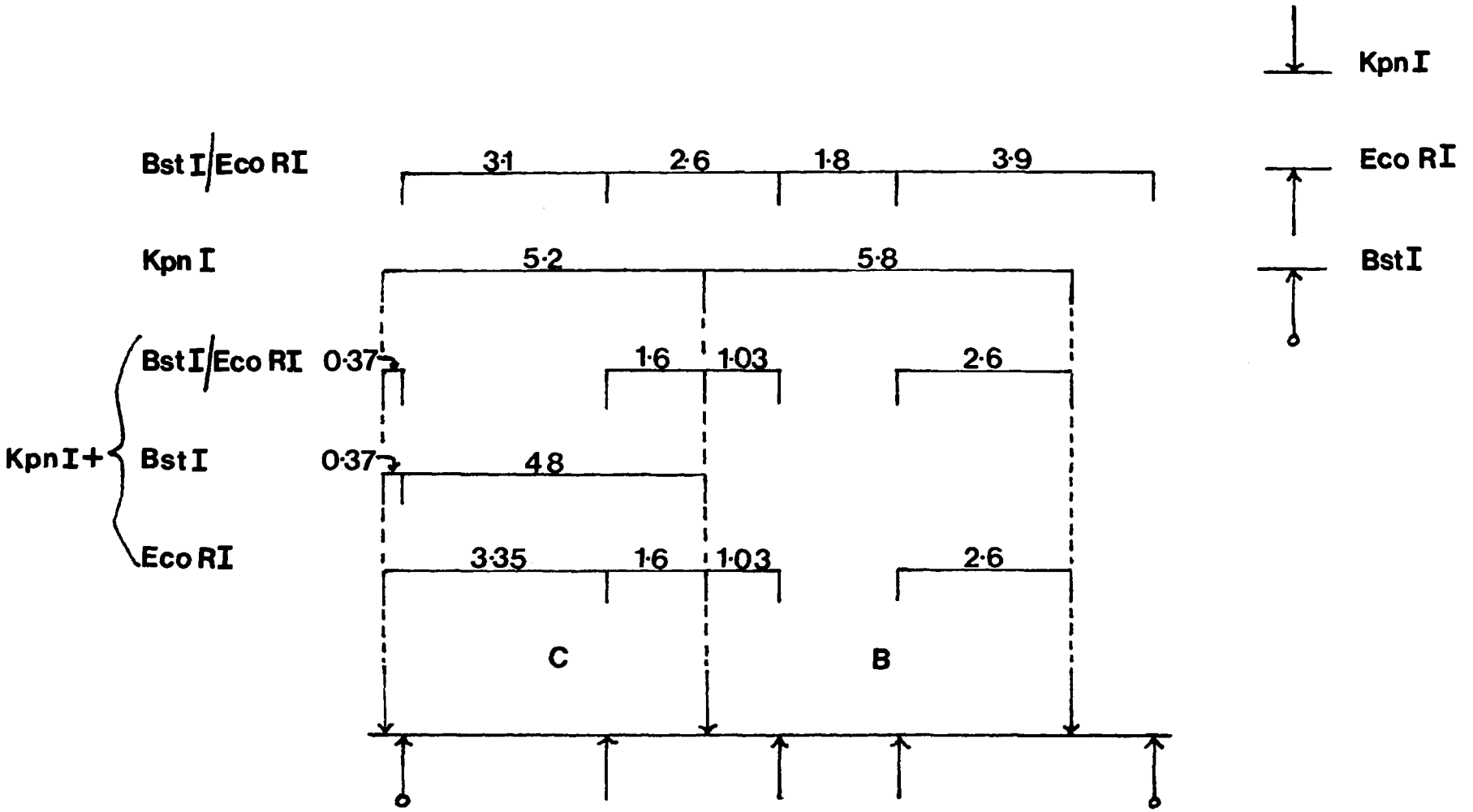

A.
EcoRI Size
EcoRI KpnI Size
KpnI Fragment Size
BstI KpnI Size
BstI Size
BstI EcoRI Size
BstI KpnI Size
EcoRI KpnI Size Fragment
27.7
pR37 2.6
pR42 1.8
11.6
4
4
1.05
0.07
3.35
1.6
1.03
2.6
1.05
0.07
4
4
11.6
E 1.05
F 0.07
C 5.2
B 5.8
E 1.05
F 0.07
D 4
D 4
A 11.6
0.37
4.8
5.8
1.05
0.07
0.37
3.55
0.37
3.55
0.37
10.9
11.4 pB74
4 pB2
4 pB7
12.6 pB11
pRB46 3.1
pR37 2.6
pR42 1.8
pRB108 3.9
pB2 4
pB7 4
pB11 11.4
1.05
0.07
0.37
3.1
1.6
1.03
2.6
1.05
0.07
0.37
3.55
0.37
3.35
0.37
10.9
5.2 C
5.8 B
1.05 E
0.07 F
4 D
4 D
11.6 A
B.
E F C B E F D D A
pRB46 pR37 pR42 pRB108 pB2 pB7 pB11
KpnI

Figure 29.

Figure 30 The Kpn I sites of unit 2 and its surrounding sequences in cDm219

The EcoRI and Bst I cleavage sites of unit 2 are shown on the lower lines together with the positions of Kpn I fragments B and C. Above this map is indicated the sizes and positions of fragments arising from Kpn I cleavage of EcoRI, Bst I or EcoRI/Bst I fragments.

Figure 30.


Kpn I
Eco RI
Bst I
Bst I/Eco RI
3·1
2·6
1·8
3·9
Kpn I
5·2
5·8
Bst I/Eco RI
0·37
1·6
1·03
2·6
Kpn I+
Bst I
0·37
4·8
Eco RI
3·35
1·6
1·03
2·6
C
B

positioned, by Bst I into a 4.8Kb fragment which was shown to be at the left end of the 11.4Kb Bst I fragment. A small part (0.37Kb) then must be in the 12.6Kb Bst I fragment. This is also the only Bst I fragment large enough to give the 10.9Kb Kpn I/Bst I fragment derived from Kpn I fragment A. The remaining portion of this Kpn fragment then must occur in the right hand 4Kb Bst I fragment giving a 0.37Kb Bst I/Kpn I fragment. These two Bst I fragments also contain the 4Kb Kpn fragment which must occur twice in cDm219. They are cleaved to fragments of 3.55 and 0.37 Kb by Bst I. It indicates that one overlaps the two 4Kb Bst I fragments, 3.55 Kb in the right hand one and 0.37Kb in the left. While the other overlaps with the left hand one and the 11.4Kb Bst I fragment, the 3.55Kb fragment occuring in the former and the 0.37Kb fragment in the latter Bst I fragment.

The positioning of the above Kpn I sites was confirmed and fragments E and F positioned by mapping Kpn I sites in the cDm 219 subclones. Kpn I cleavage of pRB108 produces fragments E and F, further digestion with Bst I gave the 0.37Kb Kpn I/Bst I fragment. Both pB2 and pB7 contain a single Kpn I site, digestion with Bst I as well gave fragments of 3.55 and 0.37Kb. The 3.55 Kb fragment of pB2 together with the 0.37 Kb fragment of pRB108 makes up one of 4Kb Kpn I fragments. While the 3.55Kb fragment of pB7 together with the 0.37Kb fragment of pB2 makes up the other. The 0.37Kb fragment of pB7 together with the 10.8Kb Kpn I/Bst I fragment of pB11 give fragment A. pB11 also gave Kpn I fragments E and F and the 0.37Kb Kpn/Bst I fragment derived from fragment C.

The orientation of fragments E and F was not determined but as they comigrate with the 1.05 and 0.07 Kpn I fragments of Dm103 (Figure 8) they have been positioned in the same orientation as in Dm103. This is supported by finding a 0.37Kb Kpn I/Bst I fragment in cDm219 which comigrates with that of Dm103. In Dm103 this fragment is adjacent to the 0.07Kb Kpn I fragment (fragment F of cDm 219). The fragment in cDm219 maps in the same position as in Dm103 supporting the argument that the fragments F and E occur as in Dm103. The Kpn I restriction map of cDm219 is shown in

Figure 29B.

Orientation of the subcloned Bst I fragments of cDm219

When using subcloned restriction fragments for restriction mapping cDm219, it was necessary in some cases to know the orientation of the inserted segment of cDm219 relative to restriction sites in the pBR322 vector. The restriction map of cDm219 for the enzymes Bst I, EcoRI and Kpn I could be constructed without knowing the orientation of the subclones. Also the restriction map of pBR322 is known (Figure 7). It has a single EcoRI site 0.348Kb from its singe Bst I site, and does not contain a Kpn I cleavage site. By knowing the cleavage sites for Kpn I and EcoRI relative to the Bst I sites of cDm219, and the position of the EcoRI site relative to the Bst I site of pBR322, the orientation of inserted cDm219 Bst I fragments into pBR322 could be determined.

The cDm219 Bst I fragments of the subclones pB2, pB7 and pB11 all contain a Kpn I site 0.37Kb away from their right hand Bst I site (Figure 29). Depending on the orientation of the inserted cDm219 Bst I fragment, a Kpn I/EcoRI digest of these subclones will give either a 0.77 or a 4.3Kb fragment. All of these subclones gave a 0.77Kb fragment, placing the right hand end of the inserted Bst I fragment nearest the pBR322 EcoRI site.

The subclone pB74 contains the 11.4Kb Bst I fragment, and has four EcoRI sites, three from the inserted Bst I fragment giving the 2.6 and 1.8 Kb EcoRI fragments of cDm219, and one in the pBR322 vector. Depending on the orientation of the inserted Bst I fragment the other two EcoRI fragments of pB74 will be either 3.5 and 7.9Kb or 4.3 and 7.1Kb. The latter was found indicating the left hand end of this Bst I fragment is nearest the pBR322 EcoRI site. Restriction maps of these far plasmids are shown in Figure 31, and their 5' to 3' orientation relative to the Dm103 insertion is indicated.

Figure 31 Restriction maps of plasmids containing the subcloned Bst I fragments, and the 3.1Kb EcoRI/Bst I fragment of cDm219

The restriction maps of various subcloned fragments are shown compared with that of the Dm103 insertion. The 5' to 3' orientation of the inserted Bst I fragments relative to that of Dm103 is indicated by the horizontal arrows. This was determined as described in the text, and the restriction fragments used for doing this are indicated below the restriction maps. Not all of the Hind III sites present in the cloned Bst I fragment of pB74 were mapped.

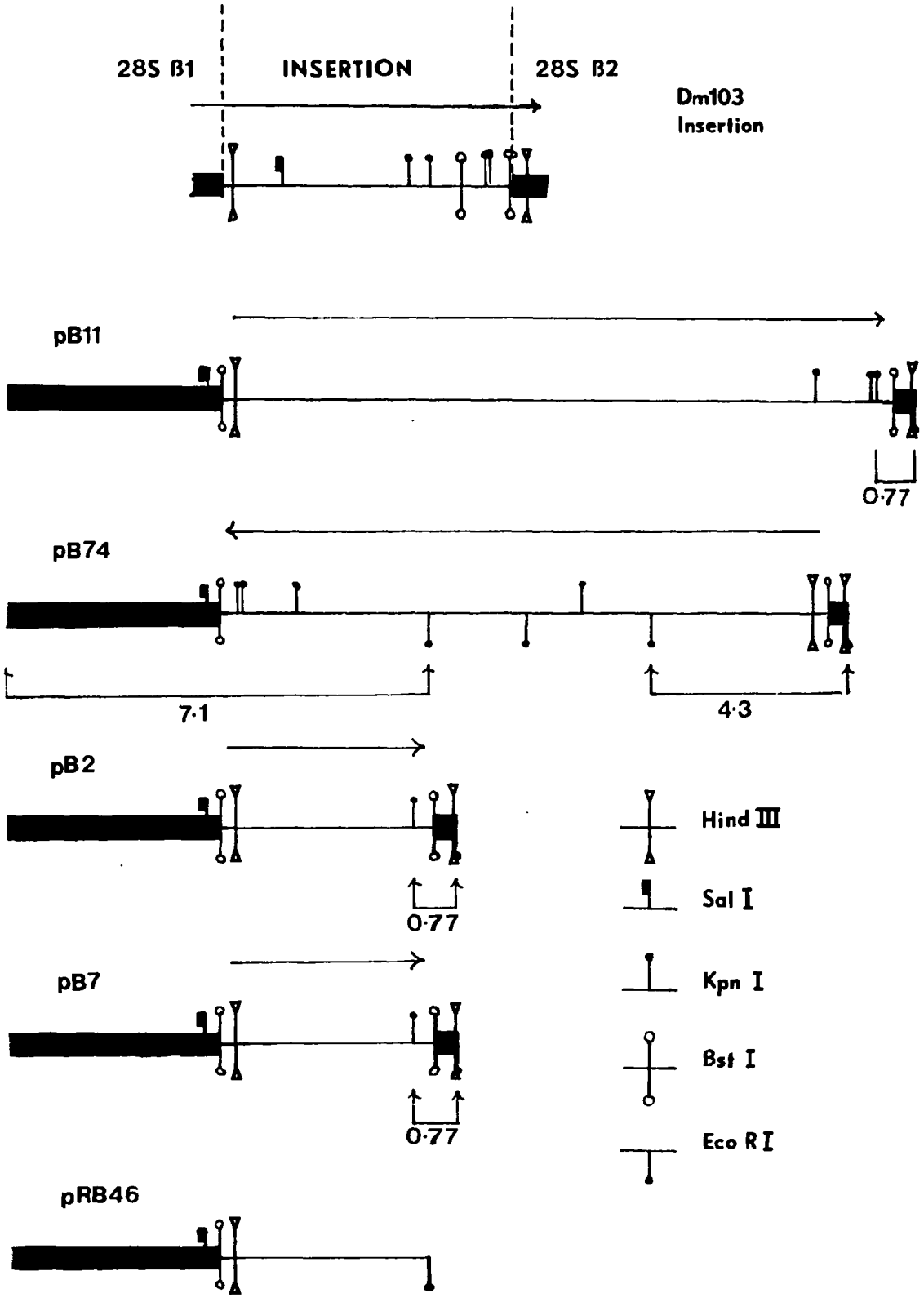

28S ß1
INSERTION
28S ß2
Dm103
Insertion
pB11
0·77
pB74
7·1
4·3
pB2
0·77
pB7
0·77
pRB46
Hind III
Sal I
Kpn I
Bst I
Eco R I

Figure 31.

The Xma I sites of cDm219

Figure 32A shows the results of Xma I cleavage of cDm219, and the results of furthercleavagewith Bst I and Bst I/EcoRI. Not shown is the result of Xma I/EcoRI digestion, which results in Xma I fragment A being cleaved to 10 and 1.6Kb Xma I/EcoRI fragments plus the 2.6 and 1.8Kb EcoRI fragments.

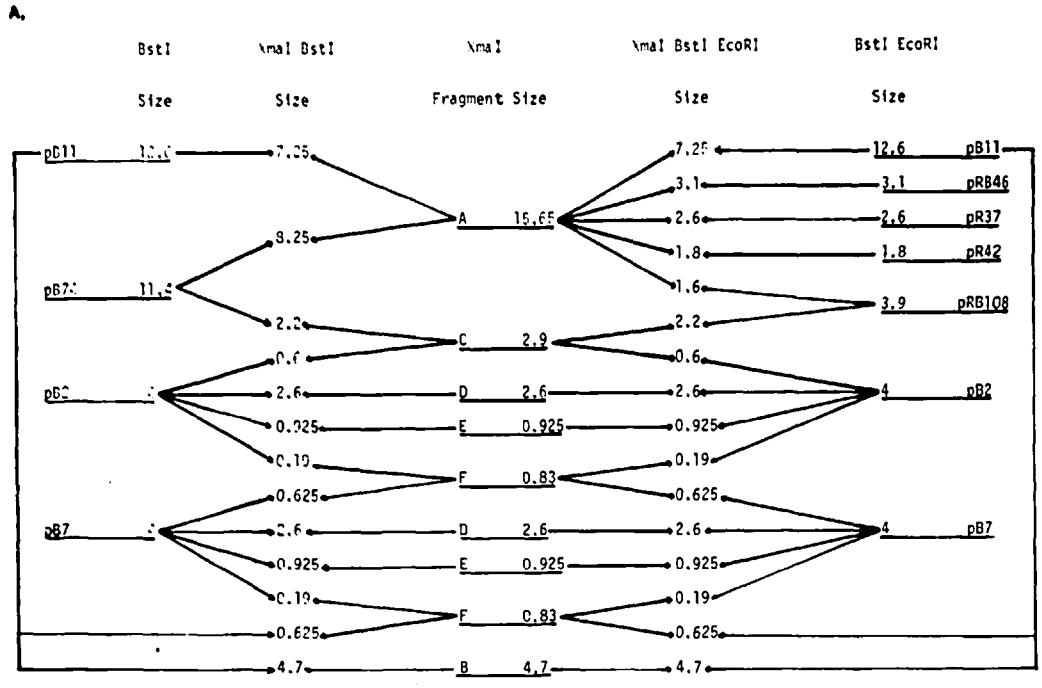
The positioning of Xma I fragment A has already been described. Fragment B must be adjacent to A as it is not cleaved by either Bst I or EcoRI and hence contains the remaining portion of ColEI (Col EI S). Fragment C is cleaved by Bst I to fragments of 2.2 and 0.6Kb. The 2.2Kb fragment corresponds to the remaining portion of the 11.4Kb Bst fragment, and so the 0.6Kb fragment must be found in the left hand 4Kb Bst I fragment. Fragment F is cleaved into fragments of 0.62 and 0.19Kb by Bst I, the 0.62Kb fragment corresponds to the remaining part of the 12.6Kb Bst I fragment.

The positions of the remaining Xma I sites was determined by using subclones of cDm219. Both pB7 and pB2 on Xma I digestion give fragments D and E. On Xma I/Bst I digestion, pB2 gives the 0.6Kb Xma I/Bst I fragment. This shows that pB2 contains the left hand 4Kb Bst I fragment and is adjacent to the 11.4Kb Bst I fragment. This has also been determined from the Sst map of cDm219. The 0.6Kb Xma I/Bst I fragment is derived from Bst I cleavage of fragment C. The other Xma I/Bst I fragment of pB2 was very small and its size was determined as follows; the orientation of the 4Kb Bst I fragment relative to the EcoRI site of pBR322 in pB2 is known and means that this Xma I site should be the closest one to the EcoRI site of pBR322. An Xma I/EcoRI digestion of pB2 gives a fragment of 0.55Kb, by subtracting the distance between the EcoRI and Bst I sites of pBR322; 0.348Kb, a size of 0.19Kb was arrived at for the Xma I/Bst I fragment of pB2. The plasmid pB7 gave similar results except that the 0.6Kb Xma I/Bst I fragment of pB2 was 0.62Kb in pB7. This fragment is derived from fragment F,

Figure 32 Mapping the Xma I sites of cDm219

(A) The sizes of fragments (in Kb) resulting from single, double or triple digests of cDm219 with the restriction endonucleases EcoRI, Bst I and Xma I. The arrows indicate the result of cleaving single digestion products with a second or third enzyme.

(B) The Xma I restriction map of cDm219, the horizontal order of the fragments is the same as the vertical order in (A). The orientation of fragments D and E was determined as described in the text.


A.
BstI Size
XmaI BstI Size
XmaI Fragment Size
XmaI BstI EcoRI Size
BstI EcoRI Size
pB11
12.6
7.25
8.25
A 15.65
7.25
3.1
2.6
1.8
1.6
12.6 pB11
3.1 pRB46
2.6 pR37
1.8 pR42
3.9 pRB108
pB7?
11.5
2.2
C 2.9
2.2
0.6
0.6
pB2
2.6
D 2.6
2.6
4 pB2
0.925
E 0.925
0.925
0.19
F 0.83
0.19
0.625
pB7
2.6
D 2.6
2.6
4 pB7
0.925
E 0.925
0.925
0.19
F 0.83
0.19
0.625
0.625
4.7
B 4.7
4.7

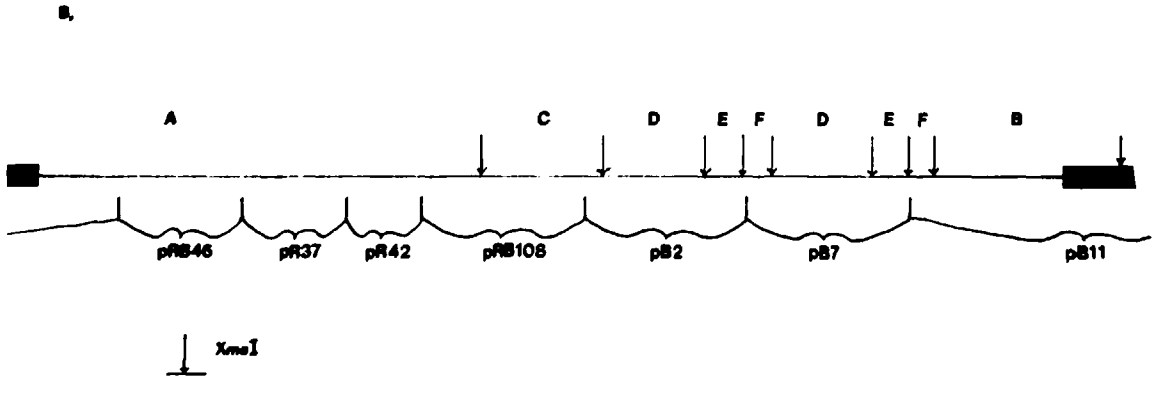

B.
A
C
D
E
F
D
E
F
B
pRB46
pR37
pR42
pRB108
pB2
pB7
pB11
XmaI

Figure 32.

the remaining part being the 0.19Kb Xma I/Bst I fragment of pB2. This Xma I/Bst I fragment is also found, in pB7 where it is part of fragment F which overlaps pB7 and the 12.6Kb Bst I fragment.

The orientation of fragments D and E within pB2 and pB7 was determined by Kpn I cleavage. This enzyme cleaves each 4Kb Bst fragment 0.37Kb from the right hand end and depending on their orientation will cleave either fragment D or E. A Xma I/Kpn I diegestion of either pB2 or pB7 results in fragment E being cleaved by Kpn I to a fragment of 0.76Kb, the other fragment migrated off the gel. This places fragment E to the right of D. The final Xma I map of cDm219 is shown in Figure 32B.

Sst I sites of cDm219

Figure 33A gives the results of cleaving cDm219 with Sst I, and the results of further digestion with EcoRI/Bst I. The result of Sst I/EcoRI digestions of cDm219 is described in the text.

The sum of the sizes of the Sst I fragment of cDm219 shown in Table 2 is 28.58Kb, a discrepancy of 6.54 Kb when compared with the contour length measurement of cDm219 (32.12 Kb) in the electron microscope. As shown below there are two copies of fragments C, E, I and J, and three copies of fragment F in cDm219. Allowing for this the sum of the fragments is 32.31 Kb, in close agreement with the length of cDm219.

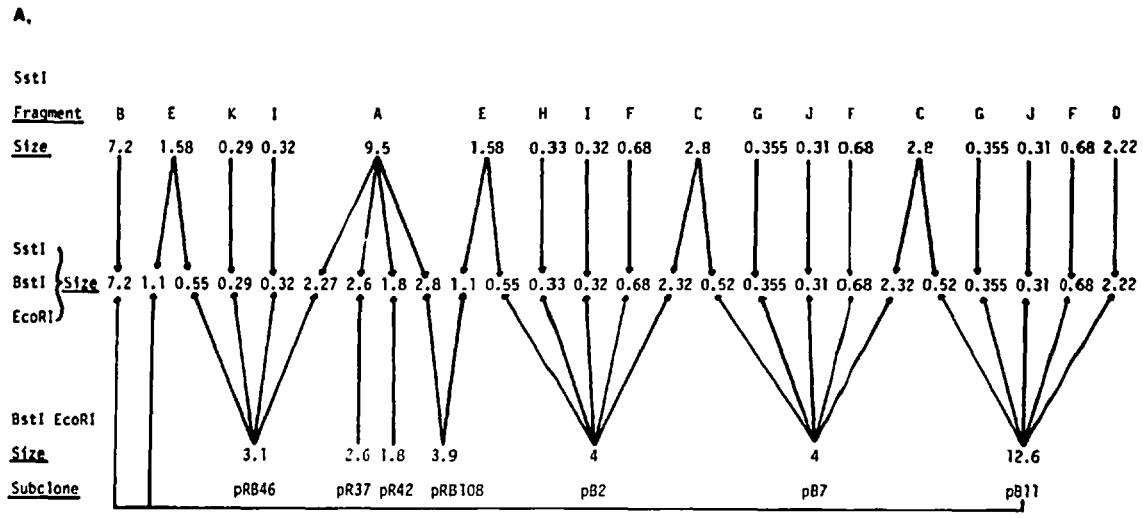
EcoRI cleaves only fragment A, resulting in the 2.6 and 1.8Kb EcoRI fragments, and EcoRI/Sst I fragments of 2.8 and 2.27Kb. The orientation of these two fragments was determind by Sst I/EcoRI cleavage of pRB46 and pRB108. The 2.27Kb fragment was found in pRB46 and the 2.8Kb fragment in pRB108. This places Sst I sites 2.27Kb to the left and 2.8Kb to the right of the EcoRI sites.

Bst I cleaves fragments C and E giving rise to fragments of 2.32, 1.1, 0.55 and 0.52Kb. Digestion of gel purified Sst I fragments with Bst I showed that the 2.32 and 0.52Kb fragments are from fragment C, and

Figure 33 Mapping the Sst I sites of cDm219

(A) The sizes of fragments (in Kb) resulting from single, double or triple digests of cDm219 with the restriction endonulceases EcoRI, Bst I and Sst I. The arrows indicate the result of cleaving single digestion products with a second or third enzyme.

(B) The Sst I map of cDm219, the horizontal order of the fragments is the same as that in (A).


A.
SstI
Fragment B E K I A E H I F C G J F C G J F D
Size 7.2 1.58 0.29 0.32 9.5 1.58 0.33 0.32 0.68 2.8 0.355 0.31 0.68 2.8 0.355 0.31 0.68 2.22
SstI
BstI Size 7.2 1.1 0.55 0.29 0.32 2.27 2.6 1.8 2.8 1.1 0.55 0.33 0.32 0.68 2.32 0.52 0.355 0.31 0.68 2.32 0.52 0.355 0.31 0.68 2.22
EcoRI
BstI EcoRI
Size 3.1 2.6 1.8 3.9 4 4 12.6
Subclone pRB46 pR37 pR42 pRB108 pB2 pB7 pB11

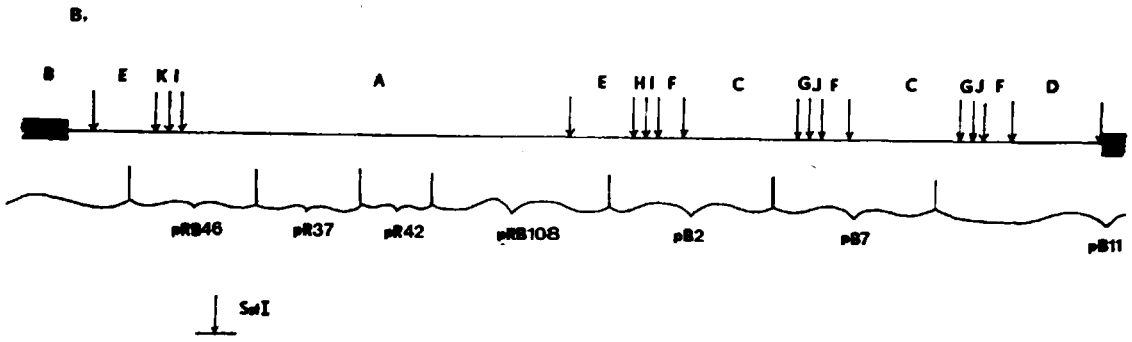

B.
B E K I A E HI F C GJ F C GJ F D
pRB46 pR37 pR42 pRB108 pB2 pB7 pB11
SmI

Figure 33.

and the 1.1 and 0.55 Kb fragments from fragment E.

Their position together with the rest of the Sst I sites was determined from digestions on subcloned fragments of cDm219. The 1.1Kb Sst I/Bst I fragment from fragment E was found in pRB108. The cloned segment in pRB108 is adjacent in cDm219 to one of the 4Kb Bst I fragments. Digestion with Sst I/Bst I of these cloned 4Kb segments showed that pB2 gave the 0.55Kb remaining of fragment E, and also fragments F, H and I, together with the 2.32 Kb double digestion fragment from fragment C. The remaining 0.52Kb portion of this fragment should be found in the 4Kb Bst I fragment adjacent to that of pB2. This is seen in an Sst I/Bst I digestion of pB7 together with fragments F, G and J and another 2.32Kb fragment. The 12.6Kb Bst I fragment of pB11 is adjacent to the 4Kb fragment of pB7. When digested with Sst I/Bst I pB11 gave the expected 0.52Kb fragment of fragment C. This digestion also gave fragments B, D, F, G and J as well as another 1.1Kb Sst I/Bst I fragment derived from fragment E. The remaining 0.55Kb was found in the adjacent 3.1Kb EcoRI/Bst I fragment cloned in pRB46 on Sst I/Bst I cleavage. Also found were fragments I and K.

The arrangement of the Sst I fragments which are not cleaved by either EcoRI or Bst I was determined by partial Sst I digestions or double digestions with Xma I. Both pB7 and pB11 have fragments F, G and J, and gave identical results upon partial Sst I digestion. The partial digestion products found were 0.64, 0.48 and 1.42Kb. The 0.64Kb fragment indicates J and G are adjacent, while the 0.98Kb fragment is close to the sum (0.99) for F and J being adjacent. The 1.42Kb fragment is F, J and G being linked by partial Sst I cleavage. This establishes their order as F J G. Their order relative to the left hand Bst I site was determined by Xma I/Sst I cleavage as Xma I cleaves 0.62 Kb from this Bst I site. Xma I cleaved fragment G, so the order relative to Bst I site is G J F. The partial Sst I fragments of pB2 were 0.63, 0.98 and 1.36Kb, as well as fragments H, I and

F. The sum of fragments H and I is 0.65Kb close to the size of the 0.63Kb partial digestion fragment and so they are adjacent in pB2. As the sum of either H and F, or I and F are so close (1.01 and 1.00Kb respectively) it is impossible to tell which of H or I is adjacent to F from the partial digestion. Their order has been placed H I F by analogy with the similar sized fragments of Dm103, pB2 and pB11. Their order relative to the left hand Bst I site of pB2 was also determined by Xma I cleavage. In this case Xma I cleaves 0.6Kb from this site, and cleaves either H or I, as the two fragments comigrated on this gel. This establishes the order in pB2 relative to the left hand Bst I site as H I F. In pRB46 fragments I and K have to be adjacent, their order relative to the 0.55Kb Sst I/Bst I fragment was obtained by a partial Sst I digestion followd by complete digestion with Bst I. In this experiment three partial fragments are seen, one of 0.55Kb due to I and K being adjacent, which comigrates with the Sst I/Bst I fragment, the others were of 0.795 and 1.15Kb. The former is close to the sum (0.84Kb) of K being linked to the 0.55Kb Sst I/Bst I fragment. The order then relative to the Bst Isite is K I. This is in contrast to the order of the equivalent fragments of Dm103, pB2 and pB11.

The 12.6Kb Bst I fragment of pB11 contains the ColEI vector of cDm219. ColEI has no Sst I cleavage sites. The only possible order for fragments B and D relative to fragment F of pB11 is F D B as this avoids the presence of Sst I sites in the ColEI vector. This allows it to be contained completely within fragment B. The final Sst I restriction map is shown in Figure 33B.

Hind III sites of cDm219

Only some of the Hind III sites of cDm219 were mapped. These were the four sites nearest the left hand end of each Bst I fragment. A Hind III/Bst I digest of cDm219 gives a 0.29Kb Hind III/Bst I fragment.

This could also be seen in Hind III/Bst I digests of pB2, pB11 and pRB46 (the left hand Bst I/EcoRI fragment of the 11.4Kb Bst I fragment) as well as the 0.346Kb Bst I/Hind III fragment of pBR322 (see Figure 7). In pRB46 the Bst I/EcoRI fragment of pBR322 containing its single Hind III site was deleted and replaced by the 3.1Kb Bst I/EcoRI fragment of cDm219. The Bst I/Hind III digest of pRB46 then places a Hind III site 0.24Kb from the left end of the 11.4Kb Bst I fragment of cDm219. In pB2, pB7 and pB11 the 0.29Kb fragment could come from either end on the inserted cDm219 Bst I fragment. The end it came from was determined by a Sal I/ Hind III digest of these plasmids, as the orientation of the inserted Bst I fragments relative to the pBR322 Sal I site was known and also Sal I does not cleave cDm219. The Sal I/Hind III digest of these plasmids resulted in a fragment of 0.57Kb, subtracting the distance between the Sal I and Bst I sites of pBR322 (0.276Kb, see Figure 7) places a Hind III 0.294Kb from the left hand Bst I site of pB2, pB7 and pB11. This Hind III site and the Sal I site of pBR322 are shown in the restriction maps of the sub-clones in Figure 31.

Pst I sites of pB74

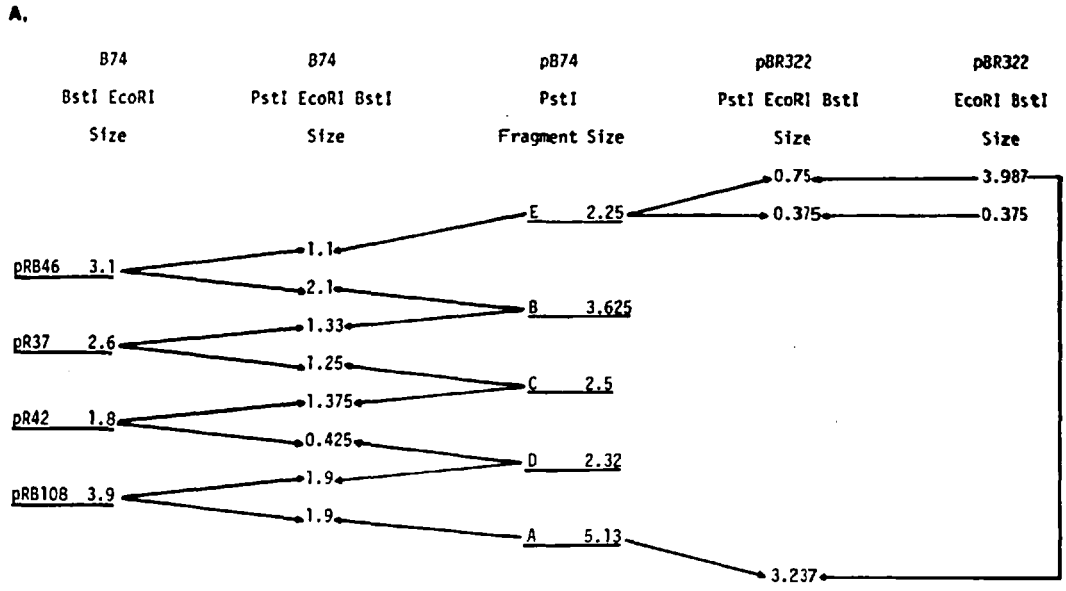
The results of cleaving pB74 are given in Figure 34. Also shown is the results of the experiments described below. In this summary fragments derived from the pBR322 vector and the 11.4Kb Bst I fragment of cDm219 are shown separately.

One of the Pst I sites of pB74 is in the pBR322 vector. A Pst I/ Bst I digest of pBR322 (see Figure 7) results in fragments of 1.123 and 3.237Kb. Double digestion of pB74 with Pst I/Bst I results in fragments A and E being cleaved resulting in the two vector fragments and fragments of 1.9 and 1.1Kb which had to be derived from the inserted Bst I fragment. The 1.1 and 1.9Kb Pst I/Bst I fragments are found in similar digests of

Figure 34 Mapping the Pst I sites in the subcloned 11.4Kb Bst I fragment in pB74

(A) The sizes of fragments (in Kb) resulting from single, double or triple digests of pB74 with the restriction endonucleases EcoRI, Bst I and Pst I. The fragments from the triple digestion of pB74 have been divided into those derived from the 11.4Kb Bst I fragment (B74) and those from the pBR322 vector.

(B) The Pst I map of pB74, the horizontal order of the fragments is the same as the vertical order in (A).


A.
B74 BstI EcoRI Size
B74 PstI EcoRI BstI Size
pB74 PstI Fragment Size
pBR322 PstI EcoRI BstI Size
pBR322 EcoRI BstI Size
pRB46 3.1
pR37 2.6
pR42 1.8
pRB108 3.9
1.1
2.1
1.33
1.25
1.375
0.425
1.9
1.9
E 2.25
B 3.625
C 2.5
D 2.32
A 5.13
0.75
0.375
3.237
3.987
0.375

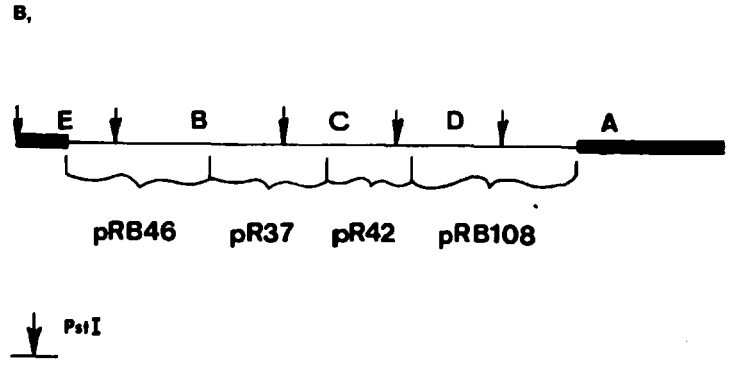

B,
E
B
C
D
A
pRB46
pR37
pR42
pRB108
PstI

Figure 34.

pRB46 and pRB108 respectively, placing Pst I sites 1.1 and 1.9Kb from the left and right ends of the 11.4Kb Bst I fragment.

EcoRI cleaves Pst I fragments B C D and E into fragments of 0.71, 1.47, 2.1, 1.33, 1.125, 1.375, 0.425 and 1.9Kb. The 0.71 Kb fragment is the smaller Pst I/EcoRI fragment of pBR322 (see Figure 7). These fragments were ordered by cleaving gel purified Pst I fragments with EcoRI and comparing the resulting fragments with Pst I/EcoRI digests of pRB46, pR37, pR48 and pRB108. Fragment B gave the 2.1 and 1.33Kb fragments which are found in pRB46 and pR37 respectively. Fragment C gave the 1.125 and 1.375 Kb fragments, the former was found in a Pst I/EcoRI digest of pR37 and the latter in pR42. Fragment D gave the 0.425 Kb fragment which is found in pR42 and one of 1.4 Kb found in pRB108. Finally fragment E on EcoRI digestion gave the 0.71Kb Pst I/EcoRI fragment of pBR322, and one of 1.475 which is the fragment from the pBR322 EcoRI site to the Pst I site 1.1Kb into the 11.4Kb Bst I fragment. In Figure 34A the 1.475Kb fragment is shown as if cleaved with Bst I resulting in a 0.348 Kb EcoRI/Bst I fragment from pBR322, and the 1.1Kb Bst I/Pst I fragment from the cDm219 Bst I fragment. The Pst I map of pB74 is shown in Figure 34B.

REFERENCES

ABELSON, J. (1979) Ann. Rev. Biochem. 48 1035-1069.

ALLET, B. and ROCHAIX, J-D. (1979) Cell. 18 55-60.

ASHBURNER, M. (1970) Advance Insect Physiol. 7 1-196.

BARNETT, T. and RAE, P.M.N. (1979) Cell 16 763-775.

BEDBROOK, J.R., KOLODNER, R. and BOGORAD, L. (1977) Cell 11 739-749.

BENTON, W.D. and DAVIS, R.W. (1977) Science 196 180-182.

BIRD, A.P. (1977) Cold Spring Harbour Symp. Quant. Biol. 42 1179-1183.

BIRNSTIEL, M.L. and GRUNSTEIN, M. (1972) FEBS Symp. 23 349-365.

BOLIVAR, F., RODRIGUEZ, R.L., GREENE, P.J., BETLACH, M.C., HEYNEKER, H.L., BOYER, H.W., CROSA, J.H., and FALKOW, S. (1977) Gene 2 95-113.

BORST, P. and GRIVELL, P.A. (1978) Cell 15 705-723.

BOS, J.L., HEYTING, L., BORST, P., ARNBERG, A.L. and VAN BRUGGEN, E.F.J. (1978) Nature 275 336-338.

BOSELEY, P.G., LUYNS, A. and BIRNSTIEL, M.L. (1978) Nucl. Acids Res. 5 1121-1137.

BOTCHAN, P., REEDER, R.H. and DAWID, I.B. (1977) Cell 11 599-607.

BRITTEN, R.J., GRAHAM, D.E. and NEUFIELD, B.R. (1974) In Methods in Enzymology 29, L. GROSSMAN and K. MOLDAVE, Eds. (New York: Academic Press) 563-418.

BROWN, D.D. and WEBER, C.J. (1968) J. Mol. Biol. 34 661-680.

CAMPBELL, G.R., LITTAU, V.L., MELERA, P.W., ALLFREY, V.L. and JOHNSON, E.M. (1979) Nucl. Acid Res. 6 1433-1447.

CHAMBON, P. (1975) Ann. Rev. Biochem. 44 613-639.

CHOOI, W.Y. (1979) Chromosoma. 74 57-74.

CHUA, N-H. and LUCK, D.J.L. (1974) In Ribosomes, M. NOMURA, A. TISSIERES and P. LENGYEL Eds. (New York: Cold Spring Harbour Laboratory) 519-539.

COCKBURN, A.F., TAYLOR, W.C. and FIRTEL, R.A. (1978) Chromosoma. 70 19-29.

CRAMER, J.H., FARRELLY, F.W. and ROWND, R.H. (1976) Mol. Gen. Genet. 148 233-241.

CRICK, F. (1979) Science 204 264-271.

CURTIS, P.J., MANTEI, N., VAN DEN BURG, J. and WEISSMANN, C. (1977) Proc. Nat. Acad. Sci. USA 74 3184-3188.

CURTIS, P.J. and WEISSMANN, C. (1976) J. Mol. Biol. 106 1061-1075.

DAVIDSON, N. and SZYBALSKI, W. (1971) In The Bacteriophage Lambola, A.D. HERSHEY Ed. (New York: Cold Spring Harbour Laboratory) 45-82.

DAVIS, R.W., SIMON, M. and DAVIDSON, N. (1971) In Methods in Enzymology, L. GROSSMAN and K. MOLDAVE Eds. (New York: Academic Press) 413-428.

DAWID, I.B. and BOTCHAN, P. (1977) Proc. Nat. Acad. Sci. USA 74 4233-4237.

DAWID, I.B. and LONG, E.O. (1979) In Specific Eukaryotic Genes, Alfred Benzon Symposium 13 (Copenhagen: Munksgaard) 103-117.

DAWID, I.B. and WELLAUER, P.K. (1976) Cell 8 443-448.

DAWID, I.B. and WELLAUER, P.K. (1977) Cold Spring Harbour Smp. Quant. Biol. 42 1185-1194.

DAWID, I.B., WELLAUER, P.K. and LONG, E.O. (1978) J. Mol. Biol. 126 749-768.

DEGELMANN, A., ROGER, H-D. and HOLLENBERG, C.P. (1979) Chromosoma 71 263-281.

DENHARDT, D.T. (1966) Biochem. Biophys. Res. Commun. 23 641-646.

DE ROBERTIS, E.M. and MERTZ, J.E. (1977) Cell 12 175-182.

DE ROBERTIS, E.M. and OLSEN, M.V. (1979) nature 278 137-143.

DIN, N., ENGBERG, J., FAFFENBERGER, W. and ECKHERT, W.A. (1979) Cell 18 525-532.

ENDOW, S.A. and GLOVER, D.M. (1979) Cell 17 597-608.

FEDOROFF, N.V. (1979) Cell 16 697-710.

FIERS, W., CONTRERAS, R., HAEGEMAN, G., ROGIERS, R., VAN DE VOORDE, A., VAN HEIVERSWYN, H., VAN VERREWEGHE, J., VOLCKAERT, G. and YSEBART, M. (1978) Nature 273 113-120.

FINNEGAN, D.J., RUBIN, G.M., YOUNG, M.W. and HOGNESS, D.S. (1977) Cold Spring Harbour Smp. Quant. Biol. 42 1053-1063.

FORSHEIT, A.B., DAVIDSON, N. and BROWn, D.D. (1974) J. Mol. Biol. 90 301-314.

GALL, J.G., COHEN, E.H. and POLEN, M.L. (1971) Chromosoma 33 319-344.

GERBI, S.A. (1976) J. Mol. Biol. 106 799-816.

GILBERT, W. (1978) Nature 271 501.

GLATZER, K.H. (1979) Chromosoma 75 161-175.

GLOVER, D.M. (1977) Proc. Nat. Acad. Sci. USA 74 4932-4436.

GLOVER, D.M. and HOGNESS, D.S. (1977) Cell 10 167-176.

GLOVER, D.M., WHITE, R.L., FINNEGAN, D.J. and HOGNESS, D.S. (1975) Cell 5 149-157.

GOLDBERG, M.L., LIFTON, R.P., STARK, G.R. and WILLIAMS, J.G. (1980) In Methods in Enzymology, In Press.

GOODMAN, H.M., OLSEN, M.V. and HALL, B.D. (1977) Proc. Nat. Acad. Sci. USA 74 5453-5475.

GRIMM, M.F. and LAMBOWITZ, A.M. (1979) J. Mol. Biol. 134 667-672.

GRUNSTEIN, M. and HOGNESS, D.S. (1975) Proc. Nat. Acad. Sci. USA 72 3961-3965.

HAGGERTY, D.M. and SCHLEIF, R.F. (1976) J. Virol. 18 654-663.

HAHN, V., LAZARUS, C.M., LUNSDORF, H. and KUNTZEL, H. (1979) Cell. 17 191-200.

HOLLENBURG, C.P. (1976) Chromosoma 57 185-197.

HSU, T.C., SPIROTO, S.E. and PARDUE, M-L. (1975) Chromosoma 53 25-36.

JEFFREYS, A.J. and FLAVELL, R.A. (1977) Cell 12 429-439.

JORDAN, E.G. (1979) Nature 281 529-530.

JORDAN, B.R. and GLOVER, D.M. (1977) FEBS Letters 78 271-274.

JORDAN, B.R., JOURDAN, R. and JACQ, B. (1976) J. Mol. Biol. 101 85-105.

JOVIN, T.M., ENGLUND, P.T. and BERTSCH, L.L. (1969) J. Biol. Chem. 244 2996-3008.

KARRIER, K.M. and GALL, J.G. (1976) J. Mol. Biol. 104 421-453.

KLECKNER, N. (1977) Cell 11 11-23.

KNAPP, G., BECKMAN, J.S., JOHNSON, P.F., FUHRMAN, S.A. and ABELSON, J. (1978) Cell 14 221-238.

KONKEL, D.A., MAIZELS, J.V. and LEDER, P. (1974) Cell 18 865-873.

KRIDER, H.M. and LEVINE, B.I. (1975) Genetics 81 501-513.

LASKEY, R.A. and MILLS, R.D. (1976) FEBS Letters 82 314-316.

LEFEVRE, Jr., G. (1976) In The Genetics and Biology of Drosophila, M. ASHBURNER and E. NOVITSKI 1A Eds. (London, New York, San Francisco: Academic Press) 31-65.

LEVIS, R., and PENMAN, S. (1978) J. Mol. Biol. 121 219-238.

LIS, J.T., PRESTIGE, L. and HOGNESS, D.S. (1978) Cell 14 901-919.

LOENING, V.E. (1968) J. Mol. Biol. 38 355-365.

LOENING, V.E., JONES, K.W. and BIRNSTIEL, M.L. (1969) J. Mol. Biol. 45 353-366.

LONG, E.O. and DAWID, I.B. (1979a) Nuc. Acid Res. 7 205-215.

LONG, E.O. and DAWID, I.B. (1979b) Cell 18 1185-1196.

MADEN, B.E.H. and TARTOF, K. (1974) J. Mol. Biol. 90 51-64.

MAIZELS, N. (1976) Cell 9 431-438.

MANDEL, M. and HIGA, A. (1970) J. Mol. Biol. 53 159-162.

MANNELA, L.A., COLLINS, R.A., GREEN, M.R. and LAMBOWITZ, A.M. (1979) Proc. Nat. Acad. Sci. USA 76 2635-2639.

MANNING, J.E., SCHMID, L.W.and DAVIDSON, N. (1975) Cell 4 141-155.

McKNIGHT, S.L. and MILLER, Jr., O.L. (1976) Cell 8 305-319.

MILLER, Jr., O.L. and BEATTY, B.R. (1969) Science 164 955-957.

MURREY, N.E., BRAMMAR, W.J. and MURREY, K. (1977) Molec. Gen. Genet. 150 53-61.

MURREY, N.E., BRUCE, S.A. and MURREY, K. (1979) J. Mol. Biol. 132 493-505.

O'FARREL, P.Z., CORDELL, B., VALENZUALA, P., RUTTER, W.J. and GOODMAN, H.M. (1978) Nature 274 438-445.

PARDUE, M-L. (1973) Cold Spring Harbour Smp. Quant. Biol. 38 475-482.

PARDUE, M-L., GERBI, S.A., ECKHARDT, R.A. and GALL, J.G. (1970) Chromosoma 29 268-290.

PATTERSON, J.T. and STONE, W.S. (1952) Evolution in the Genus Drosophila (New York: The Macmillan Company).

PEACOCK, W.J., LOHE, A.R., GERLACH, W.L., DUNSMUIR, P., DENNIS, E.S. and APPELS, R. (1977) Cold Spring Harbour Smp. Quant. Biol. 42 1121-1135.

PEARLMAN, R.E., ANDERSON, P., ENGBERG, J. and NILSSON, J.R. (1979). In Specific Eukaryotic Genes, Alfred Benzon Symposium 13 (Copenhagen: Munksgaard).

PELLEGRINI, M., MANNING, J. and DAVIDSON, N. (1977) Cell 10 213-224.

POTTER, S.S., BROREIN, Jr., W.J., DUNSMUIR, P. and RUBIN, G.M. (1979) Cell 17 415-427.

PRENSKY, W., STEFFENSEN, D.M. and HUGHES, W.L. (1973) Proc. Nat. Acad. Sci. USA 70 1860-1865.

PROCUNIER, J.D. and TARTOF, K.D. (1978) Genetics 88 67-79.

RASCH, E. (1970) In Introduction to Quantitative Cytochemistry 2, G.L. WIED and G.F. BAHR Eds. (New York: Academic Press) 357-397.

REEDER, R.H. (1974) In Ribosomes, M. NOMURA, A. TISSIERES, and P. LENGYEL Eds. (New York: Cold Spring Harbour Laboratory) 489-518.

REEDER, R.H., HIGUSHINAKAGAWA, T. and MILLER, Jr., O. (1976) Cell 8 444-454.

REEDER, R.H., SOLLNER-WEBB, B. and WAHN, H.L. (1977) Proc. Nat. Acad. Sci. USA 74 5402-5406.

RENKAWITZ, R., GERBI, S.A. and GLATZER, K.H. (1979) Mol. Gen. Genet. 173 1-13.

RIGBY, P.W.J., DIECKMAN, M., RHODES ,C. and BERG, P. (1977) J. Mol. Biol. 113 237-251.

RITOSSA, F. (1976) In The Genetics and Biology of Drosophila 1b , M. ASHBURNER and E. NOVITSKI Eds. (London, New York, San Francisc0: Academic Press) 801-846.

RITOSSA, F.and M. and SPIELGELMAN, S. (1965) Proc. Nat. Acad. Sci. USA 53 737-745.

ROBERTS, R.J. (1978) Gene 4 183-193.

ROCHAIX, J.D. and MALNOE, P. (1978) Cell 15 661-670.

ROHLENDORF, B. (1974) The Historical Development of Diptera (Edmonton: The University of Alberta Press).

ROOP, D.R., NORDSTROM, J.L., TSAIS, Y., TSAI, M.J. and O'MALLEY, B.W. (1978) Cell 15 671-685.

ROSS, J. (1976) J. Mol. Biol. 106 403-420.

RUBIN, G.M. and SULSTON, J.E. (1973) J. Mol. Biol. 79 521-530.

SANDLER, L. (1970) Genetics 64 481-493.

SHAPIRO, J.A., ADHYA, S.L. and BUKHARI, A.I. (1977) In DNA Insertion Elements Plasmids and Episomes, A.I. BUKHARI, J.A. SHAPIRO and S.L. ADHYA Eds. (New York: Cold Spring Harbour) 3-13.

SMITH, G.P. (1973) Cold Spring Harbour Smp. Quant. Biol. 38 507-1513.

SOLLNER-WEBB, B. and REEDER, R.H. (1979) Cell 18 485-499.

SOUTHERN, E.M. (1975) J. Mol. Biol. 98 503-517.

SPEAR, B.R. (1975) Chromosoma 48 159-179.

SPRADLING, A., PARDUE, M-L. and PENMAN, S. (1976) J. Mol. Biol. 109 559-587.

SPRADLING, A., PENMAN, S. and PARDUE, M-L. (1975) Cell 4 395-404.

STROBEL, E., DUNSMUIR, P. and RUBIN, G.M. (1979) Cell 17 429-439.

TARTOF, K.D. (1973) Genetics 73 57-71.

TARTOF, K.D. (1979) Cell 17 607-614.

TARTOF, K.D. and DAWID, I.B. (1976) Nature 263 27-30.

TILGHMAN, S.M., CURTIS, P.J., TIEMEIGR, D.C., LEDER, P. and WEISSMANN, C. (1978) Proc. Nat. Acad. Sci. USA 75 1309-1313.

TSUJIMOTO, Y. and SUZUKI, Y. (1979) Cell 18 591-600.

UDEM, S.A. and WARNER, J.R. (1972) J. Mol. Biol. 65 227-242.

VALENZUELA, P., VENEGAS, A., WEINBERG, F., BISHOP, R. and RUTTER, W.J. (1978) Proc. Nat. Acad. Sci. USA 75 190-194.

VANOMMEN, G-R.B., GROAT, G.S.P. and GRIVEll, L.A. (1979) Cell 18 511-524.

VOGT, V.M. and BRAUN, R. (1976) J. Mol. Biol. 106 567-587.

WALLACE, H. and BIRNSTIEL, M.L. (1966) Biochem. Biophys. Acta. 114 296-310.

WEISS, B., JACQUEMIN-SABLON, A., LIVE, T.R., FAREED, G.L. and RICHARDSON, L.L. (1968) J. Biol. Chem. 243 4543-4555.

WELLAUER, P.K. and DAWID, I.B. (1974). J. Mol. Biol. 89 379-395.

WELLAUER, P.K. and DAWID, I.B. (1977) Cell 10 193-212.

WELLAUER, P.K. and DAWID, I.B. (1978) J. Mol. Biol. 126 769-782.

WELLAUER, P.K. and DAWID, I.B. (1979) J. Mol. Biol. 128 289-303.

WELLAUER, P.K., DAWID, I.B., BROWN, D.D. and REEDER, R.H. (1976a) J. Mol. Biol. 105 461-486.

WELLAUER, P.K., DAWID, I.B. and TARTOF, K.D. (1978) Cell 14 269-278.

WELLAUER, P.K., REEDER, R.H., DAWID, I.B. and BROWN, D.D. (1976b) J. Mol. Biol. 105 487-505.

WENSINK, P.C. FINNEGAN, D.J. , DONELSON, J.E. and HOGNESS, D.S. (1974) Cell 3 315-325.

WHITE, R.L. and HOGNESS , D.S. (1977) Cell 10 177-192.

WILD, M.A. and GALL, J.G. (1979) Cell 16 565-573.

WILLIAMSON, J.H. (1976) In The Genetic and Biology of Drosophila 1b M. ASHBURNER and E. NOVITSKI Eds. (London, New York, San Francisco: Academic Press) 667-699.

YAO, M.L. and GALL, J.G. (1977) Cell 12 121-132.

YOUNG, M.W. (1980) Proc. Nat. Acad. Sci. USA In Press.