STUDIES ON AEDES AND ANOPHELES MOSQUITOES AT THE MOLECULAR LEVEL OF GENETICS

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Kevin Richard Gale

November 1986

Acknowledgements

I would like to thank my Ph.D. supervisor Julian M. Crampton for his invaluable help and advice throughout the three years of my studies. I also thank Teresa F. Knapp for her excellent technical advice.

I thank Rory J. Post, Harold Townson, John T. Ellis, Ann M. Warren, Patricia Burke and John D. Norton for advice and helpful discussions during my work in this Department.

I thank Mike W. Service of this Department for advice and the supply of all wild-caught *An. gambiae* material. I also thank C. Malcolm, R.G. Bonnington and B. Sawyer of the London School of Tropical Medicine and Hygiene for the supply of all *Anopheles gambiae* laboratory stock material. For the supply of all other material, I thank Sharon McKenna, Gill Nicholson and Carol Allen of this Department.

The work described in this thesis was funded by the Wolfson Foundation.

ABSTRACT

Section 1: Construction and screening of a genomic library for the mosquito *Aedes aegypti* (Diptera, culicidae).

A genomic library has been constructed for this important vector of arboviral disease. Total genomic DNA and various classes of RNA from *Ae. aegypti* were used to screen this library. The results obtained indicate that this species has a short period interspersion pattern of repeated sequences. Transcription of these repeats could not be detected using total cytoplasmic RNA, hnRNA or mRNA as hybridisation probes.

Section 2: Sequence organisation of ribosomal DNA in *Aedes aegypti*.

The *Aedes aegypti* genomic library was used to isolate clones containing the intact ribosomal DNA (rDNA) repeat of this species. This has been restriction mapped and the transcribed regions have been identified. The rDNA repeat is 9.0 Kb in length and is present as approximately 500 head-to-tail tandemly repeated copies. A low level of intraspecies polymorphism of *Ae. aegypti* rDNA is evident. Two restriction polymorphisms have been identified within the rDNA repeat.

Section 3: Analysis of ribosomal DNA variation within *Ae. aegypti* and between closely related species.

Four variant rDNA clones have been isolated. One of these may contain the end of a tandem array of ribosomal genes. Another variant contains a duplication of rDNA within the internal transcribed spacer region of the ribosomal repeat. Sequence analysis of this clone has identified regions at the 3' end of the 18S rRNA gene of *Ae. aegypti* which show very strong homology with the corresponding regions in other species. Some repeated sequences have been identified downstream of the 18S rRNA gene in this clone. Preliminary analysis of the two other rDNA variants indicates that one contains a duplication or insertion of DNA in the 28S rRNA coding region and one contains non-transcribed spacer homologous sequences which are not associated with rRNA coding regions.

Section 4: DNA probes for species identification of mosquitoes in the *Anopheles gambiae* complex.

DNA sequences have been isolated which distinguish four of the morphologically identical members of the *An. gambiae* species complex. Two sequence classes were obtained. Class 1 homologues are highly reiterated in the genomes of *An. arabiensis* and *An. merus*, present in low copy number in *An. melas* and were not detected in *An. gambiae* s.s. Class 1 sequences are male specific in *An. arabiensis*. Class 2 homologues are highly reiterated in the genomes of *An. merus* and *An. melas* and present in low to middle copy number in *An. gambiae* s.s. and *An. arabiensis*. Sex specificity of Class 2 homologues does not occur in the species tested (*An. gambiae* s.s. and *An. arabiensis*). Hybridisation of these species specific DNA sequences to mosquitoes squashed directly onto nitrocellulose provides a simplified method of species identification.

List of Abbreviations

cDNA	complementary deoxyribonucleic acid
rDNA	ribosomal DNA
rRNA	ribosomal ribonucleic acid
ITS	internal transcribed spacer
ETS	external transcribed spacer
NTS	non-transcribed spacer
O.D.	optical density
λ wt	lambda bacteriophage wild-type
TGE	transposable genetic element
M.wt.	molecular weight
SDS	sodium doedecyl sulphate
SSC	saline sodium citrate (1 x = 0.15 M NaCl, 0.015 M NaCitrate)
TE	Tris-EDTA buffer (10 mM Tris, 1mM EDTA)
S	Svedberg unit
60'	(60) minutes
rpm	rotations per minute
bp	base pairs of DNA
Kb	kilobase pairs of DNA
M	molar
mM	millimolar
μM	micromolar
nM	nanomolar
pM	picomolar
ml	millilitre
μl	microlitre

Table of Contents

Contents continued

Contents continued

GENERAL INTRODUCTION

In 1898, Ronald Ross described for the first time the mosquito cycle of malaria transmission. In 1901 mosquitoes were implicated in the transmission of yellow fever by Walter Reed and in 1906 T.L. Bancroft discovered the mosquito transmission of dengue. These early discoveries prompted the beginning of exhaustive research into the biology of mosquitoes and their relation to diseases.

Mosquitoes are still the subject of much research. In many tropical areas mosquito numbers and disease transmission remain unchecked. This, in some areas despite the considerable efforts of huge public health initiatives. Major problems in the control of mosquito-borne diseases stem from the vastness of the areas of the world affected and a lack of sufficient resources with which to act. In some countries the political situation and an uninformed public add further to health problems.

In recent years mosquito-borne disease has been brought under control in some areas by the combined use of insecticides to control mosquito numbers, drugs to combat parasite infections and preventative measures such as the use of mosquito nets and the destruction of mosquito breeding sites in the vicinity of populated areas. The effectiveness of two of these measures is now seriously threatened by the increasing prevalence of insecticide resistance in mosquitoes and the appearance of drug resistance in some parasites. Although a vaccine is now available which can be used to protect against yellow fever, attempts to produce a vaccine against malaria have so far been

thwarted by the immunologically evasive nature of the parasite. No effective cure or vaccine is yet available to combat the relatively new disease dengue. Even when effective drugs and vaccines are available, the expense and distribution problems involved in treating large numbers of people for long periods, makes the task of disease erradication extremely difficult. For these reasons it is vital that efforts to control mosquito-borne disease continue at all levels: vaccines, drugs and vector control, and that the strength of all available technology is brought to bear.

Although extensive information is available concerning mosquito biology, genetics and (for some species) cyto-genetics, very little is known about mosquitoes at the level of molecular genetics. This is somewhat surprising in view of the medical importance of these insects. An understanding at the molecular level of phenomena such as insecticide resistance and susceptibility of mosquitoes to infection by the parasites they transmit may eventually lead to effective new methods of mosquito control and interruption of disease transmission. Genetic control of insect pests has been explored [1] and further information regarding molecular genetics, rather than just classical genetics may eventually make the genetic control of mosquitoes possible. One of the most powerful techniques for the study of gene function in higher eukaryotes is that of 'P' element transformation of *Drosophila melanogaster* [2]. The development of a system for the genetic transformation of mosquitoes in a similar manner (either using 'P' element DNA vectors or by developing mosquito DNA vectors) would prove invaluable for the analysis of gene function in mosquitoes, the isolation of otherwise unselectable genes by 'transposon-tagging' and may eventually provide a means of genetically altering the mosquito.

It is impossible to predict the precise benefits which might arise from a study of the molecular genetics of the mosquito, but such a study is sure to add further to the armoury of weapons against the mosquito and will also be of interest to researchers in the broader field of eukaryote molecular genetics.

The experiments described in this thesis represent some of the first work to be done specifically on mosquito molecular genetics. Three sections of experiments are described. The first section details the construction of a basic tool for the study of the *Aedes aegypti* mosquito genome; a large genomic library. Some information regarding the organisation of repeated sequences in the genome of this important vector of arboviral disease is presented. In the second section, this genomic library is used to isolate a number of clones representing a particular family of repeated sequences; the ribosomal RNA genes. The variation of ribosomal genes within *Aedes aegypti* and between closely related *Aedes* species is investigated and the genomic library is used to isolate specific variants. The viability of screening ribosomal genes to isolate putative transposable genetic elements is explored. The implications of the results of these experiments for further work on mosquito molecular genetics is considered.

The third section of this thesis describes a short project in which some of the most basic techniques of molecular genetics are applied to a fundamental problem in entomology; the identification of individual species within a species complex. The isolation of DNA probes for species identification of mosquitoes in the *Anopheles gambiae* complex is described, together with a preliminary study of the feasibility of using DNA probes for routine species identification.

1.1. *Aedes aegypti*; Classification, Basic Biology and Vectorial Importance

Classification

Mosquitoes comprise the Dipteran family Culicidae. *Aedes aegypti* belongs to the genus Aedes, sub-genus Stegomyia. Mosquito systematics are reviewed by Knight and Stone [3] and Knight [4].

The name now used *Aedes aegypti* was originally proposed by Linnaeus in 1762, although this was only adopted in the 1930's. *Ae. aegypti* has formerly been referred to as *Culex fasciatus*, *Stegomia fasciatus*, *Stegomyia calupus* and *Culex argenteus*. More recently, *Ae. aegypti* has been classified into three forms [5].

1. *Ae. aegypti* (Linnaeus) sensu-stricto (the type form) is brownish or blackish in colour, distributed widely, but absent from inland tropical Africa.

2. *Ae. aegypti* (Walker) s.sp.formosus is black, confined to tropical Africa and is the only form found inland in tropical Africa.

3. *Ae. aegypti* (Theobald) var. queenslandensis is paler than *Ae. aegypti* s.str. Originally identified in Northern Australia, its distribution is similar to that of *Ae. aegypti* s.str.

Basic Biology (see [6] for full review)

Aedes aegypti is the most studied single mosquito species and is commonly used in the testing of insecticides and repellants. Distributed almost worldwide within tropical and sub-tropical zones, *Ae. aegypti* is an extremely hardy species which is easy to rear in laboratory colonies. Eggs may be stored dry for several months after initial conditioning and remain viable after storage for a year or more.

Aedes aegypti will breed in almost any unpolluted temporary water. Typical breeding sites include tree holes, leaf axils, temporary pools, domestic recepticles and even uncovered toilet cisterns. *Ae. aegypti* is therefore particularly well adapted to breeding in and around human habitats. Up to 140 eggs are deposited singly onto wet surfaces or directly onto sheltered water. A minimum of three days is required for maturation of eggs. After eclosion, the first instar larva goes through three ecdyses to fourth instar larva before pupation and subsequent emergence of the imago. The time taken for development from egg to adult is largely dependant on ambient temperature, ranging from a minimum of 7 days in the warm season (28-32°C), to 20 days in the cool season (23-26°C). The size of adult produced is not adversely affected by a moderately reduced temperature [7] although development will not proceed below approximately 17°C. The food requirements of the filter-feeding larvae are not stringent, almost any source of living bacteria and fungi being suitable. After emergence, 1-2 days is required for the female to become receptive to mating. Only the first mating is usually effective due to the polypeptide 'matrone' which is introduced into the female reproductive tract by the accessory gland fluid of the first mating partner [8].

Three days are required for oogonal development between the females first blood meal and oviposition. The total life cycle of *Aedes aegypti* has a duration of between 16 and 30 days.

Adult nourishment of both sexes for flight and metabolism is usually derived from plant juices (moist sugar cubes serve for laboratory colonies), although a blood meal is essential for ovarian development and egg production by the female. Under favourable conditions in laboratory colonies (70% relative humidity, 25°C), 50% of females may survive up to 60 days, the oldest survivors living for 100 days. Male longevity is about half that of females [9]. Survival of females in natural populations is much lower (e.g. 2-3 ovipositions or 10 days on average) [10]. Longevity is of great importance in determining the vectorial capacity of a species [11]. *Aedes aegypti* is active mainly by day. Consequently this species has a high degree of contact with man and is a particularly annoying pest.

Vectorial Capacity

Aedes aegypti is the main vector responsible for the transmission of yellow fever (Y.F.) and dengue fever (D.F.) in urban areas. It also plays a role in the transmission of many less important arboviruses of man and other animals and in the transmission of *Dirofilaria*. Factors contributing to the high vectorial capacity of *Aedes aegypti* include: high anthrophilly, hardiness, mobility, the ability to support a wide range of pathogens, high longevity, high biting frequency and a high population density in many areas.

Yellow fever (causative agent: flavivirus) occurs as two forms: urban Y.F. (affecting humans) and sylvan or forest Y.F. (affecting

humans and monkeys). *Aedes aegypti* is the prime vector of urban Y.F., but plays only a minor role in the transmission of sylvan Y.F. (of which *Aedes simpsoni* and *Aedes africanus are the main vectors). The distribution of Y.F. is similar to that of Aedes aegypti*, with the exception that Y.F. is not endemic in Asia, even though *Ae. aegypti* is still prevalent. Urban Y.F., which formerly caused large epidemics in urban centres of Africa, America and Europe, is now largely controlled by the use of efficient environmental measures and vaccines. The possibility remains however that re-invasion of urban areas by vectors could occur and if vaccination is halted, this could lead to fresh epidemics with catastrophic results. One particular cause for concern is that Y.F. may spread to India where no preventative measures exhist [12].

Dengue fever (causitative agent: flavivirus, four serotypes DEN 1-DEN 4) is transmitted mainly by *Ae. aegypti*, but also by other steganomous species including *Aedes albopictus, Aedes scutellaris* and *Aedes polynesiensis*. The distribution of D.F. is similar to that of *Ae. aegypti*, with the most severely affected areas at present being S.E. Asia and the Caribbean. D.F. occurs in two forms; one with haemorrhagic and shock manifestations (dengue haemorrhagic fever, or D.H.F.) and one without these more severe manifestations (called simply dengue fever, D.F). Outbreaks of D.F. may involve 0.5-2 million people annually. Although debilitating, this form is rarely fatal. Large areas in Asia experience frequent outbreaks of D.H.F. which result in significant mortality (e.g. 7%), especially amongst children. D.H.F. is thought to represent a hypersensitivity reaction to simultaneous or sequential infection by more than one dengue virus serotype.

Aedes aegypti also contributes to the transmission of several other less widespread arboviral diseases afflicting humans and other animals. These include Japanese B, St. Louis and Murray Valley encephalitis, West Nile virus, Bwamba fever, Chikungunya, Semliki Forest virus, Bunyamwera, Ntaya, Uganda S virus, Zika virus, fowl pox, Venezuela equine encephalitis and enzootic hepatitis. *Aedes aegypti* also transmits *Dirofilaria immitis* (dog filariasis).

1.2. Construction of an *Aedes aegypti* Genomic Library

A representative genomic library contains all the DNA sequences of an organism present as a number of DNA fragments, each inserted into a bacterium or bacteriophage. Information concerning the arrangement and expression of repeated sequences may be gained by screening genomic clones with various DNA and RNA hybridisation probes [13]. Specific genes or other DNA sequences may be isolated from a genomic library when homologous probes (e.g. an evolutionarily conserved gene from a different species) are available. Alternatively, differential screening of a genomic library may be used to isolate DNA sequenes with particular properties; differing expression in various tissues [14] or a particular species distribution [15] for example. If the genomic library is constructed using randomly sheared or partially digested insert DNA, a series of overlapping fragments is obtained. This allows the isolation of sequences adjacent to a region already cloned. A genomic library may therefore be considered a basic tool of molecular genetics and as such, it was considered worthwhile to construct a large genomic library for *Ae. aegypti*. This library has been used for many of the experiments described in this thesis and should remain an aid to future work.

1. The bacteriophage vector λgtWESλB

λgtWESλB is a replacement vector derived by Leder *et al* [16]. It is well suited for cloning intermediate sized EcoRI fragments (7-

15Kb). Digestion of λgtWESλB (40.4 Kb total) with EcoRI generates three restriction fragments; a 21.7Kb 'left arm', a 4.9 Kb 'stuffer' fragment (which contains no essential sequences) and a 13.8 Kb 'right arm'. After separation of the stuffer fragment, the annealed arms and EcoRI-generated insert DNA are ligated under conditions favouring the formation of concatamers. This ligation mixture is an efficient substrate for *in vitro* packaging. Insert fragments in the range 2-15 Kb can be accommodated along with the arms in the phage particle. Plating of the recombinant phage obtained by *in vitro* packaging onto *E. coli* host cells (e.g. LE 392) yields an amplified stock which may be stored for several years at 4°C without appreciable decrease in titre. Bacteriophage vectors allow sizeable portions of complex genomes to be screened easily. The size of inserts in λgtWESλB genomic libraries makes cloning intact eukaryotic genes possible, facilitates chromosome-walking in reasonable-sized steps and is conducive to restriction enzyme mapping (unlike cosmid-sized inserts).

2. Preparation of vector DNA

λgtWESλB bacteriophage DNA was prepared as described in Materials and Methods. A small aliquot of this DNA was digested with EcoRI to verify the strain. 100 µg of vector DNA was then digested overnight with a 2-fold excess of EcoRI. A small aliquot of DNA was checked for 100% digestion. The remainder was then phenol extracted, ethanol precipitated and redissolved in T.E. buffer (10 mM Tris pH 7.5, 1 mM EDTA).

The cohesive ends of the λ arms were annealed (60' in 10 mM $MgCl_2$ at 42°C). Annealing was checked by gel electrophoresis. Two 10-40%,

14 ml sucrose gradients (in 1M NaCl, 20 mM Tris-HCl pH 8, 5 mM EDTA) were prepared using a gradient former (BRL GA 0793). Half of the digested vector was carefully layered onto each gradient and centrifuged for 16 hours at 32,000 rpm (Sorvall TST41 swingout rotor, 'g' at 'r.av' = 1.3 x 10^5). The gradients were fractionated from the base of the tube (0.7 ml fractions) using a peristaltic pump. 20 µl of every other fraction was analysed by gel electrophoresis. Fractions 1-5 were found to contain λ arms and in fraction 7 the stuffer fragment was visible. Fractions 1-4 were pooled, dialysed against T.E. and the DNA ethanol precipitated. The DNA was redissolved in 20 µl of T.E. and a small aliquot checked by gel electrophoresis. The final yield of purified arms was estimated to be 10 µg.

3. Preparation of genomic insert DNA

The insert DNA used was 7-15 Kb size fractionated *Aedes aegypti* (Bangkok strain) genomic DNA, obtained from a mixture of complete and partial EcoRI digests. (This strain of *Ae. aegypti*, originally from Bangkok, Thailand, is resistant to the insecticide DDT). The advantages of using size fractionated, partially digested DNA are:

1. Precise excision of insert DNA is possible.
2. Elimination of small fragments (e.g. 2-7 Kb) from the genomic library reduces the number of clones which have to be screened to screen a given proportion of the genome.
3. Partial digestion favours cloning of regions of the genome sparse in EcoRI sites and of regions rich in EcoRI sites.

4. Ligation of unrelated genomic DNA fragments during the ligation reaction will mainly generate inserts too large to be packaged. Multiple inserts will therefore be largely avoided. (These give misleading results in the analysis of flanking regions and when 'chromosome walking').

The main disadvantages of this cloning system are that EcoRI sites within DNA are not evenly distributed and that EcoRI cuts infrequently relative to the size of DNA fragments cloned. Long stretches of highly repetitive DNA may contain no EcoRI sites and thus would not be cloned [17]. The proportion of *Aedes aegypti* DNA which is present as fragments larger than 15 Kb after complete digestion with EcoRI is approximately 2-5% (as estimated from an ethidium bromide stained gel of EcoRI restricted genomic DNA). Use of tetracutter restriction endonucleases or mechanical shearing of genomic DNA is required to generate a truly representative genomic library.

4. Preparation of high molecular weight genomic DNA

The method of Bingham *et al* [18], with some modifications, was used to prepare pure, high molecular weight genomic DNA. (A detailed protocol is given in Materials and Methods). DNA was prepared from 2-400 fourth instar *Ae. aegypti* (Bangkok strain) larvae. The size of DNA obtained was greater than 100 Kb as estimated by gel electrophoresis. Elimination of mechanical shearing of DNA is important in the cloning system employed here as DNA fragments with non-EcoRI compatible termini are unable to ligate properly and so act as inhibitors of concatamerisation and packaging.

5. Titration of EcoRI against *Aedes aegypti* DNA

To determine the quantity of restriction endonuclease required to produce a given level of digestion of genomic DNA, the batch of enzyme to be used was titrated against the genomic DNA to be used as follows.

20 µg of *Ae. aegypti* DNA in 300 µl of EcoRI reaction buffer was aliquoted into 9 tubes on ice: 60 µl in tube 1, 30 µl in 2-9. 10 units of EcoRI were added to tube 1 and after mixing, 30 µl from tube 1 was transferred to tube 2. After mixing, 30 µl from '2' was transferred to '3' etc. No enzyme was added to tube 9. All tubes were then incubated at 37°C for 1 hour. The reaction was stopped by the addition of EDTA to 10 mM. The digests were then analysed by gel electrophoresis. The result of this experiment is shown in Figure 1.2a. The unit activity of the enzyme was estimated as 1 µg of genomic DNA per hour under the conditions of the assay.

Fig. 1.2 A

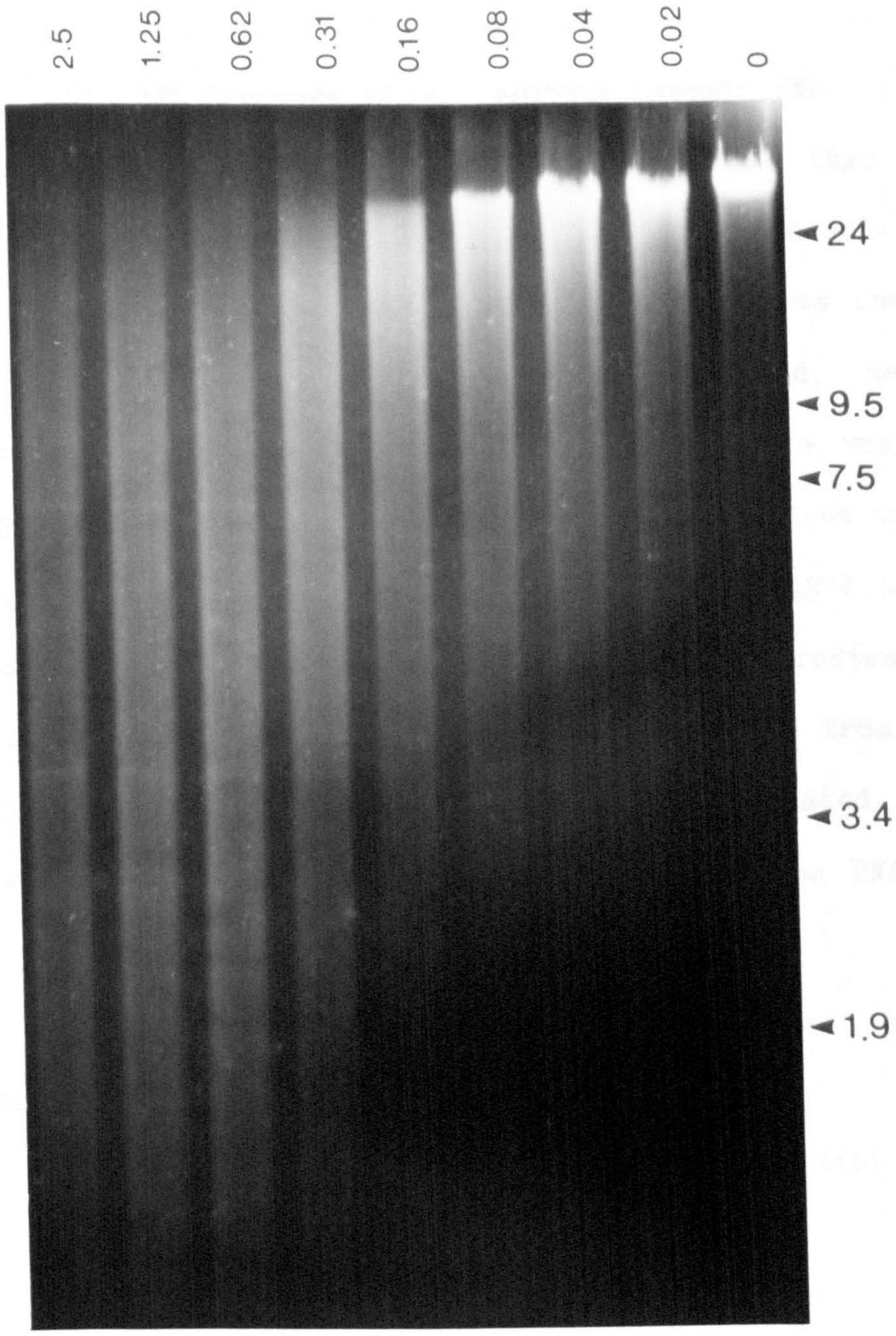


Figure 1.2A; Titration of the restriction endonuclease EcoR1 against *Ae.aegypti* genomic DNA.

6. Digestion and size fractionation of genomic DNA for insertion into vector

Four EcoRI digests were performed under conditions which gave 10%, 25%, 50% and 100% cleavage of *Ae. aegypti* genomic DNA. 25 µg of DNA was used per reaction at a concentration equal to that used in the titration experiment. Reactions were stopped by the addition of EDTA to 10 mM and 0.5 µg of DNA from each reaction was checked by gel electrophoresis. The digests were then pooled, gently phenol extracted and ethanol precipitated. The precipitate was redissolved overnight at 4°C, in 200 µl of T.E. and a small aliquot checked by gel electrophoresis. The DNA was then heated (68°C, 10') prior to size fractionation through a single 10-40% sucrose gradient which was fractionated as described in '2' above. 100 µl from every third fraction was ethanol precipitated and size fractionated through 0.8% Agarose to determine the size distribution of the DNA (see Figure 1.2b).

Results;

Fraction no.	Size distribution (Kb)
2	8.4 - 11
5	4.8 - 7.5
8	2.1 - 3.2
11	less than 2.0

No DNA was visible in fractions 14, 17 and 20. Fractions 1-3 were pooled, dialysed against T.E. and ethanol precipitated. The DNA was redissolved overnight in 20 µl of T.E. and a small aliquot (2µl) checked by gel electrophoresis (see Figure 1.2c). The final yield of

Fig. 1.2

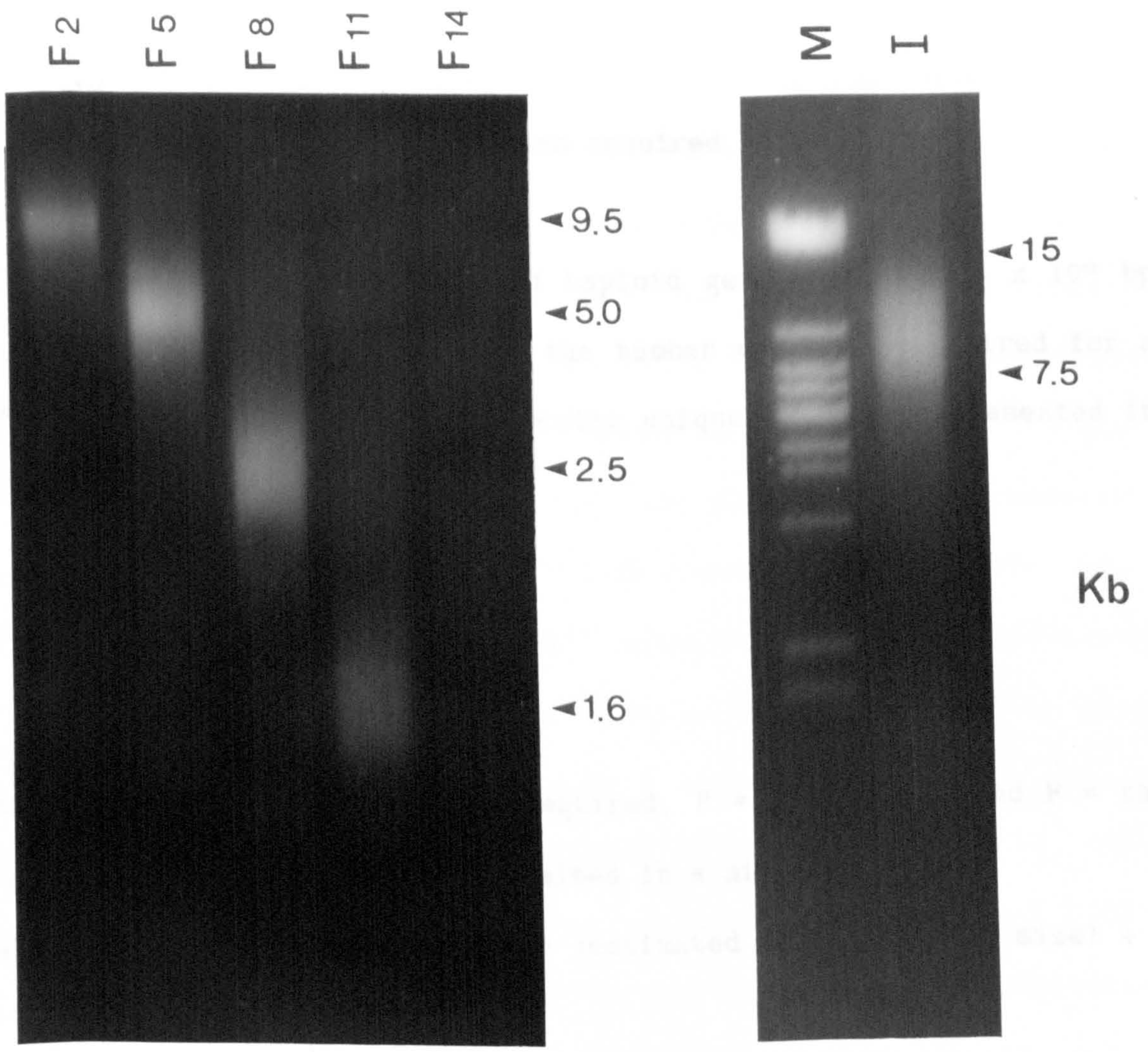


Figure 1.2B and 1.2C; Preparation of size fractionated, EcoR1 partial digested *Ae.aegypti* genomic DNA. A; size fractionated DNA, B; size distribution of pooled fractions. M= λ marker, I= insert DNA.

7-15Kb, size fractionated *Ae. aegypti* genomic DNA was estimated to be 10 µg.

7. Estimate of the number of clones required

Using a value for *Ae. aegypti* haploid genome size of 8×10^8 bp (see section II of this thesis), the number of clones required for a 99% probability of having a particular unique sequence represented in the library was calculated [19].

Using $$N = \frac{\ln (1 - P)}{\ln (1 - F)}$$

Where N = the number of clones required, P = probability and F = the fraction of the *Aedes* genome contained in a single clone.

Values used; P = 0.99, F = 10^4 bp (estimated average insert size) ÷ 8×10^8 (genome size).

$$N = \frac{\ln (1 - 0.99)}{\ln (1 - [10^4 \div 8 \times 10^8])} = 3.68 \times 10^5 \text{ clones}$$

This value is calculated assuming random fragmentation of the genome. This is not true for the method used here (EcoRI digestion), but the value of 3.68×10^5 clones serves as a minimum number of clones for which to aim.

8. *In vitro* packaging

Packaging extracts were prepared and used as described by Hohn [20]. The protocol used is detailed in Materials and Methods.

The lysogenic *E. coli* strains BHB2688 (Eam) and BHB2690 (Dam) were used to produce the freeze-thaw lysate and the sonicated extract respectively. BHB2688 has an amber mutation in the 'E' gene (coding for phage head protein) and so accumulates viral capsid components following induction. BHB2690 has an amber mutation of the 'D' gene (required for insertion of λ DNA into the phage head and subsequent maturation of the phage particle) and so accumulates 'pre-heads' when induced. Both strains have the cIts857 mutation (temperature sensitive λ repressor) and Sam 7 (an amber mutation in the 'S' (lysis) gene which causes the accumulation of capsid components in SuIII$^-$ cells after induction). Both have 'b' region deletions which remove the DNA attachment site, thereby reducing the packaging of endogenous λ DNA molecules. In addition, both BHB2688 and BHB2690 have inactivated generalised recombination systems (red 3 and recA respectively) which minimises recombination between endogenous DNA and the added recombinant λ DNA. When the two extracts are mixed in the presence of cofactors and λ wild-type DNA, 10^7 - 10^8 plaque forming units (pfu's) are produced per microgram of DNA.

The packaging buffers used here contain putrescine (see Materials and Methods), which favours the packaging of λ genomes close to the size of λ wild type DNA. The packaging of recombinants containing large insert DNA fragments is therefore favoured.

9. Optimisation of sonicated extract: freeze-thaw lysate ratio

In two separate experiments, the packaging efficiency of the packaging mixes was determined 'a', using a set volume of freeze-thaw lysate (FTL) and varying the volume of sonicated extract (SE) and 'b', using a set volume of SE and varying the FTL.

Results;

Reaction				Efficiency (pfu/μg λ wt)
a.	10 μl FTL	+	4 μl SE	6×10^6
	" "	+	6 μl "	8.3×10^6
	" "	+	10 μl "	4.2×10^6
	" "	+	20 μl "	2.0×10^6
b.	6 μl SE	+	10 μl FTL	2.3×10^7
	" "	+	12 μl "	2.3×10^7
	" "	+	16 μl "	1.8×10^7
	" "	+	20 μl "	1.9×10^7

No phage plaques were obtained when a 10^{-3} dilution of a packaging reaction (10 μl FTL + 6 μl SE) without added λ wt DNA was plated. No increase in packaging efficiency was obtained when the concentration of ATP in the reaction was increased or when the reaction was incubated for 1½ hours instead of 1 hour. The lower efficiences observed for 'a' are due to deterioration of packaging buffers used in this experiment.

The volumes found to give best results were; 10 μl of FTL + 6 μl of SE. The packaging efficiency obtained was 2.3×10^7 pfu/μg λ wt DNA.

10. Determination of optimum ligation conditions

The most efficient substrate for *in vitro* packaging is a concatamer of the form [insert-annealed λ arms]$_n$. The formation of this is optimal when the EcoRI cohesive ends of each of the three reacting species (left arm, insert and right arm) are present in equimolar amounts. Therefore, a ratio of two of each of the arms to one insert molecule is required (e.g. a 2:1 molar ratio of annealed arms to inserts). The absolute concentration of reacting species is also important. The ligation conditions must favour intermolecular concatamerisation rather than intramolecular self-circularisation. The theoretical optimum conditions for ligation were calculated as described by Maniatis *et al* [21] using a value of 10 Kb for average insert DNA size (workings not shown).

The values obtained were; molarity of insert DNA = 8.52×10^{-9} (57 µg/ml), molarity of annealed arms = 2 x molarity of inserts = 1.74×10^{-8} (400 µg/ml).

Although these values serve as a good theoretical approximation, several inaccuracies are inherrent in practice. DNA concentrations as estimated by ethidium bromide fluorescence are inaccurate. Some DNA molecules may be damaged and therefore unable to react. The calculations used do not take into account the changing concentrations of reactants with time.

The actual optimum ratio of arms : insert was determined empirically using the theoretical optimum as a guide. The molar ratio of arms : insert DNA was varied and the efficiency of each reaction product as a substrate for *in vitro* packaging determined.

Results;

Ratio of arms : insert	Efficiency (pfu/μg)
4:1	2.5 x 10^4
2:1	7.5 x 10^4
1:1	4.5 x 10^4
0.5:1	6.0 x 10^4

In this case, the theoretical and actual optima were the same, a 2:1 molar ratio of λ arms : insert DNA.

11. Large-scale ligation and packaging

Reaction: 0.34 μg of *Ae. aegypti* 7-15Kb, EcoRI restriction fragment insert DNA and 2.4 μg of annealed λ arms were ligated in a reaction volume of 11.2 μl using 5 units of T4 DNA ligase. The reaction was incubated overnight at 20°C.

The final concentrations of λ arms and insert DNA were 214 μg/ml and 30 μg/ml (a 2 : 1 molar ratio). The theoretical optimal concentrations could not be attained in the minimum possible reaction volume.

The ligation reaction was packaged using 14 separate packaging reactions (190 ng DNA/reaction). Controls were performed to determine packaging efficiency for λ wt DNA and the background of plaques from a packaging reaction with no added insert DNA.

Results

Reaction	Efficiency (pfu/µg DNA)
a, 250 ng λ wt/reaction	3.60×10^7
b, 500 ng λ wt/reaction	3.26×10^7
c, 100 ng ligated arms only/reaction	1×10^4

The total number of pfu's in the combined packaging of the large-scale ligation was 1.36×10^6 (This corresponds to an efficiency of 5×10^5 pfu/µg total DNA).

From 'c', the background of non-recombinants in the genomic library was estimated to be approximately 2%. The number of recombinants obtained was therefore approximately 1.30×10^6 (3.6 x the value of the minimum number as calculated in '7' above)

12. Amplification

Half of the recombinants obtained were kept at 4°C (with added chloroform) as an unamplified stock. The other half were plated onto two BRL bioassay trays (22.5 cm x 22.5 cm). The top agar from these plates was then scraped off into phage buffer and kept overnight at 4°C to allow the phage to diffuse out of the agar. The agar was then spun down (10,000rpm, 20', 4°C) and the supernatent (the amplified phage stock) decanted. The titre of the phage stock (total volume 30 ml) was determined to be 2.2×10^{10} pfu/ml.

Although low abundance, slow growing recombinants may be lost during amplification of the library, these should still be represented in the unamplified stock.

13. Evaluation of the representation of ribosomal DNA sequences in the *Aedes aegypti* genomic library

To assess the representation of a particular sequence family (in this case the major ribosomal RNA genes) in the amplified library stock, the ribosomal DNA (rDNA) composition of the library was compared with that of the original genomic DNA as described by Phillips *et al* [22].

Amplified library DNA was prepared as follows: The phage particles from 5 ml of amplified library stock (2.2 x 10^{10} pfu/ml titre) were precipitated by the addition of 1.25 ml of a solution containing 20% polyethylene glycol 6000, 2.5 M NaCl. After standing at 4°C for 20 minutes, the precipitate was collected by centrifugation (10,000 r.p.m. 4°C, 15') and resuspended in 500 µl of T.E. buffer. DNA was prepared from this by extracting the suspension with phenol three times and choroform once, followed by ethanol precipitation and washing with 70% ethanol. A yield of 12.5 µg of recombinant DNA was obtained.

2 µg of *Aedes aegypti* (Bangkok) genomic DNA and 5 µg of genomic library DNA were digested with the restriction enzyme EcoRI, size fractionated through 0.8% Agarose and transferred to nitrocellulose as described in materials and methods. This Southern blot was then hybridised with radiolabelled *Ae. aegypti* rDNA (λ Aar7 total insert DNA, see Section 2 of this thesis). A photograph of the stained gel and corresponding autoradiograph are shown in Figure 1.2d.

This clearly shows that all rDNA bands visible for the digested genomic DNA are also visible on the genomic library DNA digest. Some difficulty was encountered in obtaining complete digestion of the library DNA and this is evident from the strong partial digestion band

Fig.1.2 D

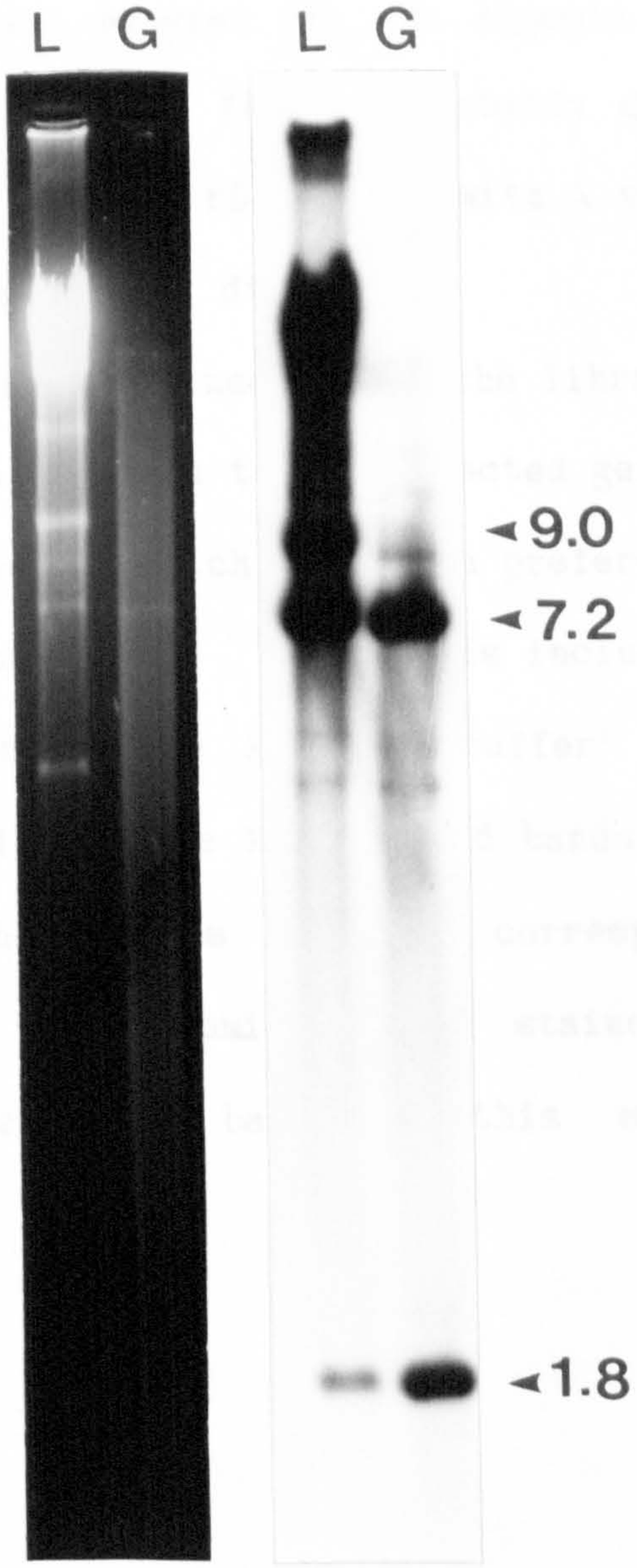


Figure 1.2D; Representation of ribosomal DNA in the *Ae.aegypti* genomic library; Ethidium bromide stained gel and corresponding autoradiograph of 5µg of genomic library DNA and 2µg of genomic DNA digested with EcoR1,transferred to nitrocellulose and probed using a cloned rDNA repeat. L= genomic library DNA, G= genomic DNA.

of 9.0 Kb on the Southern transfer (the 7.2 Kb and 1.8 Kb bands are products of the cleavage of this DNA species). No other qualitative differences are evident between the two digests. Some hybridisation is observed to the λ arms. This is probably due to a low level of contamination of the gel-cut rDNA probe with λ vector DNA and is also a consequence of the partial digestion.

The ethidium bromide stained gel of the library DNA shows a series of distinct bands in addition to the expected genomic DNA smear. These bands represent sequences which have been preferentially cloned and/or grow fastest as recombinants. These bands include a 4.9 Kb band which probably corresponds to the λgtWES 'stuffer' fragment (present in small amounts relative to the λ arms) and bands of 7.2, 8.4, 9.5 and 14 Kb. One of these bands (7.2 Kb) corresponds to a repetitive sequence band on the genomic digest stained gel and may also correspond to the rDNA band of this size visible on the autoradiograph.

1.3. Organisation and Expression of Repeated Sequences in the Mosquito *Aedes aegypti*: Studies Involving Screening of Genomic Clones

Introduction

The genomic organisation and expression of eukaryotic repeated DNA sequences has been studied for a variety of organisms (reviewed in [23]). Each species has a distinct complement of repeated sequences and the precise arrangement of these may differ considerably from species to species. Most organisms may be classified into one of two groups; those with a short period interspersion (S.P.I.) pattern of repeated sequences and those with a long period interspersion (L.P.I.) pattern of repeats. DNA from organisms with a S.P.I. pattern (e.g. sea urchin [24], *Xenopus* [25] and all mammals studied to date [26]) consists largely of short stretches of unique DNA sequence 1-2 Kb in length separated by short repeated sequences 100-300 bp in length. The frequency of occurrance of these short repeats is variable; some regions of DNA may contain significantly fewer than average repeats whereas other regions may contain significantly more. Long stretches of tandemly repeated sequences may be devoid of short interspersed repeats. Organisms with a long period interspersion pattern of repeated sequences include *Drosophila* [27], *Apis* [28] and *Chironomus* [29]. Repeated sequences in these organisms are much longer (typically 5 Kb) and are interspersed with much longer stretches of unique DNA (e.g. 30 Kb). As for short interspersed repeats, the

frequency of interspersion of long repeats is likely to vary. Long interspersed repeats are found in addition to short interspersed repeats in organisms with a S.P.I. pattern [30]. In the mouse, Young [31] has estimated that there are 70 different sequence families of long dispersed repeats which are present in 3-100 copies per haploid genome and represent approximately 17% of the total DNA. *Drosophila* DNA is composed of about 12% middle repetitive DNA, of which one quarter codes for tandemly repeated genes (ribosomal RNA and histones) and the remainder is composed of approximately 40 sequence families of dispersed long repeats [32].

The short interspersed repeats of humans are dominated by a single sequence family; the 'Alu' sequences. Sequences homologous to 'Alu' are also found in other mammals. 'Alu' repeats comprise one half to one third of all short repeats in the human genome and their counterparts are equally abundant in other mammalian genomes [33]. In humans there are approximately 300,000 copies of the 300 bp long 'Alu' sequence, equivalent to 3% of the total genome. Different members of the 'Alu' sequence family are similar, but not identical in DNA sequence.

Long and short dispersed repeated sequences differ in their structure and transcription. Long repeated sequences frequently resemble the retrovirus integrated provirus. They may have long terminal repeats (225-400 bp in length), which in turn may be flanked by short inverted repeats. Homology has been reported between the open reading frames found within long dispersed repeats (the L1 rodent long dispersed repeats for example) and the open reading frames of retroviruses [34]. Short repeated sequences do not possess these characteristics and do not encode proteins (the proposed reverse

transcriptase for example of long dispersed repeated sequences). Both short and long dispersed repeats are considered to be mobile genetic elements due to their variable locations within the genome and both are flanked by a short direct repeat (7-20 bp for 'Alu'-like sequences, 4-6 bp for long dispersed repeats) which is thought to represent a duplication of DNA produced upon integration of the element into a new site [23].

Transcription of long dispersed repeats such as the 'Copia'-like elements in *Drosophila melanogaster* by RNA polymerase II generates abundant polyadenylated transcripts of discrete sizes [35]. 'Alu' sequences are heterodisperse in mRNA and hnRNA as a result of transcription by RNA polymerase II as part of other transcription units (i.e. within introns) or due to 'promoter leakage' which leads to the transcription of repeats adjacent to transcribed regions [36].

'Alu' sequences are also transcribed by RNA polymerase III as discrete transcription units [37]. A small RNA species (7S) which contains extensive homology to 'Alu' RNA is a vital component of a cytoplasmic complex involved in the transport of newly synthesised proteins across membranes [38]. 'Alu' transcripts constitute a considerable proportion of hnRNA in mammalian cells [39] and are also present in reduced, but still significant amounts in cytoplasmic mRNA [40].

Transcription of repeated DNA sequences into hnRNA has been reported for all eukaryotes studied. The presence of short interspersed repeated sequence transcripts in mRNA however has been reported only in mammals. The short interspersed repeats of the sea urchin for example, are expressed in hnRNA, but are not present in sea urchin mRNA [41].

An estimate of the genome complexity (quantity of DNA per haploid genome) of an organism may be made from the reassociation kinetics of sheared DNA in solution [42]. The proportion of the genome which is highly repetitive, moderately repetitive or unique may be estimated from the reassociation kinetics of these fractions [43]. For genomes of mammalian sequence complexity (around 3 x 10^9 bp or DNA per haploid genome), highly repeated is considered to be 10^5 - 10^6 copies per haploid genome and middle repetitive to be 10^2 - 10^4 copies per haploid genome. The copy number of a specific sequence may be estimated from the reassociation kinetics of the purified sequence or its RNA transcript to denatured genomic DNA in large excess [44]. Similarly, the sequence complexity of a mRNA or hnRNA population may be estimated by hybridisation of the RNA or its cDNA transcripts to genomic DNA in excess [45].

Solution hybridisation studies of this type have been undertaken for the mosquito. Analysis of genomic complexity of *Aedes albopictus* [46], *Aedes aegypti* (A.M. Warren unpublished results) and *Culex pipiens quniquefasciatus* (J.M. Crampton unpublished results) has been performed. In addition, Spradling *et al* [46] have estimated ribosomal gene copy number and sequence complexity of hnRNA and mRNA of *Aedes albopictus*.

In this section, I describe some information regarding the genomic organisation and expression of repeated sequences in the mosquito *Aedes aegypti* obtained by a different approach: screening the *Ae. aegypti* genomic library described in Section 1.2 with genomic DNA, various classes of RNA and a ribosomal DNA probe. The results of these experiments are compared to those obtained by the other methods described and their implications are discussed.

Experiments and results

Approximately 10,000 clones from the *Aedes aegypti* genomic library amplified stock were plated onto a B.R.L. 'bioassay plate' measuring 22.5 cm x 22.5 cm. Three nitrocellulose filter replicas (20 cm x 20 cm) of this plate were taken (see Materials and Methods). After reincubation of the plate (3 hours, 37°C), a further three replicas were taken. One replica filter was hybridised with each of the following radiolabelled probes:

a. total genomic DNA

b. total cytoplasmic RNA

c. poly A^+ cytoplasmic RNA

d. nuclear RNA

e. ribosomal DNA

Probes a-d were all prepared from *Aedes aegypti* (Bangkok strain) fourth instar larvae (the same material that was used for the construction of the genomic library). The ribosomal DNA used was the insert of clone λAar7 which contains an entire 18S/28S ribosomal DNA repeat of *Aedes aegypti* (see Section 2).

Nucleic acids were prepared and labelled as described in Materials Methods. Probes were labelled to specific activities of 10^8 dpm/µg for DNA, 2-5 x 10^7 dpm/µg for RNA, and were used at a concentration of 50 ng per ml of hybridisation solution. Hybridsations using complex probes (a-d above) were carried out over the weekend (60 hours). Homogeneous probes ('e' for example) were usually hybridised overnight (18-24 hours).

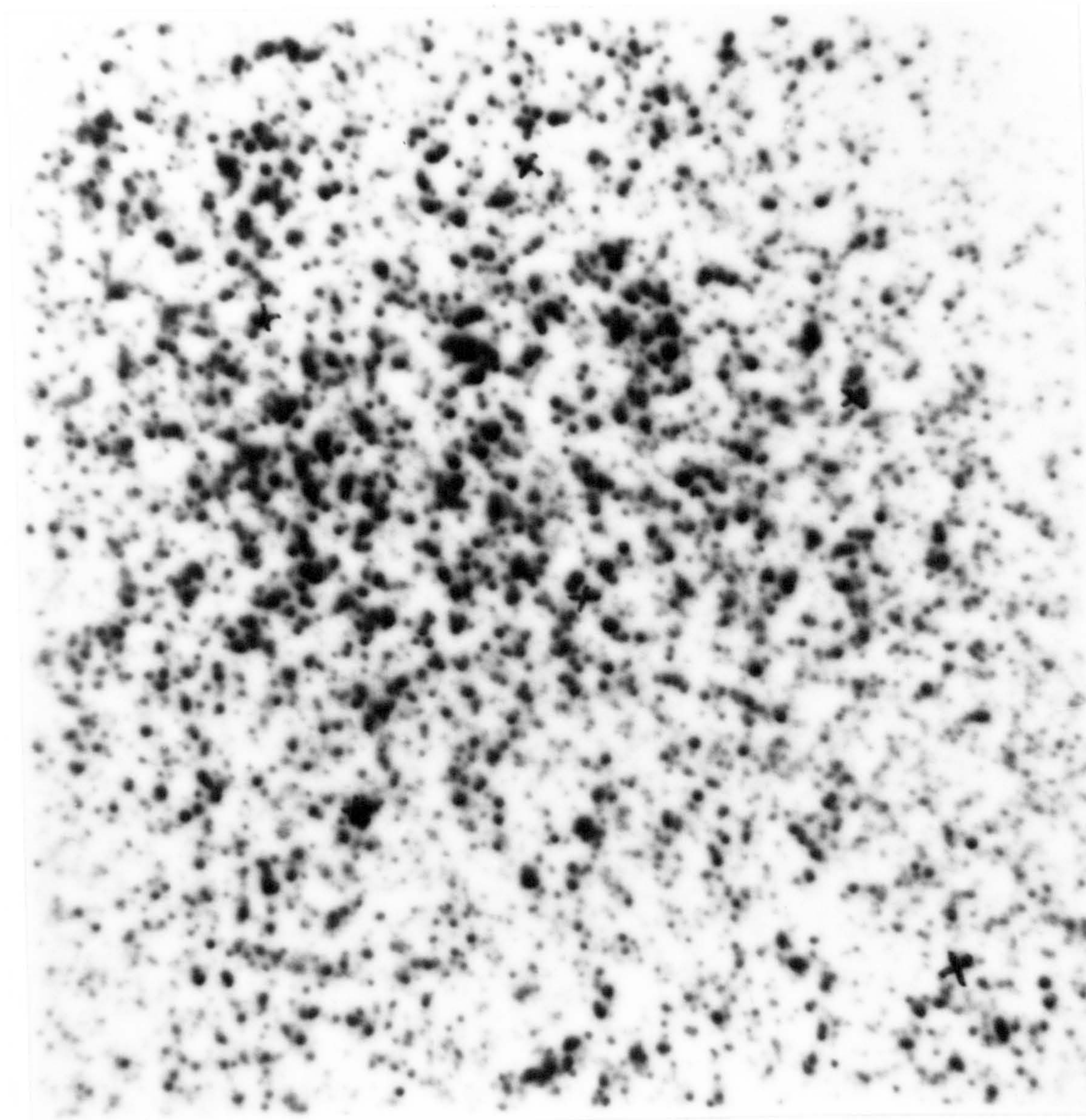


Fig. 1.3A

Figure 1.3A; Autoradiograph of approximately 7100 clones from the *Ae. aegypti* genomic library screened using radiolabelled *Ae.aegypti* genomic DNA.

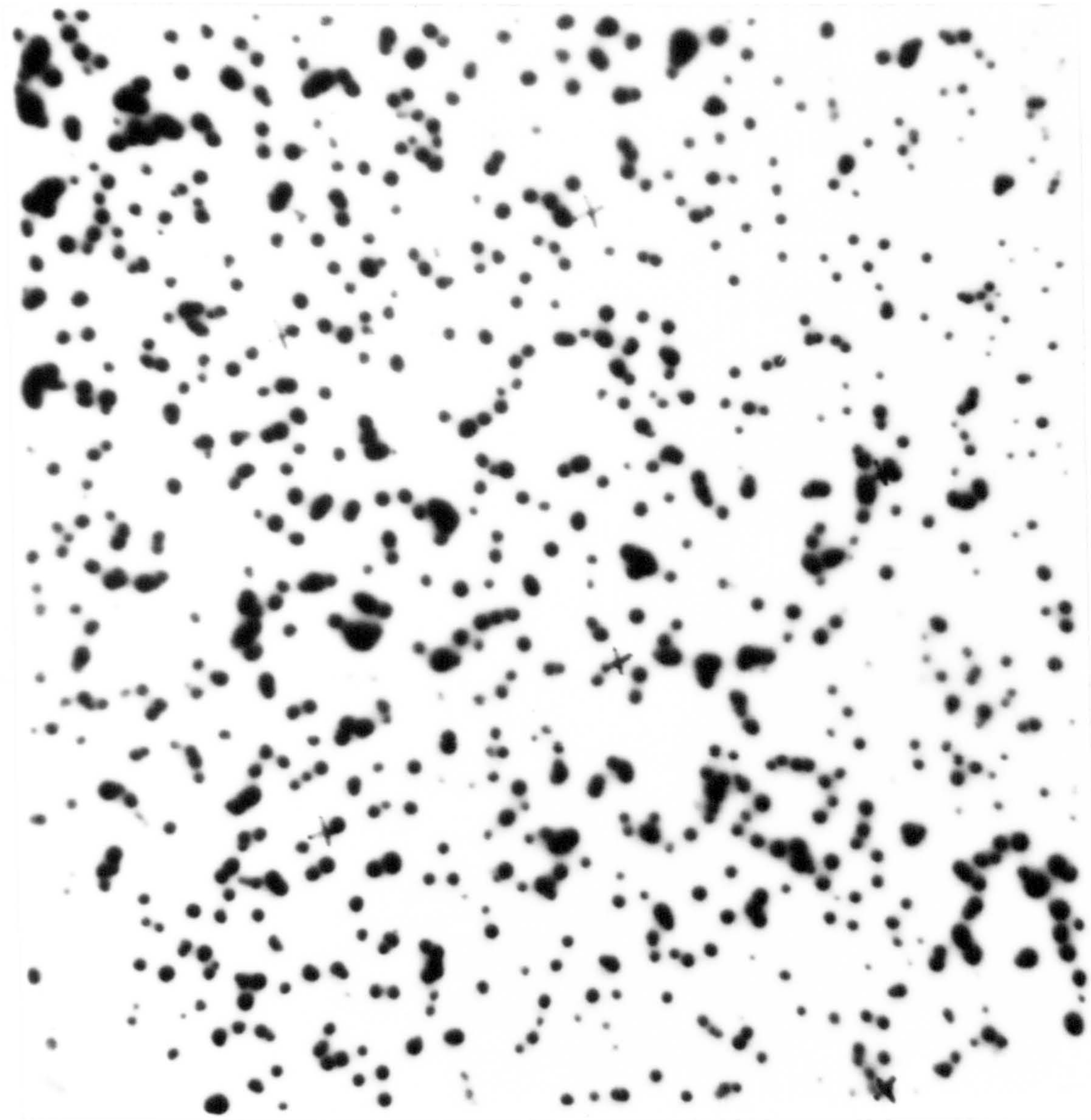


Fig.1.3 B

Figure 1.3B; Autoradiograph of approximately 7100 clones from the *Ae. aegypti* genomic library screened using radiolabelled *Ae.aegypti* total cytoplasmic RNA.

When screened using total genomic DNA, the majority of clones in the genomic library gave a detectable hybridisation signal. (Figure 1.3A). Of 7900 clones screened, approximately 7100 (90%) were positive with this probe.

When probed using total cytoplasmic RNA, a much smaller proportion of clones (120, 14%) showed hybridisation signal. (Figure 1.3B). These clones gave very strong autoradiographic signals, even after short exposures. Clones which gave a hybridisation signal with the total cytoplasmic RNA probe corresponded to clones which gave a medium strength hybridisation signal with the total genomic DNA probe.

When replica filters were screened using poly A^+ RNA (selected by oligo-dT column chromatography) or nuclear RNA prepared by guanidine hydrochloride extraction from purified nuclei (see Materials and Methods), the pattern of hybridisation was qualitatively identical to that obtained using total cytoplasmic RNA.

When RNA probes derived from Mos20A *Ae. aegypti* cell-line (which was established in 1969 and resembles epithelial cells [47]) were used to probe replica filters, results were identical to those obtained using larval RNA for all three RNA classes (results not shown).

Cellular RNA consists mainly of ribosomal RNA (80 - 85%), therefore clones containing ribosomal RNA coding regions (rDNA) will give strong hybridisation signals with the RNA probes used here. (Even $polyA^+$ selected RNA and nuclear RNA preparations are likely to contain significant proportions of rRNA). To determine the proportion of RNA positive clones which contain rDNA, a mosquito ribosomal DNA repeat was used to probe one of the replica filters.

Almost all clones which showed hybridisation signal with the RNA probes were positive with this ribosomal DNA probe. Of the 1120

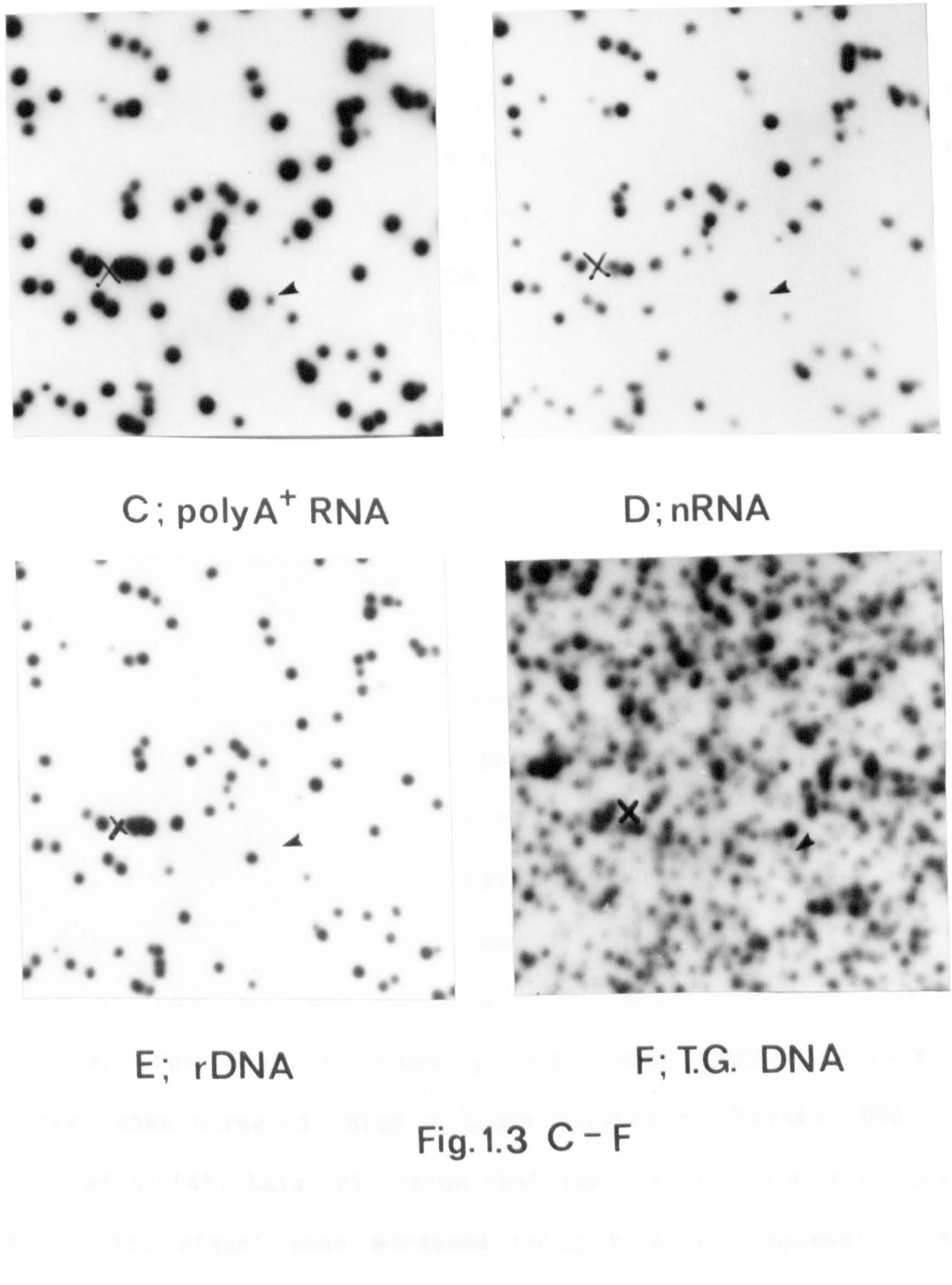


Fig. 1.3 C - F

Figure 1.3C-1.3F; Sections of replica phage lifts showing the hybridisation properties of a highly expressed, non-ribosomal DNA clone (arrowed). C; polyA$^+$ selected RNA probe, D; nuclear RNA probe, E; ribosomal DNA probe, F; total genomic DNA probe. Probes used for C, E and F were prepared from *Ae.aegypti* (Bangkok strain) larvae. The probe used for D was prepared from Mos20A *Ae.aegypti* cell-line.

clones positive with the RNA probes, only one was not positive with rDNA. This clone showed relatively weak hybridisation signal with the total cytoplasmic and nuclear RNA probes, but stronger hybridisation signal with poly A^+ RNA. No differences in intensity of hybridisation of this clone were evident between the RNA probes derived from larvae and cell-line material. This clone did not correspond to a positive signal on the genomic DNA screening autoradiograph. (These results are shown in Figure 1.3 C-F)

Discussion

In a genomic library constructed for an organism with a short period interspersion pattern of repeated sequences, most recombinants containing large inserts would be expected to contain several short repeats. In the case of a human genomic library containing DNA fragments 10-20 Kb in length, approximately 95% of recombinants contain at least one repeated sequence and most contain several [13]. A similar proportion of clones in this human genomic library were positive when screened using a human repetitive fraction DNA probe. Gergen *et al* [48] have calculated that for a clone to give a positive hybridistion signal when screened using a multi-component, complex hybridisation probe, it must contain some DNA sequence which comprises at least 0.1% of the total probe. For practical purposes, using reasonable probe concentrations, hybridisation times and autoradiography exposures, the authors show that the limit may be as high as 0.5% of the total probe. The minimum copy number of sequences which give detectable hybridisation signal when total genomic DNA is

used to screen a genomic library is determined by the sequence complexity of the genomic DNA used. For example, 0.1% of a genomic DNA probe is equivalent to 3000 copies per haploid genome of 1 Kb sequence in human DNA (3 x 10^9 bp of DNA per haploid genome), but only 800 copies of a 1Kb sequence in sea urchin DNA (8 x 10^8 bp of DNA per haploid genome). Similar criteria apply to the detection of sequences homologous to complex RNA probes. In this case the complexity of the RNA probe determines the minimum transcript copy number which gives detectable hybridisation signal.

When the *Ae. aegypti* genomic library (insert size; 7-15 Kb, background of non-recombinants; 2%) was screened using total genomic DNA, 90% of clones showed hybridisation signal. Therefore 90% of clones must contain some sequence which represents 0.1% or more of the total genomic DNA probe. Spradling *et al* [46] estimated the genome complexity of the mosquito *Aedes albopictus* to be approximately one third of that of the mammalian genome (e.g. 8 x 10^8 bp) although the proportions of highly repeated (10%), middle repetitive (10-15%) and unique DNA (75-80%) were similar. Similar results have been obtained for the species *Culex pipiens quinquefasciatus* and *Aedes aegypti* (J.M. Crampton, A.M. Warren, personal communications). A mosquito genomic DNA clone containing 15 Kb of unique sequence is homologous to only 0.002% of a total genomic DNA probe. Therefore clones containing unique sequences only will not give detectable hybridisation signal when screened using a total genomic DNA probe. Hence 90% of the clones in the mosquito genomic library must contain some repetitive DNA sequence. This is strong evidence for the presence of a short period interspersion pattern of repeated sequences in the *Aedes* genome.

No estimate of the length of repeated sequences can be made from this data presented here, although Hogan [49] has reported unpublished results communicated by S. Penman that on average, mosquito repeats are 300 bp long with an interspersion interval of 2000 bp. For a 300 bp sequence to comprise 0.1% of the mosquito genome, the reiteration frequency would be 2700 copies per haploid genome.

The finding that *Aedes aegypti* has a short period interspersion pattern of repeated sequences supports the theory that a S.P.I. pattern is usually associated with a relatively large genome size in higher eukaryotes. *Dictyostelium* is the only eukaryote known to have S.P.I. and a small genome size [50]. The species *Drosophila* [32] and *Chironomus* [29] which are both dipterans, and *Apis* [28] which is a hymenopteran, do not have S.P.I. patterns of repeated sequences and have small genome sizes (9×10^7 bp, 1.8×10^8 bp and 2.7×10^8 bp respectively). *Aedes aegypti* and all other species studied, including *Musca* [28], a dipteran, and *Antheraea* [51], a lepidopteran, have a S.P.I. pattern of repeats and relatively large genomes (8×10^8 bp or greater).

Tashima *et al* [13] have demonstrated that when mammalian genomic libraries are screened using cDNA transcribed from cytoplasmic messenger RNA, the proportion of clones showing hybridisation signal is similar to that obtained using a genomic DNA probe. This is because 'Alu'-like sequences are transcribed and represent significant proportions of total cytoplasmic, hnRNA and poly A^+ cytoplasmic RNA. For a human genomic library containing 20Kb inserts, 95% of clones give a hybridisation signal with cDNA transcribed from poly A^+ RNA of human cells [13]. The authors show that this is due to the presence of repetitive sequences in a sub-set of the cDNA molecules.

When screened using high specific activity mosquito total cytoplasmic RNA, poly A$^+$ cytoplasmic RNA or nuclear RNA, only 14% of *Ae. aegypti* genomic clones gave detectable hybridisation signal. The vast majority of these (99.9%) correspond to ribosomal DNA clones, as visualised using the mosquito ribosomal DNA probe. This result indicates that transcripts of the dispersed repeats which show hybridisation signal with the genomic DNA probe do not constitute a significant proportion (more than 0.1%) of any of the RNA probes used here. A similar result was obtained when cDNA transcribed from poly A$^+$ RNA of *C. p. quinquefasciatus* was used to screen a genomic library of this species. (J.M. Crampton, personal communication).

Spradling *et al* [46] report that although the level of transcription of repetitive sequences into hnRNA is similar for the mosquito and mammals (about 30%), the level of repetitive sequence transcripts in mRNA is higher in mammals (less than 10% of their mosquito mRNA preparation was repetitive compared to 30% for mammalian mRNA). The nuclear RNA probe used in the present studies to screen the *Ae. aegypti* genomic library does not detect transcription of repeated sequences. This may be due to the presence of a large proportion of contaminating ribosomal RNA in the hnRNA and poly A$^+$ cytoplasmic RNA probes. This would substantially decrease the sensitivity of these as probes for the detection of clones containing non-rDNA transcribed sequences. Further work, including the use of more highly purified hnRNA (Spradling *et al* separated nucleoli and nucleoplasm by sucrose gradient centrifugation) which could be verified by assaying for rRNA and determining the molecular weight distribution of the RNA, is needed to clarify the degree of transcription of mosquito repeated DNA sequences.

The finding that almost all positives obtained using the RNA probes contain rDNA has several important implications. Firstly, it explains why no differences were observed when replica filters were screened using cell-line and larval RNA probes; rRNA is highly expressed in all tissues and evolutionarily conserved. Secondly, the proportion of clones containing rDNA (14%) is far larger than expected. In *Aedes aegypti*, ribosomal genes are reiterated approximately 500 times per haploid genome (see Section 2 of this thesis). The rDNA repeat is 9.0 Kb in length and rDNA therefore represents approximately 0.6% of the total *Aedes* genome. Ribosomal RNA coding sequences are therefore greatly over-represented in the *Ae. aegypti* genomic library. This may be due to the organisation of *Ae. aegypti* rDNA with respect to EcoRI recognition sites. Clonable rDNA fragments 7-15 Kb in length are produced on complete and partial digestion of *Aedes aegypti* genomic DNA (see Section 2.9).

Other abundant cellular RNAs (excluding abundant mRNAs) such as 5S rRNA, transfer RNAs and small nuclear RNAs are shown to be under-represented or possibly absent from the genomic library by the rDNA screening results. (DNA coding for these should comprise at least a small proportion of clones which give hybridisation signal with the RNA probes). For example, 5S rRNA, which is present in equimolar ratio with 18S, 5.8S and 28S rRNA in higher eukaryotes, represents approximately 2% of cellular RNA [52]. The genes coding for 5S rRNA and their non-transcribed spacer regions are short and tandemly repeated (*Drosophila* has 160 copies, 373 bp in length, for example [53]). There is a low probability that a tandem repeat this length will be cut by EcoRI (hexacutters would cleave random DNA once every 4096 bp on average) and if this is the case, 5S coding sequences will

not generate clonable-sized EcoRI fragments. This illustrates the advantages of using tetracutter enzymes (which would cleave random DNA once every 256 bp on average) or mechanical shearing to produce insert DNA for genomic libraries.

All rDNA-containing clones correspond to positives obtained with the total genomic DNA probe (medium intensity signal). As ribosomal genes represent 0.6% of the genomic DNA probe, this indicates that the sensitivity of screening is higher than 0.6%.

The very small proportion of clones (0.1%) which show hybridisation signal with the complex RNA probes, but do not contain rDNA and do not show hybridisation signal with genomic DNA is expected for low copy number, highly expressed genes. For example, in the sea urchin (which has a similar genome size to the mosquito; 8 x 10^8 bp) only 2.7% of unique sequences are transcribed. These unique sequence transcripts are composed of approximately 8500 different mRNA species of which fewer than 100 are abundant (e.g. 1000 copies per cell or higher) [52]. The number of clones screened in this experiment represents only a small proportion of the *Aedes* genome (see Section 1.2) and therefore very few low copy number, highly expressed sequences will have been screened. The clone which clearly does contain a non-rDNA, highly expressed sequence, gives a stronger hybridisation signal with poly A^+ selected cytoplasmic RNA than total cytoplasmic RNA and may therefore represent a sequence which is highly expressed as polyadenylated mRNA in mosquito cells and larval tissues.

Future Perspectives

The presence of an S.P.I. pattern of repeated sequences in the *Aedes aegypti* genome has some important implications for further studies. A genomic clone containing a specific DNA sequence is likely to also contain at least one dispersed repeat. This is especially true for clones with large inserts. The presence of a repetitive sequence will drastically alter the hybridisation characteristics of a clone and may mask the hybridisation characteristics of the sequence of interest. In such an instance, analysis of different segments within the genomic DNA fragment may be necessary. The presence of S.P.I. repeats will make 'chromosome walking' (by looking for overlapping genomic clones) considerably more difficult for similar reasons.

When screening *Aedes* genomic libraries for DNA sequences with specific hybridisation characteristics (species-specific DNA probes for example; Section 4), short DNA fragments should be used for both the library insert DNA and probe DNA to avoid the masking of one repetitive DNA sequence by the hybridisation characteristics of another closely linked repeat.

When screening *Aedes* genomic libraries using RNA probes, the transcription of repeated sequences is not sufficient to mask that of very highly expressed genes. To screen genomic clones for sequences transcribed at lower levels, RNA probes devoid of ribosomal RNA, or the use of cDNA probes of higher specific activity is required. It is unlikely that rRNA or rRNA cDNA could be removed from any class of RNA or cDNA probe with an efficiency which would prevent this strong hybridisation to rDNA clones. To screen large numbers of highly expressed genes (e.g. 100) in the *Aedes* genomic library, a large

number of amplified clones (e.g. several genome equivalents) must be screened. At very high plaque density, the hybridisation signal from rDNA-containing clones is likely to mask signals from much less abundant sequence classes which may also be highly expressed. One technique which would circumvent this problem is to design a genomic library which would be devoid of rDNA clones. This is possible from the information obtained in Section 2 of this thesis. Two restriction enzymes (SalI and XhoI) have been identified which do not cut within the tandemly repeated rDNA. These enzymes would therefore not generate rDNA fragments of clonable size (e.g. 2-20 Kb for bacteriophage vectors) when used to digest *Ae.aegypti* genomic DNA. A genomic library constructed using either or both of these enzymes to generate genomic insert DNA should therefore contain very few rDNA clones.

Further information regarding the lengths and interspersion frequences of repeated sequences may be obtained by analysing the inserts of a number of genomic clones. The hybridisation properties of some specific repeated sequence from *Anopheles* mosquitoes are described in Section 4. Hybridisation of specific cloned repeats to Northern transfers of various RNA classes may be the most sensitive method for assessing the transcription of repeated sequences.

Section 2: Sequence Organisation of Ribosomal DNA in *Aedes aegypti*

2.1. Introduction

The genes encoding the structural RNA components of ribosomes are among the most extensively studied DNA sequences. In higher eukaryotes, ribosomal genes are present in multiple copies and may undergo amplification in some cell types (in the *Xenopus* oocyte [54], for example). This allowed the isolation of pure ribosomal genes for detailed biochemical analysis prior to the advent of recombinant DNA technology. Ribosomal RNA is the most abundant RNA species of cells, representing 80-85% of total cellular RNA, present in complex ordered structures (the ribosomes) which are central to the process of mRNA translation. Ribosomes may be purified from other cellular components on the basis of their sedimentation rates (ribosomal sub-units and RNA components are named according to their sedimentation coefficients in Svedberg units).

This section describes the cloning and characterisation of the major rDNA repeat from *Aedes aegypti*. The genes encoding the 5S rRNA component (which is present in equimolar amounts with the larger RNAs in the ribosome) are found in a separate gene cluster in higher eukaryotes (reviewed in [53]) and are not included in this study.

In most of the eukaryotes studied, the genes encoding the 18S, 5.8S and 28S ribosomal RNAs (or their equivalent) are organised as

tandemly repeated arrays in the nucleolar organiser region(s) of the chromosomes [55]. The integrated ribosomal DNA (rDNA) repeats in higher eukaryotes are usually present in 50-500 copies per haploid genome, arranged in a tandemly repeated, head-to-tail manner [23]. The rDNA repeat consists of coding regions and internal and external transcribed spacer regions separated by a non-transcribed spacer region. The length of the repeat varies from 8.4 Kb in the primitive diptera *Sciara coprophila* [56] and *Chironomus tentans* [57] to 38 Kb in *Dictyostelium discodium* [58] and 44 Kb in humans [59] and mice [60]. This large variation in size is due mainly to variation in the lengths of spacer sequences, particularly of the non-transcribed spacer (NTS) region. The rRNA is transcribed as a single precursor molecule of approximately 8 Kb in most eukaryotes [61-63], although the rRNA precursor is longer (around 13 Kb) in mammals [64]. Pre-rRNA is cleaved to yield mature 18S, 5.8S and 28S rRNAs. In insects the 28S rRNA undergoes additional cleavage to yield two sub units (28S α and 28S β) [65]. In some insect species, the 5.8S rRNA also undergoes further cleavage. In *Drosophila melanogaster*, the 5.8S rRNA is cleaved to give 5S and 2S rRNA molecules [66].

Comparisons of ribosomal DNA sequence between different species have shown that while regions within rRNA coding sequences are evolutionarily very highly conserved [67], the sequence of the spacer regions is highly variable even between closely related species [68]. The inference originally drawn from this finding was that spacer regions have no sequence-specific function. Recent studies on the effect of the N.T.S. region on transcription of *Xenopus* rRNA however, have shown that N.T.S. sequences of this

species contain transcriptional enhancers [69,70] which effectively load RNA polymerase I onto the rRNA promoter.

In addition to the observed inter-species variation of ribosomal DNA spacer sequences, many species exhibit a high degree of rDNA intra-species polymorphism. In *Locusta migratoria*, extensive restriction site polymorphism has been observed in the long (11 Kb) non-transcribed spacer region [71]. Within *Xenopus laevis*, different rDNA repeats vary in length. This is due to differences in the number of repeated sequences which comprise the non-transcribed spacer of this species [72].

In most of the insects studied, there is an additional source of length heterogeneity: a proportion of the genes contain insertions of non-ribosomal DNA. The position, sequence and length of non-rDNA insertions within a species is variable [71, 73-78]. In *Drosophila melanogaster*, roughly half of the 28S genes contain insertions of two distinct types (Type I and Type II) which are not homologous in sequence. Type I insertions are flanked by a direct repeat of 11 or 14 bp which represents a duplication of rDNA at the site of insertion [79]. Type I insertions are also found outside ribosomal DNA [80,81] and may be derived from a transposable genetic element. Type II insertions are not flanked by repeats [82] and have not been reported outside rDNA.

In some lower eukaryotes, all of the 28S rRNA genes contain insertions of non-rDNA [83,84]. These cannot therefore inactivate the genes and may be termed 'introns' in view of their obvious analogy to the introns of eukaryotic protein-coding genes. There is however considerable evidence that the insertions in *D. melanogaster* rDNA inactivate the rDNA repeat [85]. Some rDNA insertions generate

short deletions of rDNA coding sequence which would preclude them from yielding functional rRNA [86]. In the species studied, intron-containing rDNA repeats are clustered together [87] and appear to be present in a transcriptionally inactive state [88,89]. In support of this, it is observed that the proportion of rDNA repeats which contain insertions is lower in dipterans with fewer rDNA gene copies, i.e. interrupted repeats may be selected against in this situation [56,90].

Ae. aegypti has a single nucleolus organiser located on chromosome 3 [91]. The ribosomal gene copy number for a closely related species *Aedes albopictus* has been estimated as 430 per haploid genome [46]. Studies by Shine and Dalgarno [65] on the heat-dissociation of rRNA of *Ae. aegypti* and *Ae. albopictus* have shown that the 26S rRNA of these species (eqivalent to 28S rRNA of mammals) is composed of two polyribonucleotide chains of similar molecular weight to the 18S rRNA, which are hydrogen-bonded together in the ribosome. A small RNA (approximately 150 nucleotides) is also hydrogen-bonded to the 26S rRNA (this occurs with mammalian 28S and 5S rRNA).

Of particular interest in the study of *Ae. aegypti* ribosomal genes presented here, was the degree and the source of ribosomal gene heterogeneity. Also, to establish whether mosquito ribosomal genes contain insertions of non-ribosomal DNA, and if so, to determine whether these are related to transposable genetic elements. A number of eukaryote transposable elements (distinct from rDNA introns) have been found inserted into ribosomal genes [92-94]. The possibility of locating transposable elements inserted into mosquito ribosomal genes is examined.

2.2. Preliminary studies on *Ae. aegypti* rDNA organisation by Southern blotting

Experiments and Results

Aedes aegypti (Bangkok strain) genomic DNA was prepared from 400 pooled fourth instar larvae as described in Materials and Methods. A number of different restriction endonucleases were used to digest 4 µg aliquots of this DNA to completion. Each sample was then fractionated through 0.8% Agarose and the gel blotted onto nitrocellulose. This filter was then hybridised with nick-translated pDm238 clone DNA (kindly supplied by D. M. Glover), which contains an intact ribosomal DNA repeat from *Drosophila melanogaster* [73]. The resulting autoradiograph (not shown) indicated a number of discrete bands of hybridisation signal. Four of the enzymes used (HbaI, PvuII, SacI and HindIII) all gave a single band of hybridisation signal corresponding to a length of 9.0 Kb. Two enzymes (EcoRI and PstI) gave two bands of sizes which when summed equalled 9.0 Kb. One enzyme (HincII) generated a more complex set of bands which when summed equalled significantly more than 9.0 Kb. Two enzymes (SalI and XhoI) gave high molecular weight smears on the autoradiograph with no distinct bands.

Conclusion

For four different restriction enzymes to generate a band of the same size on the autoradiograph (9.0 Kb), each must cut once within the tandemly repeated rDNA unit and this must be 9.0 Kb in length. Similarly, enzymes which give two bands, which when summed equal 9.0 Kb must cut twice within the ribosomal repeat. The enzymes which do not generate discrete bands do not cut the ribosomal repeat. The result obtained with HincII (a complex set of bands) is explained by the results of subsequent experiments.

2.3. Restriction mapping the *Aedes aegypti* rDNA repeat by Southern blotting

Experiment and Results

Restriction enzymes which cut the rDNA repeat once or twice were used singly and in combination to restrict *Ae. aegypti* genomic DNA which was then electrophoresed through 0.8% Agarose, blotted onto nitrocellulose and probed using pDm238 clone DNA. The resulting autoradiograph is shown in Figure 2.3. The sizes of the bands generated are as follows:

EcoRI + HindIII	:	7.0 + 1.8 Kb
EcoRI	:	7.2 + 1.8 Kb
EcoRI + SacI	:	6.2 + 1.8 Kb
SacI	:	9.0 Kb
HindIII	:	9.0 Kb
HindIII + PvuII	:	8.0 Kb
PvuII	:	9.0 Kb
PvuII + SacI	:	7.0 + 2 Kb

From this data it was possible to construct the preliminary restriction map shown in Figure 2.3.

Figure 2.3; Restriction mapping of the *Ae.aegypti rDNA repeat by Southern blotting. Restricted genomic DNA was size fractionated through 1% Agarose, transferred to nitrocellulose and probed using a cloned Drosophila* rDNA repeat (pDm328, kindly supplied by D.M.Glover). Fig. 2.3a shows the resulting autoradiograph and fig. 2.3b shows the derived restriction map.

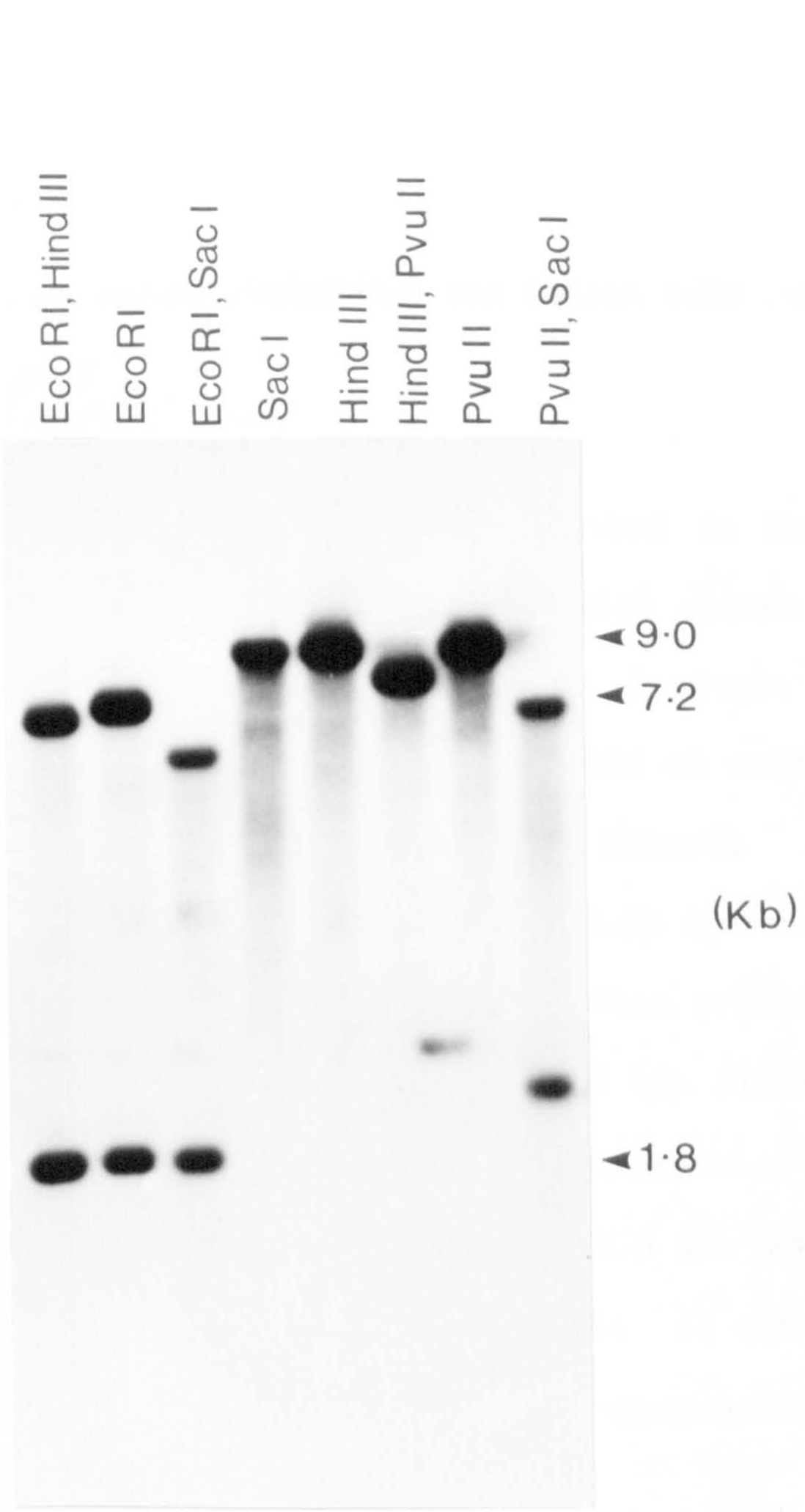


Fig.2.3a

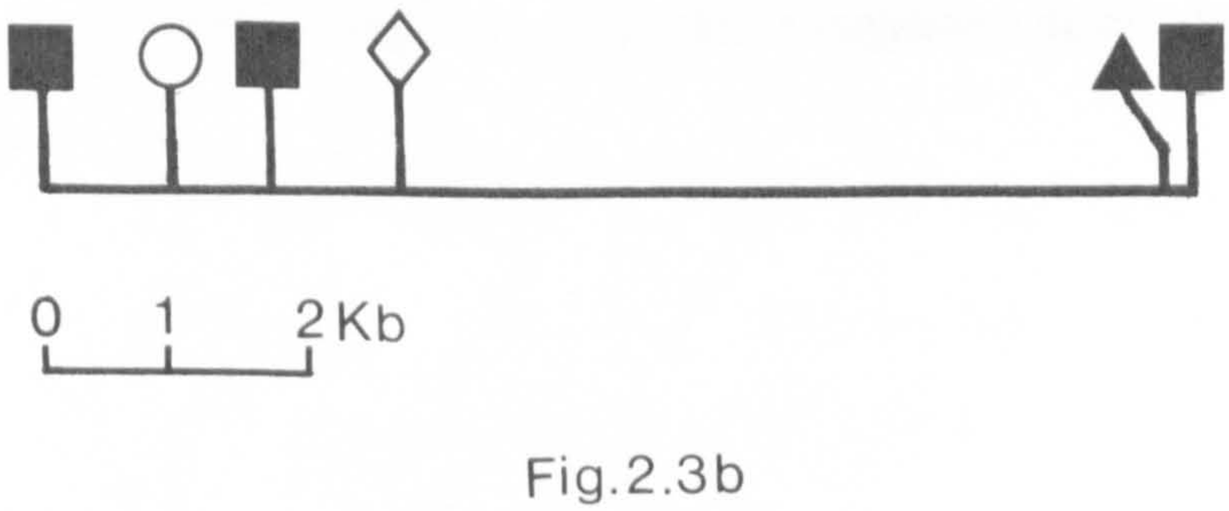


Fig.2.3b

2.4. Isolation of clones containing the intact rDNA repeat of *Aedes aegypti*

The *Aedes aegypti* genomic library described in Section 1 was screened using pDm238 clone DNA. Ten positive clones were picked and replated to give isolated phage plaques. A single fresh plaque from each clone was then used to inoculate 500 ml cultures and the DNAs prepared as described in Materials and Methods. 1 µg of each clone DNA was restricted with EcoRI, run in 0.8% Agarose and the gel blotted onto nitrocellulose. This blot was then probed with pDm238 to confirm the clones as ribosomal. Nine of the clones were found to contain ribosomal DNA. Of these, three contained a 7.2 Kb insert fragment and six contained a 7.2 Kb and a 1.8 Kb insert fragment. From the preliminary restriction mapping data, it was deduced that the latter six clones contain sequences representing the entire *Aedes aegypti* ribosomal DNA repeat.

The clones containing the intact *Ae. aegypti* rDNA repeat were designated λAar1, 3, 5, 7, 8 and 9.

The clones containing the 7.2 Kb fragment only were designated λAar4, 6 and 10.

2.5. Restriction mapping of ribosomal clones

One of the ribosomal clones (λAar 3) containing the intact *Ae. aegypti* rDNA repeat (7.2 Kb and 1.8 Kb EcoRI fragments) was restriction mapped as follows. 100 µg of clone DNA (corresponding to 20 µg of insert DNA) was digested with EcoRI and size fractionated through 0.8% low melting Agarose. Preparative well-formers, each equivalent to 30 normal wells of 40 µl volume were used. After electrophoresis (25 volts, 15 hours) the two insert DNA bands were excised and the DNA extracted as described in Materials and Methods. A final yield of 7.8 µg of the 7.2 Kb restriction fragment and 2.2 µg of the 1.8 Kb fragment was obtained (50% recovery). These two DNA fragments were then restriction mapped separately using the following enzymes singly and in combination; SacI, HindIII, PvuII, HincII, KpnI and PstI. The restriction map obtained is shown in Figure 2.5 a. This map confirms the findings of the data obtained in the preliminary mapping experiment using Southern blotting data and provides a more detailed physical map of the *Ae. aegypti* rDNA repeat.

Insert fragments were isolated from three additional clones, λAar1, λAar7 and λAar9, which all contain the intact *Ae. aegypti* rDNA repeat (7.2 Kb and 1.8 Kb EcoRI fragments). These insert DNAs, together with λAar3 insert DNA, were subjected to restriction enzyme digestion with the enzymes already used to map λAar3. When all digestions were run in parallel, only one difference was detected between the four cloned rDNA repeats. λAar9 was found to contain an extra HincII site. This results in the cleavage of the 3.4 Kb HincII/EcoRI fragment found in λAar1, λAar3 and λAar7 into a 2.0 Kb

Hinc II fragment and a 1.4 Kb Hinc II/EcoRI fragment in λAar9. The complexity of the HincII digestion of genomic rDNA (described in Section 2.2) stems partly from this restriction site polymorphism. This is discussed in more detail in a subsequent section (Section 3.2).

Figure 2.5; Restriction mapping and identification of transcribed regions of the *Ae.aegypti* rDNA repeat. Fig.2.5a shows the restriction map of the insert DNA from the clones λAar1, λAar3 and λAar7. The extra Hinc II site in λAar9 is also shown (inverted symbol). Regions of the rDNA repeat which show hybridisation signal with total cytoplasmic RNA from *Ae.aegypti* are marked T.C.RNA^{+}. Regions which show hybridisation signal with *Xenopus* 18S and 28S rDNA coding regions are also shown. Fig.2.5b shows the corresping approximate transcription unit of *Ae.aegypti* rDNA derived from this information and information from other sources (see text).

Key to restriction enzyme sites;

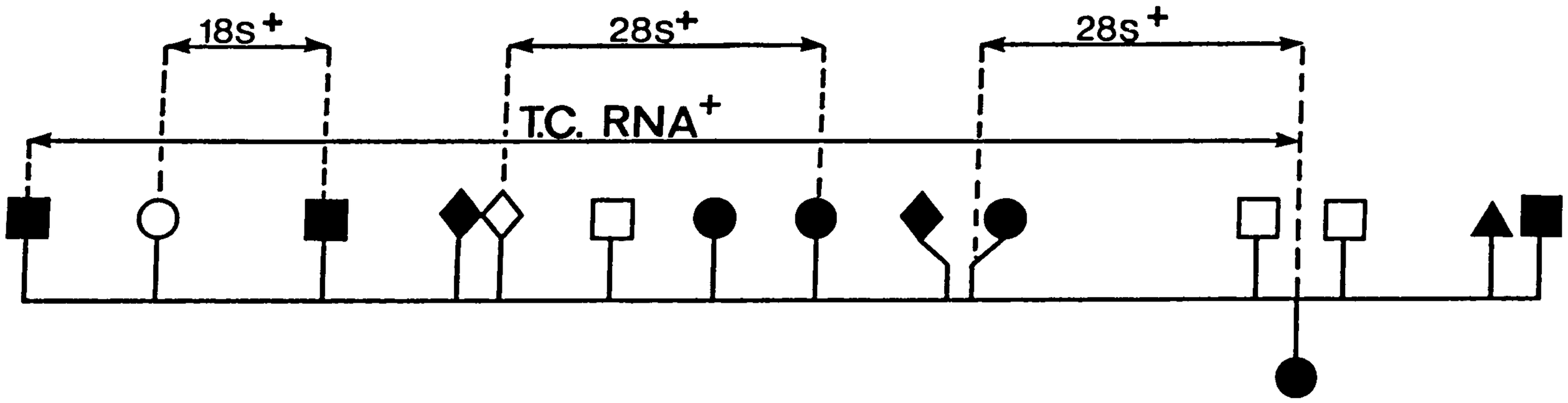


Fig. 2.5

2.6. Identification of transcribed regions within the rDNA repeat

The gels used for the restriction mapping experiments described in Section 2.5 were blotted onto nitrocellulose. These Southern transfers were then hybridised with kinase labelled, base cleaved, total cytoplasmic RNA prepared from *Aedes aegypti* (Bangkok) larvae. (Protocols for extraction and labelling of RNA are described in Materials and Methods). Figure 2.6 shows the results of one of these experiments. The ethidium bromide stained gel shown in Figure 2.6a was used in the restriction mapping of the 7.2 Kb EcoRI fragment from clone λAar3. Figure 2.6b shows the autoradiograph of the corresponding Southern transfer probed with total cytoplasmic RNA. Hybridisation is observed to all restriction fragments except the 1.2 Kb KpnI/EcoRI fragment in lane 5 of the figure. This fragment therefore contains little or no rRNA coding sequence and corresponds to non-transcribed spacer DNA. All other restriction fragments must contain at least some sequence transcribed into cytoplasmic RNA. This non-transcribed region was observed to extend to the extra Hinc II site in λAar 9 (data not shown). Hybridisation with this probe was observed for both restriction fragments which comprise the 1.8 Kb EcoRI fragment of the cloned rDNA repeats (data not shown). The hybridisation properties of the restriction fragments with the total cytoplasmic RNA probe are shown in Figure 2.5.

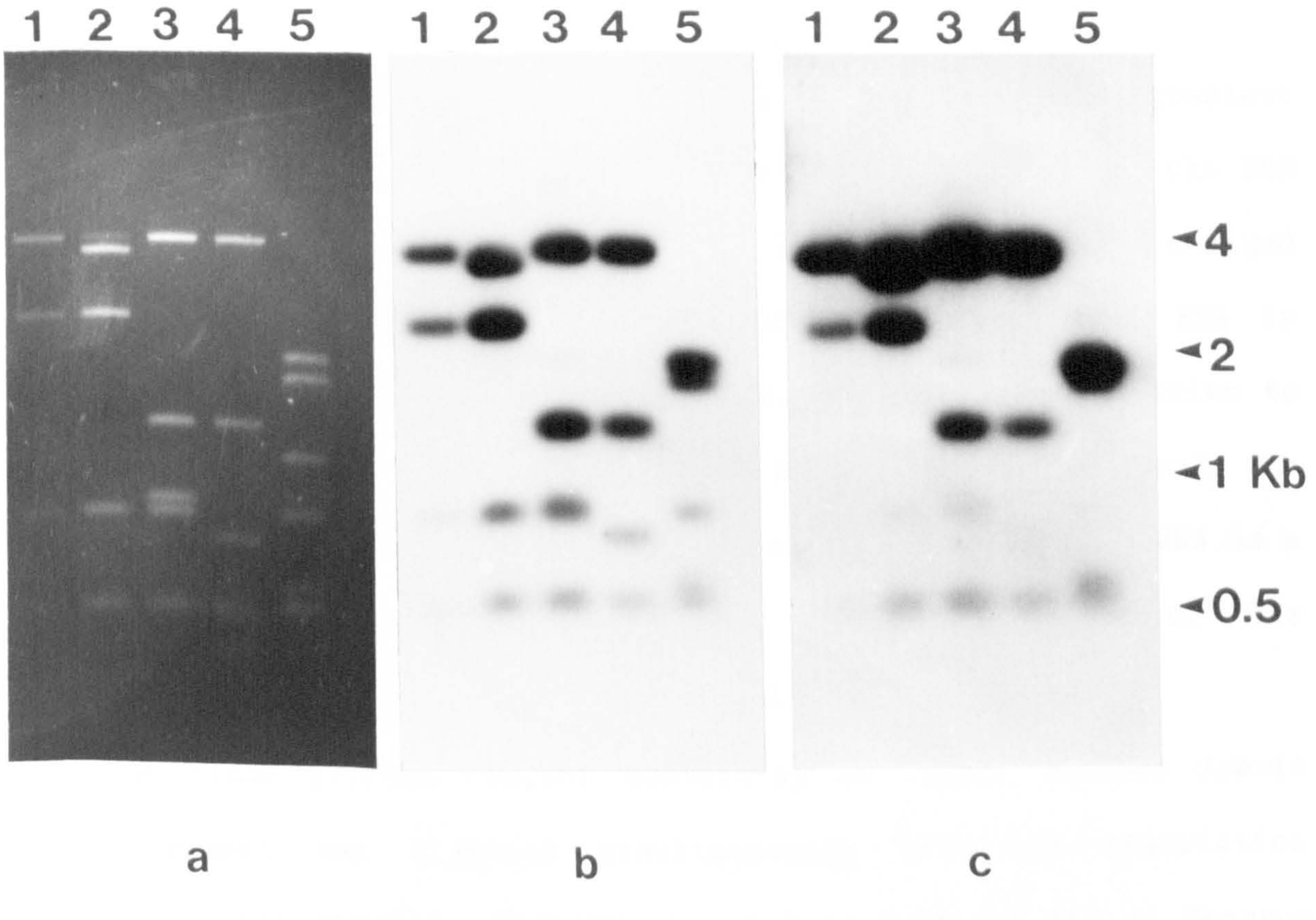


Fig. 2.6

Figure 2.6; Identification of transcribed regions within the *Ae.aegypti* rDNA repeat. Fig. 2.6a shows a 1% minigel used in the restriction mapping of the 7.2Kb EcoR1 fragment from λAar3. Fig. 2.6b shows the result of probing the corresponding Southern transfer with *Ae.aegypti* total cytoplasmic RNA. Fig. 2.6c shows the result of re-probing this Southern transfer using *Xenopus* 28S rDNA. Lane 1; Hinc II, Lane 2; Hinc II + Hind III, lane 3; Hinc II + Sac I, lane 4; Hinc II + Sac I + Pst I, lane 5; KpnI + Hinc II.

2.7. Distinction of 18S and 28S rRNA encoding restriction fragments

Attempts to separate 18S and 28S rRNA by sucrose gradient centrifugation were unsuccessful due to the dissociation of the 28S rRNA into the α and β sub-units [14]. The ethidium bromide stained gel of fractionated and unfractionated total cytoplasmic RNA is shown in Figure 2.7a. It was not possible to allocate identities to the three bands visible in Figure 2.7a (one doublet and one singlet). It was therefore not possible to use 18S or 28S rRNA as a hybridisation probe to identify coding restriction fragments. The alternative strategy employed was as follows.

The clone pXlr101 [95,96] containing an intact *Xenopus laevis* rDNA repeat was digested simultaneously with the restriction endonucleases HindIII and BamHI (see Figure 2.7b for map of *Xenopus laevis* rDNA with respect to these two enzymes). The 3.9 Kb BamHI restriction fragment containing the external transcribed spacer, 18S rRNA gene, internal transcribed spacer and 5.8S rRNA gene, and the 3.0 Kb BamHI fragment which contains only 28S rRNA coding sequences, were purified from 1% low-melting Agarose. These restriction fragments were then nick-translated and used to re-probe the original *Ae. aegypti* rDNA restriction mapping Southern blots. Figure 2.6c illustrates the result of reprobing the λAar3 7.2 Kb rDNA mapping gel with radiolabelled *Xenopus* 28S rDNA. This shows that most restriction fragments on this gel contain some *Xenopus* 28S rRNA homologous sequence. No hybridisation signal was observed for the 7.2 Kb *Ae. aegypti* rDNA fragment when probed with the *Xenopus* transcribed spacer/18S rRNA/5.8S rRNA fragment (result not shown).

However, this probe did show hybridisation to the 1.8 Kb EcoRI rDNA fragment of *Ae. aegypti*. No *Xenopus* 28S rDNA homologous sequences were detected in the 1.8 Kb *Ae. aegypti* ribosomal DNA fragment (results not shown). The hybridisation characteristics of the *Ae. aegypti* rDNA restriction fragments with respect to the two *Xenopus* rDNA probes are shown in Figure 2.5a. Using this data, the results of Shine and Dalgarno [65] and estimates of approximate lengths of the 18S, 5.8S and 28S rRNA genes, which do not vary significantly amongst higher eukaryotes [97], it was possible to construct the approximate transcription map of *Ae. aegypti* rDNA shown in Figure 2.5b. The estimates of rRNA gene size used are: 18S- 2.0 Kb, 5.8S- 0.15 Kb, 28S- 4.0 Kb.

Fig. 2.7A

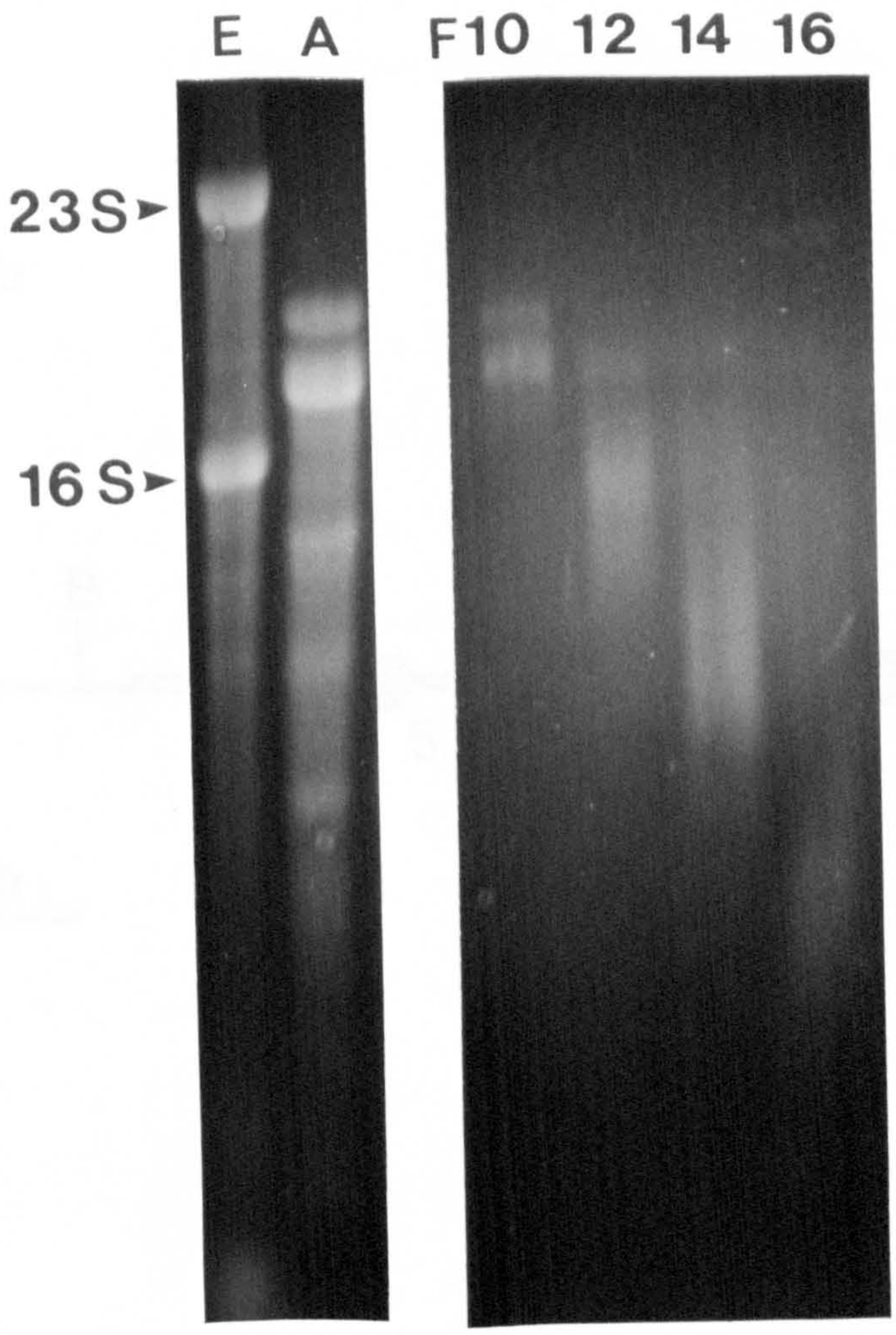


Figure 2.7A; Ethidium bromide stained gel of guanidine hydrochloride extracted total cellular RNA from *Ae.aegypti* (Bangkok strain) larvae and sucrose gradient fractionated RNA from the same source. E= *E.coli* (MC1060) total RNA. This was a gift from J.T.Ellis of this department. A= *Ae.aegypti* total cellular RNA. The identities of the 23S and 16S rRNA bands of *E.coli* are indicated. Fractions are numbered from the base of the gradient.

Fig. 2.7b

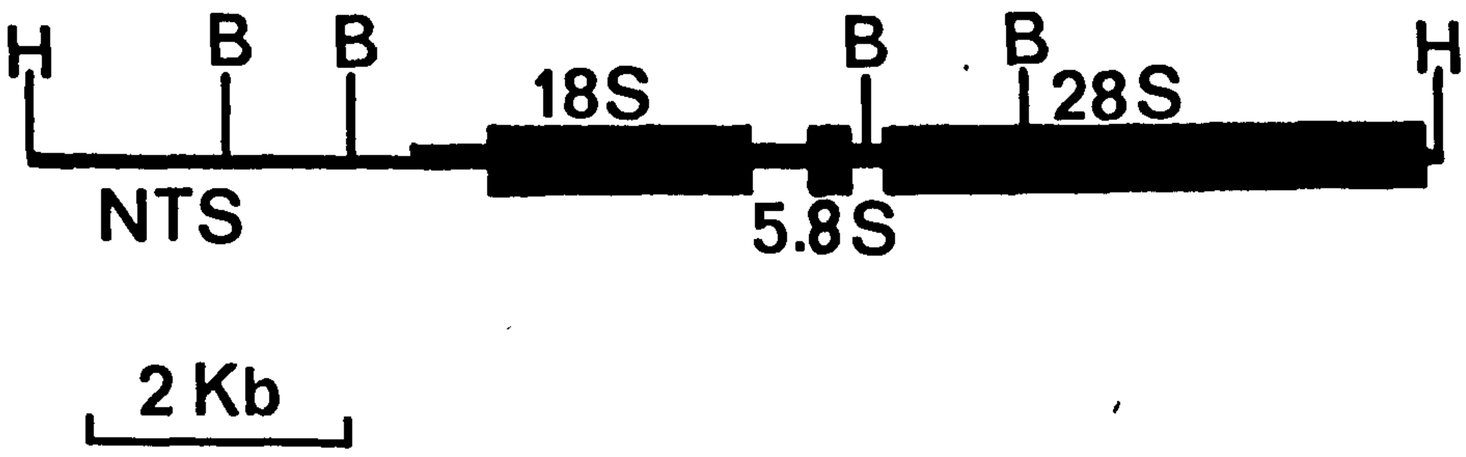


Figure 2.7b; Transcription of the *Xenopus laevis* rDNA contained in the clone pXlr101 showing the positions of the Hind III (H) and BamHI (B) recognition sites.

2.8. Estimate of rDNA copy number in *Aedes aegypti*

Known quantities of EcoRI digested, cloned *Ae. aegypti* rDNA were run in parallel with a known quantity of *Ae. aegypti* (Bangkok) genomic DNA on a 0.8% Agarose gel. The DNA was then transferred to nitrocellulose and the resulting Southern blot probed with nick-translated *Ae. aegypti* rDNA (clone λAar7 total insert DNA). Figure 2.8 shows 'a', the ethidium bromide stained gel and 'b', the autoradiograph of the corresponding Southern transfer.

The relative intensities of the 7.2 Kb rDNA bands obtained for cloned DNA and genomic DNA were determined using a 'Chromoscan 3' densiometric scanner. (The settings used were: aperture width: 0.3 μm, maximum absorbance: 2 O.D. units, valley to valley background correction, maximum background ramp rate: 5, peak to trough thresholds: height- 10, width- 20, noise- 8).

Results

Amount of clone rDNA (ng)	Peak integral
10	2362
20	2972
40	8818
80	17948
2 μg of genomic DNA	2778

The values obtained for cloned DNA were used for linear regression analysis. The amount of rDNA corresponding to the integral obtained for the genomic DNA band was 11.7 ng. Therefore,

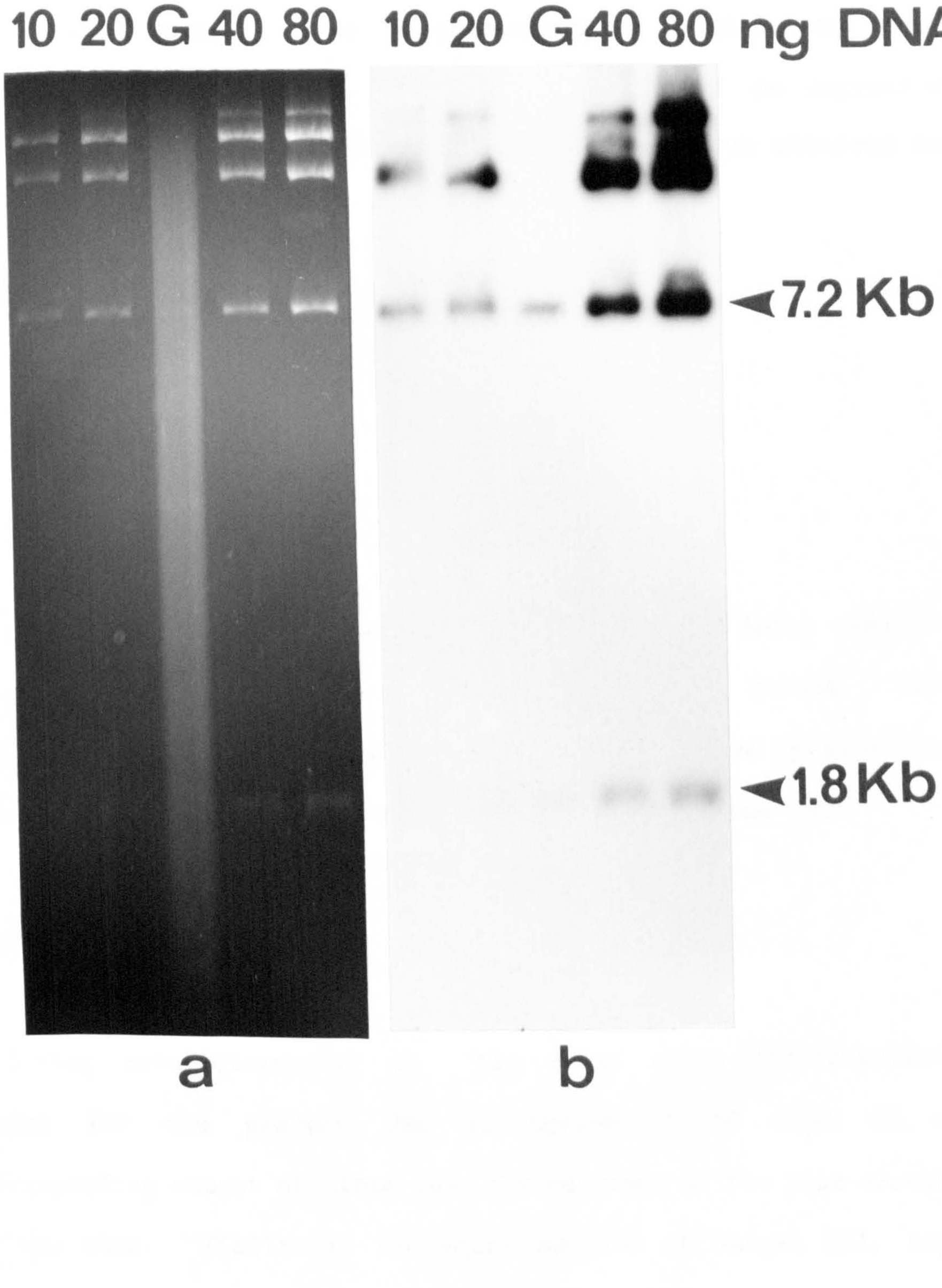


Figure 2.8; Estimate of rDNA copy number in *Ae.aegypti*. Known quantities of λAar7 cloned rDNA and *Ae.aegypti* genomic DNA were digested with EcoRI and size fractionated through 0.8% Agarose. Fig.2.8a shows the ethidium bromide stained gel. Fig.2.8b shows the result of probing the corresponding Southern transfer with λAar7 insert DNA. The values shown for nanograms of DNA correspond to the amount of clone insert DNA on the gel. G= genomic DNA (2μg) from *Ae.aegypti* (Bangkok strain) larvae.

in 2 μg of genomic DNA, 11.5 ng corresponds to the 9.0 Kb *Ae. aegypti* rDNA repeat. This is equivalent to 0.57% of the total genomic DNA. Using a value for genome complexity of *Ae. aegypti* of 8 x 10^8 bp (see Section 1 of this thesis), the value obtained for rDNA copy number

$$= \frac{\text{genome size x rDNA proportion of genome}}{\text{rDNA repeat size x 100}}$$

$$= \frac{8 \times 10^8 \times 0.57}{9 \times 10^3 \times 100} = \underline{507}$$

The estimate of ribosomal DNA copy number in *Aedes aegypti* obtained by this method is 507 copies per haploid genome. The value obtained for rDNA copy number in *Ae. albopictus* by solution hybridisation kinetics is 430 copies per haploid genome [46].

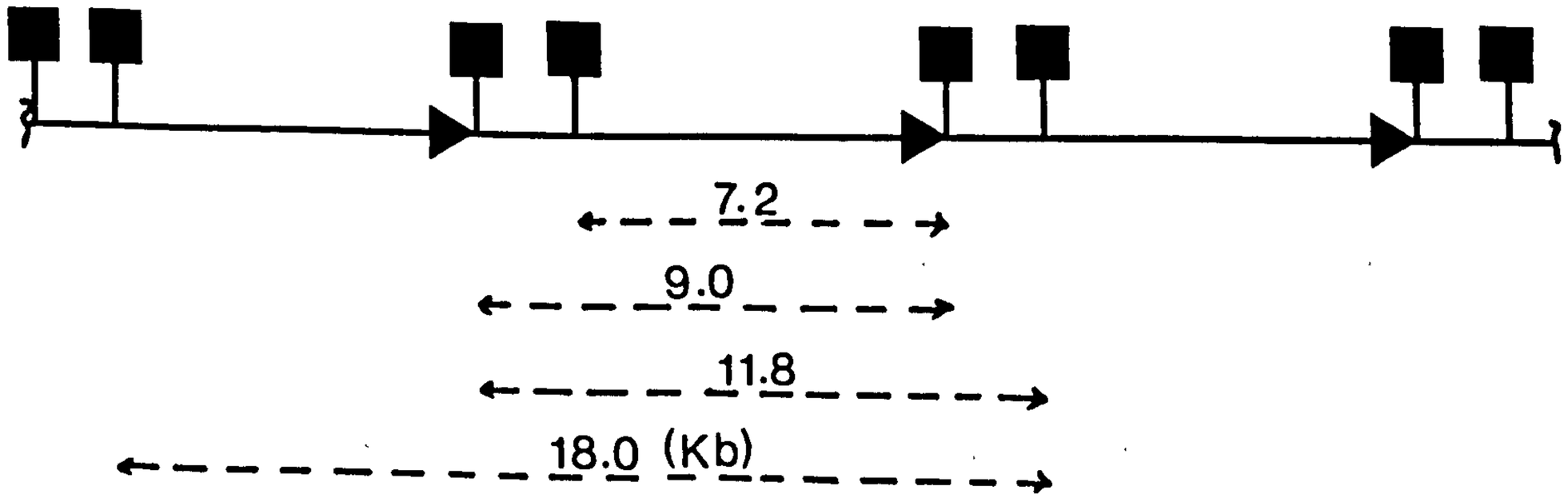
Limits of Accuracy

During autoradiography, at the point where hybridisation signal for the genomic DNA is equivalent to that of a corresponding amount of clone DNA, the response of the film should be the same. Similarly, for equal amounts of target DNA, any effects of probe depletion or target DNA saturation should be identical. The filter-bound target DNAs are identical for cloned and genomic DNA sequences. Some differences in transfer efficiency during blotting and hybridisation efficiency may occur due to the other DNA fragments present in the genomic digest. The precision of measurements (e.g. DNA concentrations, molecular weights and

integral readings) is probably high relative to the possible errors in accuracy caused by system errors.

2.9. Ae. aegypti rDNA is tandemly repeated in a head-to-tail manner

The preliminary studies by Southern blotting described in Section 2.1. and 2.2. of this thesis indicate that the genes encoding rRNA of *Ae. aegypti* are tandemly repeated as in all other higher eukaryotes studied. The simplicity of the rDNA fragment pattern on Southern transfers also suggested the repeats were in a simple head-to-tail configuration. To confirm this, the Southern transfer of the gel shown in Figure 1.1 (a titration series of EcoRI against *Ae. aegypti* genomic DNA) was probed using *Ae. aegypti* rDNA. The partial digestion bands which were observed in addition to the 7.2 Kb and 1.8 Kb limit digestion bands were 9.0, 11.8 and 18 Kb in length (results not shown). These bands correspond to those expected for a head-to-tail tandem repeat of the *Ae. aegypti* rDNA unit. (See below).



2.10. Section 2: Summary

Preliminary experiments using a cloned *Drosophila* rDNA repeat to probe Southern transfers of *Ae. aegypti* genomic DNA indicate that the rDNA repeat of *Ae aegypti* is 9.0 Kb in length and that individual rDNA repeats exhibit a high degree of homogeneity with respect to length and the position of restriction enzyme recognition sites within the rDNA. The preliminary mapping data (Section 2.3) together with the partial digestion experiment (Section 2.8) demonstrates that as in all other higher eukaryotes, the rDNA repeats are arranged in a head-to-tail tandemly repeating manner.

The restriction mapping of cloned rDNA repeats confirms the largely uniform length of the *Ae. aegypti* rDNA repeat and provides a more detailed physical map of the DNA. (shown in Figure 2.5 a). The complexity of the HincII digestion observed in the preliminary experiment of Section 2.2. is explained in part by a restriction site polymorphism in clone λAar9 which contains an extra HincII site, not present in λAar1, λAar3 or λAar7.

Transcription mapping as described in Sections 2.6 and 2.7 allows the allocation of identities to the various restriction fragments and the approximate positioning of the transcription unit as shown in Figure 2.5 b. Although the short length of the *Ae. aegypti* rDNA repeat precludes the presence of large insertions in all the repeats, the presence of short introns in all of the rDNA units cannot be excluded from the results presented here.

The estimate of rDNA repeat copy number obtained in Section 2.8 (approximately 500 copies per haploid genome) is similar to the

estimate reported for the closely related species *Aedes albopictus* of 430 copies per haploid genome [46]. Ribosomal DNA comprises approximately 0.6% of the total *Ae. aegypti* genome.

Section 3: Analysis of rDNA Variation Within *Aedes aegypti* and Between Closely Related Species

3.1. Variation between different species

High molecular weight genomic DNA from *Aedes aegypti*, six other mosquito species (see below) and an *Ae. aegypti* cell-line was digested with EcoRI and size fractionated through 0.8% Agarose. The DNA was then transferred to nitrocellulose and hybridised with the rDNA repeat from *Ae. aegypti* (total insert DNA from clone λAar3). The result is shown in Figure 3.1.

Classification of the mosquitoes used [3,4] is given below.

Order : Diptera, family : Culicidae

SPECIES	SUB-FAMILY	GENUS	SUB-GENUS
Ae. aegypti	Culicinae	Aedini	Stegomyia
Ae. albopictus	"	"	"
Ae. scutellaris	"	"	"
Ae. horrescens	"	"	"
Ae. polynesiensis	"	"	"
Ae. togoi	"	"	finlaya
Culex pipiens quinquefasciatus	"	Culex	culex
Anopheles gambiae	Anophelinae	Anopheles	cellia

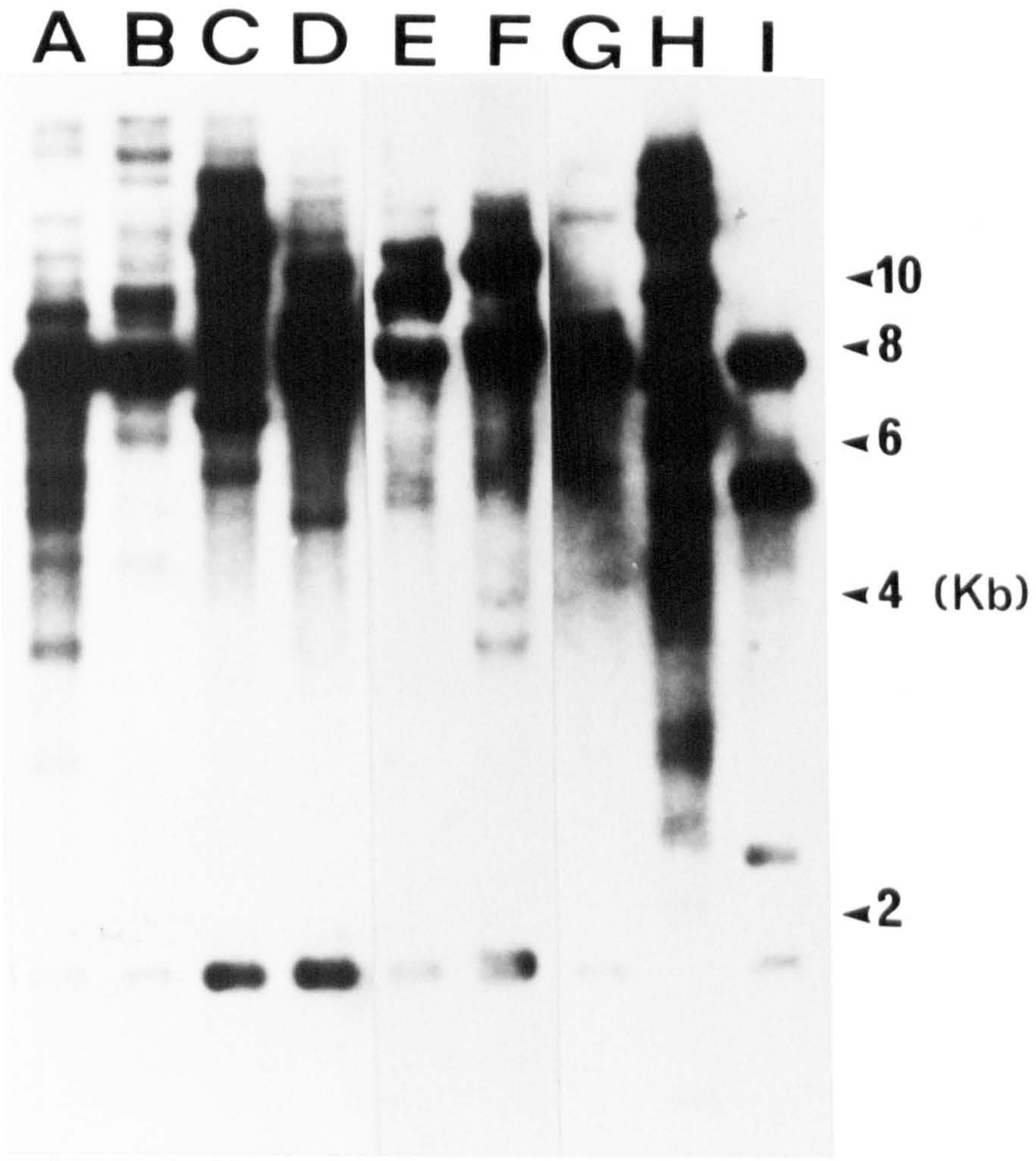


Fig. 3.1

Figure 3.1; Ribosomal DNA variation between different species of mosquito. Genomic DNA (4μg) from each species was restricted with EcoR1, size fractionated through 0.8% Agarose and transferred to nitrocellulose. This Southern transfer was then probed using λAar3 total insert DNA. A; *Ae.aegypti* (Bangkok strain), B; *Ae.aegypti* (Mos20 cell-line), C; *Ae.albopictus*, D; *Ae.scutellaris*, E; *Ae.polynesiensis*, F; *Ae.horrescens*, G; *Ae.togoi*, H; *An.gambiae* (Warwick cell-line), I; *C.pipiens quinquefasciatus* (Apo strain).

The *Ae. aegypti* cell line used was Mos20A, originally derived by Varma and Pudney [47]. The *An. gambiae* DNA was kindly supplied by J.D. Norton from the Department of Biological Sciences, Warwick.

The species *Ae. scutellaris*, *Ae. horrescens* and *Ae. polynesiensis* are members of the morphologically very similar *Aedes scutellaris* group.

Figure 3.1. shows that when *Aedes aegypti* genomic DNA is hybridised with an *Ae. aegypti* rDNA probe of high specific activity (greater than 10^8 dpm/μg) it is possible to distinguish minor bands of hybridisation signal in addition to the major bands visible in Figure 2.3. Interestingly, when the two *Ae. aegypti* DNAs are compared, although the mijor bands of hybridisation are identical, the array of minor bands is very different between the larval and cell-line DNAs. The positions and relative intensities of these minor bands are highly reproducable when the DNAs are digested under different conditions (e.g. using a large excess of enzyme and long incubation time) and are therefore not artifacts of partial digestion. The significance of this intra-specific variation in rDNA and its possible sources is discussed in subsequent sections.

The different species shown in Figure 3.1. show a large degree of variation in their rDNA 'fingerprint'. Most exhibit a more complex array of rDNA bands than *Ae. aegypti* with more than two major rDNA bands. Although it is very difficult to quantitate variation in rDNA hybridisation patterns, some correlation is evident between the degree of variation observed and the evolutionary relatedness of the species as judged by their classification. The 1.8 Kb EcoRI fragment (which contains 18S rRNA coding sequences in *Ae. aegypti*) is visible in all of the *Aedes* species shown and the *Culex* species (subfamily

culicinae), but is absent from the *Anopheles* digest (subfamily anophelinae).

In *Ae. albopictus*, the major rDNA bands are 12.5 Kb and 1.8 Kb in length. The 12.5 Kb band is probably longer in this species due to a longer non-transcribed spacer region. The three *scutellaris* group species (*Ae. scutellaris*, *Ae. polynesiensis* and *Ae. horrescens*) exhibit two or three major rDNA bands in addition to the common 1.8 Kb EcoRI fragment. These species show more variation between species in their rDNA fingerprints than in their external morphological characteristics. The rDNA pattern of *Ae. togoi* is very similar to that of *Ae. aegypti*, although *Ae. togoi* is classified in a different sub-genus and is morphologically quite distinct. Both species show bands at 7.2 Kb and 1.8 Kb with minor additional bands of hybridisation. The *Culex* rDNA fingerprint also shows strong bands at 7.2 Kb and 1.8 Kb and has two additional bands of 5.2 Kb and 2.2 Kb.

The *Ae. aegypti* rDNA probe used in this experiment shows a more complex rDNA fingerprint in *Ae. aegypti* than the *Drosophila* rDNA probe used previously (Section 2.3). This may be explained by the homologous nature of this probe (forming more homologous duplexes with the filter-bound rDNA). Also, the *Ae. aegypti* rDNA probe should detect both the highly conserved rDNA coding regions and the less conserved spacer sequences. Heterologous rDNA probes are unlikely to detect spacer sequences, even in quite closely related species. The rDNA fingerprint of *Culex* shown in Figure 3.1 for example, is probably incomplete in comparison with the pattern which would be obtained using an homologous (*Culex*) rDNA probe. It should be noted that the long non-transcribed spacer regions of mammalian rDNA often contain highly repeated interspersed sequences [99] which would give

highly misleading results if the whole rDNA repeat was used as a probe against genomic DNA.

Ribosomal DNA Variation within *Aedes aegypti*

3.2. The HincII restriction polymorphism

The extra HincII recognition site in the non-transcribed spacer region of λAar9 is indicative of sequence polymorphism within this region of the *Ae. aegypti* rDNA repeat. This polymorphism is evident on the Southern transfer of HincII digested *Ae. aegypti* genomic DNA probed with the mosquito rDNA repeat. (Figure 3.2a). The sizes of the HincII rDNA bands predicted from the restriction map of clones λAar 1, 3 and 7 are: 5.0 Kb, 2.6 Kb, 0.56 Kb and 0.88 Kb (see Figure 3.2b). With the exception of the low molecular weight bands, these bands are visible on the autoradiograph shown. The extra HincII site present in clone λAar9 generates 3.0 Kb and 2.0 Kb fragments from the 5.0 Kb fragment observed in the other three clones. Bands corresponding to these molecular weights are also visible on the autoradiograph shown in Figure 3.2a. The molar ratio of the 5.0 Kb and the 3.0 Kb rDNA bands as estimated by densiometric scanning of the autoradiograph and correction for the difference in molecular weight is 3 : 1. This is also the relative abundance of the two cloned rDNA types examined. This ratio is an indication of the approximate frequencies of the two rDNA types within the *Ae. aegypti* genome.

In addition to the bands predicted from the restriction maps of the cloned ribosomal repeats, an extra band (2.3 Kb) is also visible on the autoradiograph shown in Figure 3.2a. One other quite frequent polymorphism must therefore occur with respect to the HincII recognition sites within the rDNA repeat. The three bands of molecular weights 2.6 Kb, 2.3 Kb and 2.0 Kb could not be resolved using densiometric scanning techniques.

Figure 3.2; Polymorphism of Hinc II recognition sites within *Ae.aegypti* ribosomal DNA. Fig. 3.2a; Hinc II restricted *Ae.aegypti* genomic DNA probed using *Ae.aegypti* rDNA (insert from clone λAar7). Fig. 3.2b; Map of *Ae.aegypti* rDNA showing the position of Hinc II recognition sites. The extra Hinc II recognition site in clone λAar9 is marked '. The 0.56Kb Hinc II fragment was not visible on the autoradiograph. The 0.88Kb Hinc II fragment was visualised, but at insufficient intensity to reproduce on the exposure shown.

Fig. 3.2a

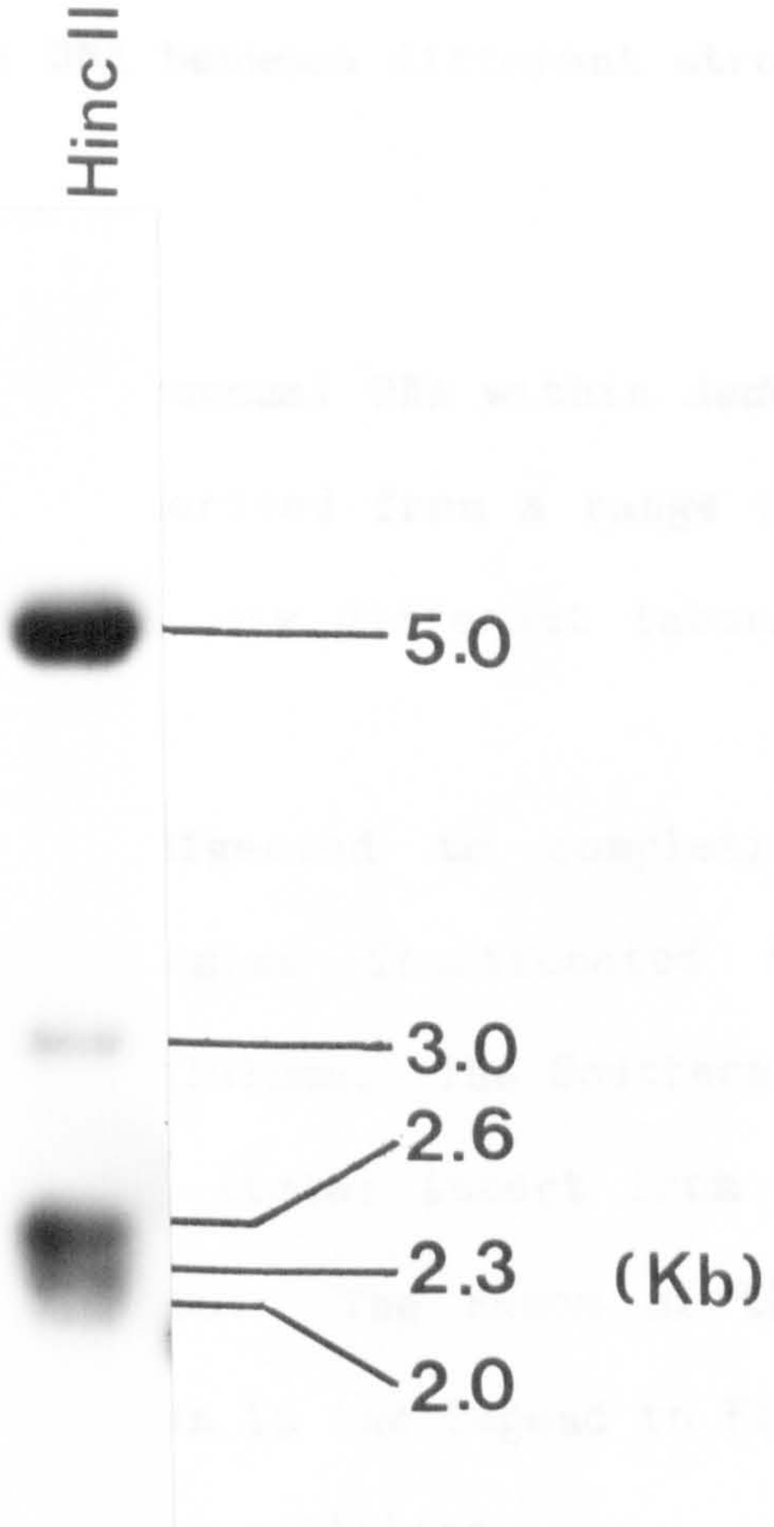
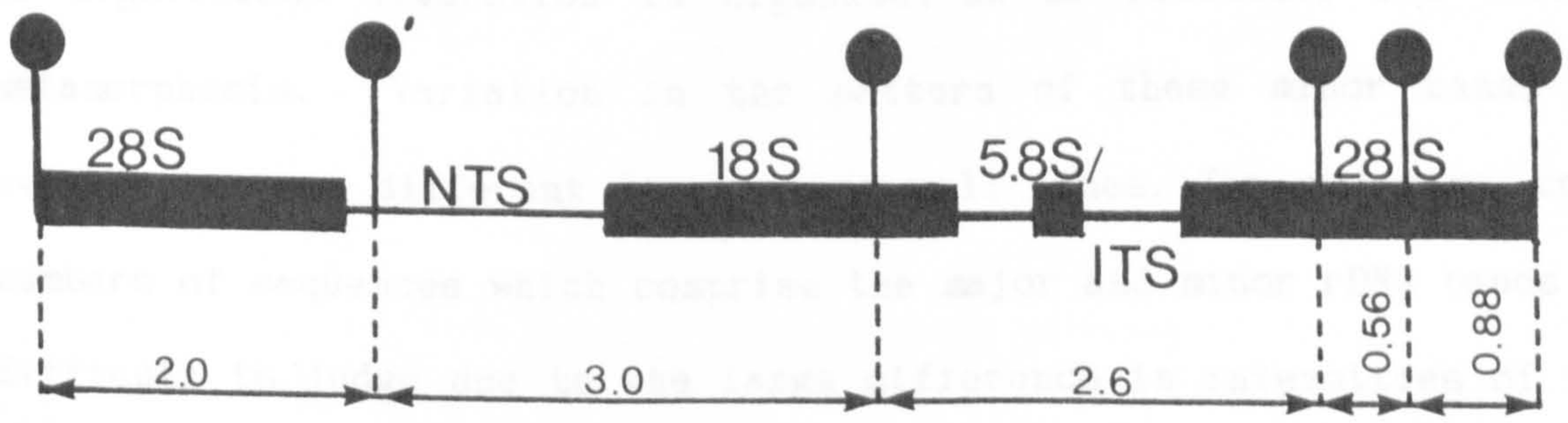


Fig. 3.2b



3.3. Variation of ribosomal DNA between different strains of *Aedes aegypti*

To investigate variation of ribosomal DNA within *Aedes aegypti*, a comparison was made between DNAs derived from a range of life cycle stages (larvae, pupae and adults), six different laboratory stocks, and two different cell-lines.

DNA from each source was digested to completion with the restriction endonuclease EcoRI, size fractionated through 0.8% Agarose and transferred to nitrocellulose. The Southern transfer was then probed with *Ae. aegypti* rDNA (total insert from clone λAar3). The result is shown in Figure 3.3a. The names of the laboratory stocks and cell-lines used are given in the legend to Figure 3.3a and indicate the source of the colony or cell-line.

The results show that the major 7.2 Kb and 1.8 Kb rDNA bands (when visible) are common to all strains used. No variation is evident between the minor rDNA bands of genomic DNA from larvae, pupae and adults of the Bangkok strain. This indicates that there is no significant alteration in organisation of ribosomal DNA during metamorphosis. Variation in the pattern of these minor bands is evident between different strains and cell-lines. The relative copy numbers of sequences which comprise the major and minor rDNA bands is difficult to judge due to the large difference in intensities of the bands (e.g. 3 O.D. units and 0.1 O.D. units respectively) and the non-linear response of the film over such a range. For the Bangkok strain pupae DNA (which lacks smearing of the main rDNA band due to non-specific degradation), over 95% of the total signal on the autoradiograph is present in the 7.2 Kb rDNA band. The copy number

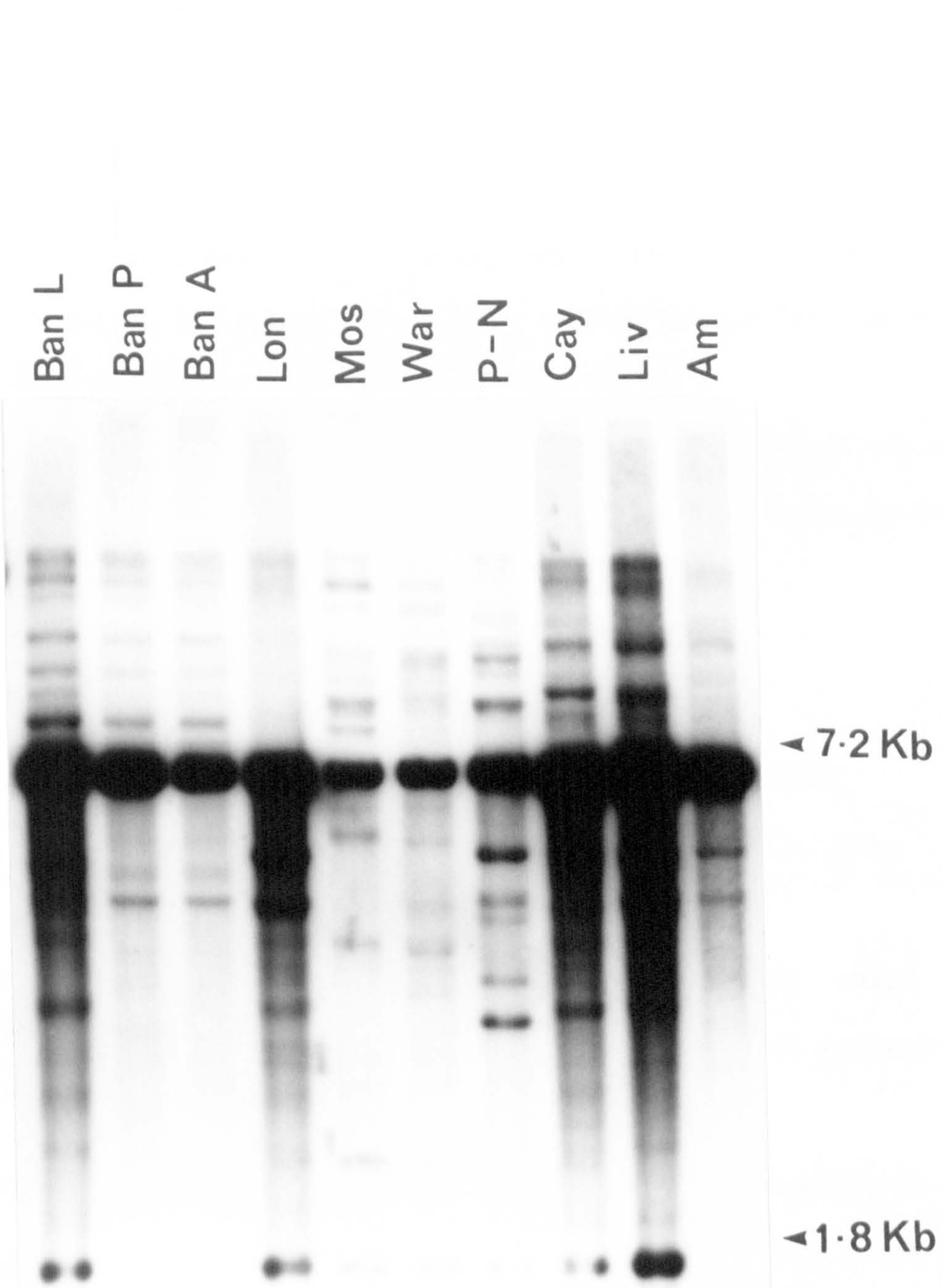


Fig. 3.3a

Figure 3.3a; Variation in rDNA fingerprint between different strains of *Ae.aegypti*. 4µg of DNA from each strain was digested to completion with EcoR1, size fractionated through 0.8% Agarose and transferred to nitrocellulose. Shown is the autoradiograph of this Southern transfer probed with *Ae.aegypti* rDNA (total insert DNA from clone λAar3). Lanes 1-3; Bangkok strain larval, pupal and adult DNA, lane 4; London strain larval DNA, lane 5; Mos20A cell-line DNA, lane 6; Warwick cell-line DNA, lane 7; Porto-Novo strain larval DNA, lane 8; Cayenne strain larval DNA, lane 9; Liverpool strain larval DNA, lane 10; American ss strain larval DNA.

of the minor rDNA bands is obviously much lower than that of the major rDNA species which are present in approximately 500 copies per haploid genome (Section 2.8).

To indicate the molecular weight range and variation of minor bands observed in Figure 3.3a, the sizes of the minor bands on two of the digests (Bangkok and Mos20A) are listed below.

STRAIN	SIZES OF MINOR RIBOSOMAL DNA BANDS (Kb)
Bangkok	55, 29, 20, 11.2, 10.0, 9.0, 5.6, 5.0, 4.4, 3.6.
Mos 20A	60, 50, 26, 21, 12.8, 10.8, 9.7, 9.5, 8.6, 6.0, 4.4.

The possible sources of the minor rDNA bands which show variability between different *Ae. aegypti* strains are as follows:

1. Restriction polymorphisms with respect to EcoRI recognition sites within a small proportion of the rDNA repeats. Extra sites would generate low molecular weight minor bands whereas destroyed EcoRI sites would generate larger rDNA fragments of 7.2 + 1.8 Kb multiples.
2. Variation in the length of the rDNA repeat. This could arise from duplication or deletion of any part of the repeat (e.g. polymorphism in N.T.S. length) or from the insertion of non-rDNA (e.g. a transposable genetic element) into a repeat unit.
3. The presence of sequences homologous to some part of the rDNA repeat outside the tandemly repeated rDNA. Pseudogenes inserted at diverse loci or interspersed repeats present within the rDNA repeat (see [49] for example) are two possiblesources of this.
4. Polymorphism in the length of the EcoRI fragments produced at the ends of tandemly repeated rDNA. Each end may generate a restriction

fragment containing some ribosomal sequence joined to non-ribosomal DNA.

5. Experimental artifacts. As mentioned previously, the molecular weights and relative intensities of the minor rDNA bands are highly reproducable for a number of different Southern transfers, using different digestion conditions. (See lanes 1-3 of Figure 3.3 for example). These are therefore unlikely to be due to partial digestion or EcoRI-star activity. The smearing of the major 7.2 Kb rDNA band may however give some misleading results. Non-specific degradation of DNA during the extraction process may act on specific 'hypersensitive sites' within the DNA to produce discrete break-down products [100]. The two bands of 6.2 Kb and 5.0 Kb clearly visible in lanes 4, 9 and 10 of Figure 3.3 may be the result of this phenomenon.

6, Methylation of mosquito DNA. This could affect the susceptibility of sites within the rDNA to cleavage by restriction endonucleases. Two studies on the methylation of insect DNA [101] and mosquito DNA in particular [102] indicate that mosquito DNA is not methylated.

Ribosomal DNA fingerprint of individual mosquitoes

DNA was prepared from ten individual *Ae. aegypti* Bangkok strain adults (5 males, 5 females) as detailed in Materials and Methods. The DNAs were digested with EcoRI, size fractionated through 0.8% Agarose and transferred to nitrocellulose. This Southern transfer was then hybridised with nick-translated total insert DNA from clone λAar7 which contains an intact *Ae. aegypti* rDNA repeat. The ethidium bromide stained gel and corresponding autoradiograph are shown in Figure 3.3b. This shows that approximately 1 μg of high molecular weight genomic DNA, which shows little non-specific degradation was obtained from each individual. Slightly less DNA was obtained from the males due to their smaller size. The Southern transfer shows the expected 7.2 Kb and 1.8 Kb main rDNA bands, but unfortunately a large proportion of partial digestion products (9.0 Kb, 10.8 Kb, 17.0 Kb, 18.8 Kb and 26 Kb) is also evident. This is despite the use of an eight-fold excess of enzyme, overnight incubation, and careful washing and precipitation of DNAs prior to digestion.

With the exception of the degree of partial digestion, all individuals have the same major rDNA bands. Some minor rDNA bands are also just visible on some of the more completely digested samples. Although the quality of the DNA used here is poor compared to that obtained by the Caesium chloride gradient method (see Materials and Methods) and the resolution of the Southern transfer is correspondingly poor, these results are evidence that variant rDNA repeats occur as a fraction of the total rDNA repeats of an individual rather than as all the rDNA repeats of a fraction of the individuals. A better method for extracting DNA from individuals (or

of digesting the DNA obtained by this method) and analysis of a larger number of individuals using other enzymes (e.g. HincII) in addition to EcoRI is required to verify this hypothesis.

3.4. Isolation of clones containing variant rDNA from the *Ae. aegypti* genomic library

To determine the source of the minor rDNA bands discussed in Section 3.3, it was decided to search for clones in the *Aedes aegypti* (Bangkok) genomic library (see Section 1) which might correspond to rDNA variants. In addition to ascertaining the extent to which the possible causes discussed in Section 3.3 contribute to rDNA variation, it was also noted that screening for rDNA variants may represent a potentially useful method of isolating transposable genetic elements. No mosquito transposable genetic elements have been reported to date, although such sequences are of considerable interest to workers in the field of mosquito molecular genetics.

Two approaches were taken to isolate rDNA variants. Firstly, a larger number of *Ae. aegypti* rDNA clones (isolated by screening the genomic library with the *Drosphila* rDNA probe) were analysed with respect to insert size. Variants were picked simply by the criterion of size (deviation from the 7.2 + 1.8 Kb insert sizes). Secondly, the genomic library was screened to selectively isolate only variant ribosomal clones. This method was devised to select clones which contained only a part of the discrete 7.2 Kb EcoRI rDNA fragment which is present in all non-variant rDNA clones.

Figure 3.3b; Ribosomal DNA fingerprint of individual mosquitoes. DNA from five male and five female *Ae.aegypti* (Bangkok strain) adults was digested with EcoRI, size fractionated through 0.8% Agarose and transferred to nitrocellulose. The ethidium bromide stained gel and autoradiograph of the corresponding Southern transfer probed using *Ae.aegypti* rDNA (insert from clone λAar7) are shown. Lanes 1-5; males, lanes 6-10; females. The restriction enzyme was omitted from the digestion shown in lane 10.

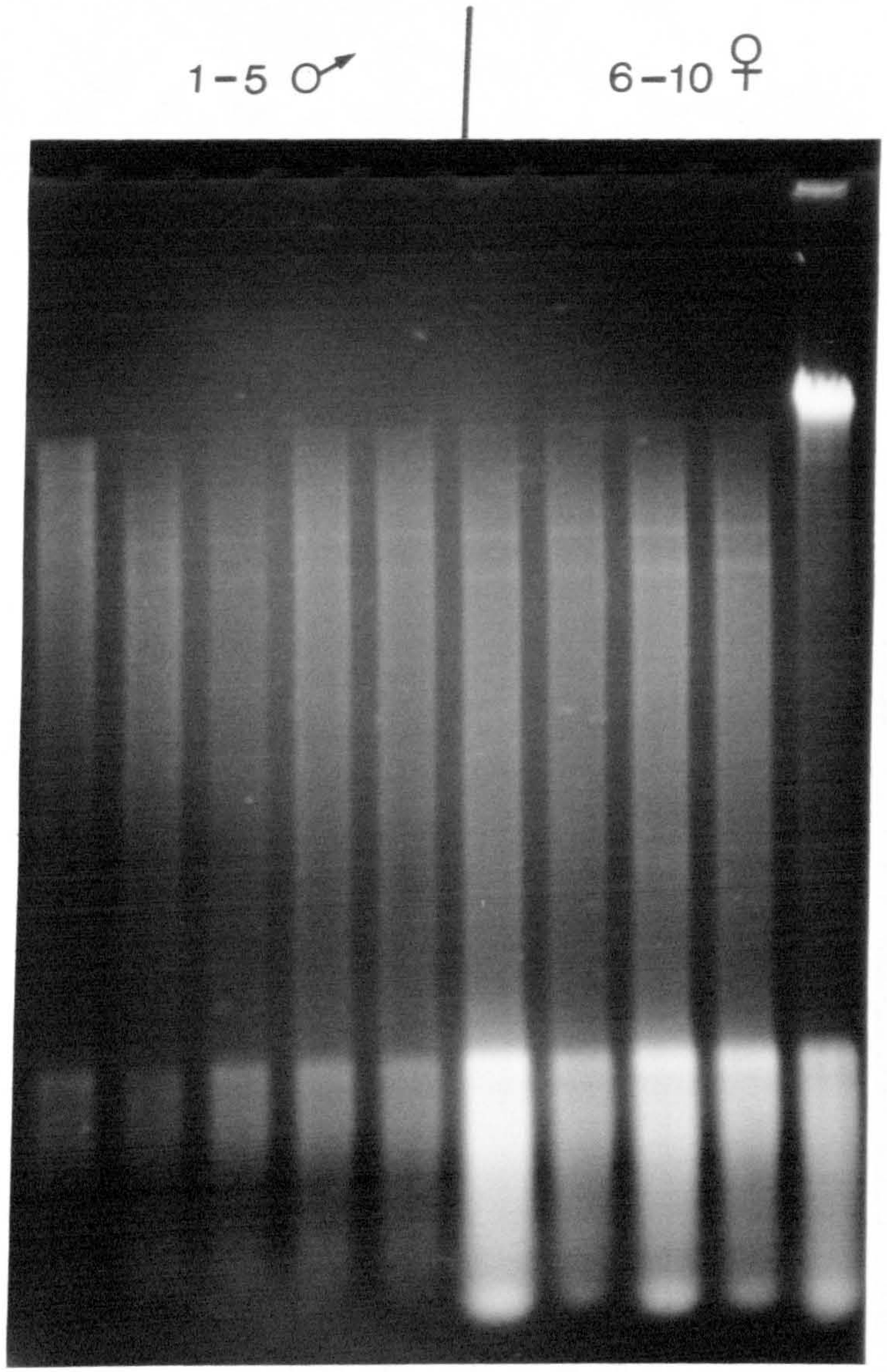
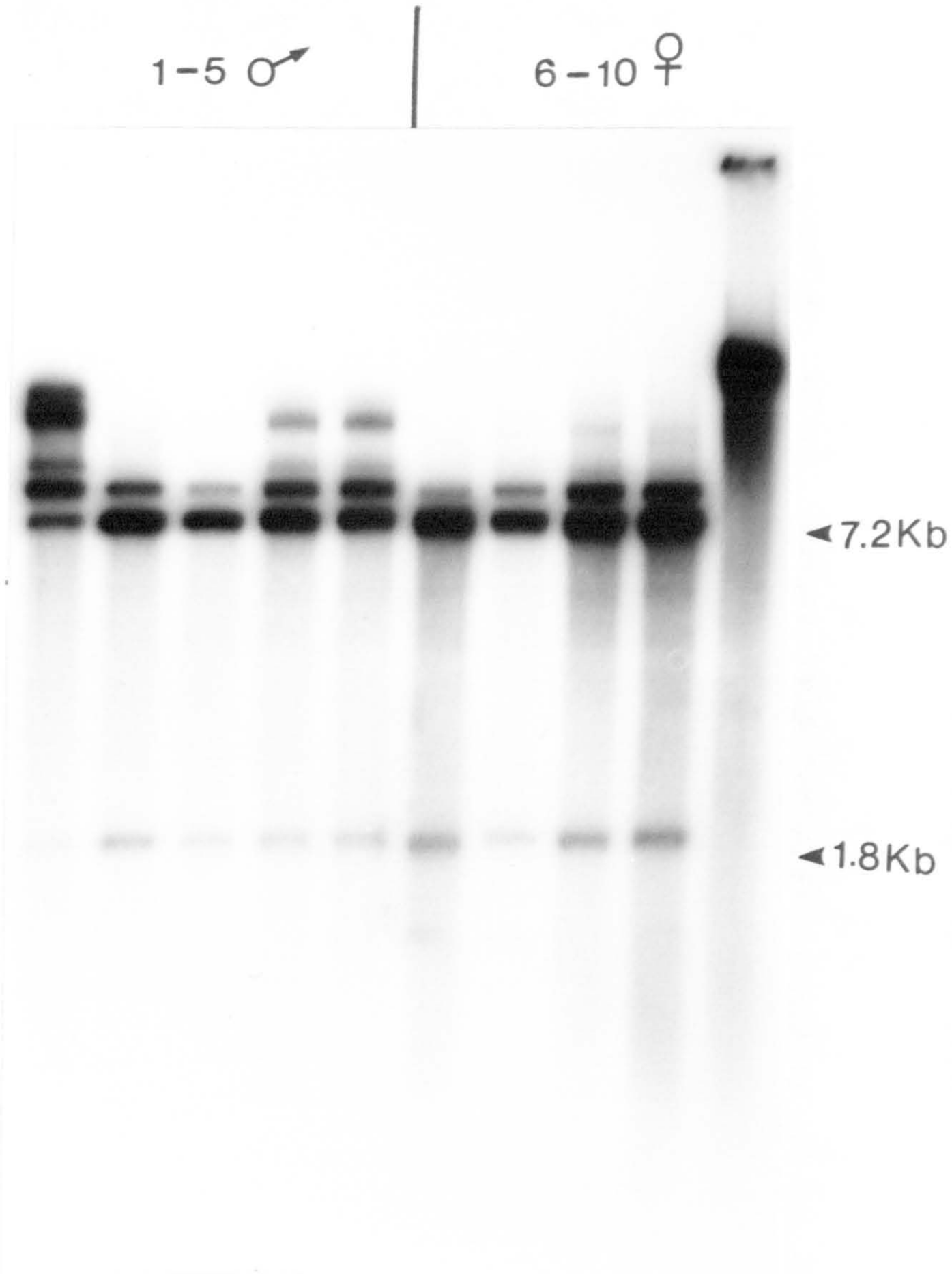


Fig.3.3b



A; Analysis of 30 additional, randomly picked rDNA clones

Thirty clones which showed hybridisation signal with the *Drosophila* rDNA probe were picked. Phage minipreps (see Materials and Methods) were performed and the DNA digested with EcoRI to determine insert sizes. The gel used for insert size determination was blotted and probed with *Drosophila* rDNA to confirm the clones as ribosomal. Sixteen clones were identified which contained 7.2 Kb and 1.8 Kb rDNA fragments (the fragments expected from the major rDNA class), seven were found to contain only the 7.2 Kb EcoRI fragment and two contained insert fragments of approximately 9.0 Kb. These two clones containing inserts of variant length EcoRI fragments of rDNA were designated λAar29 and λAar35.

B; Screening of the *Ae. aegypti* genomic library for specific rDNA variants

Approximately 8000 clones from the *Ae. aegypti* genomic library (amplified stock) were plated onto a single BRL bioassay tray. Three nitrocellulose replicas were taken of the plate (see Materials and Methods). A replica filter was then hybridised with one of the following hybridisation probes:

a. pDm238 clone DNA (*Drosophila* rDNA)

b. 1.2 Kb EcoRI/KpnI restriction fragment DNA from the clone λAar4 (which contains the 7.2 Kb *Ae. aegypti* rDNA fragment only), corresponding to non-transcribed spacer sequences.

c, 1.7 Kb EcoRI/KpnI restriction fragment DNA from λAar4 which contains internal transcribed spacer sequences.

Probes 'b' and 'c' were isolated by digesting λAar4 total clone DNA with EcoRI and KpnI followed by extraction of the appropriate restriction fragments from 1% low-melting Agarose as described in Materials and Methods. These restriction fragments correspond to the right and left ends respectively of the 7.2 Kb EcoRI fragment of *Ae. aegypti* rDNA (see Figure 2.5a). Within the *Ae. aegypti* genomic library (which contains EcoRI- fragmented *Ae. aegypti* DNA) the two ends of this restriction fragment would normally be cloned in unison. Separation of the two ends of the 7.2 Kb restriction fragment within the genomic library requires the presence of an extra EcoRI site within the rDNA. This situation could arise either via a sequence polymorphism or the insertion into the rDNA of a sequence containing an EcoRI recognition site. The pDm238 probe was used to distinguish clones containing 18S/28S coding sequences.

Results

The *Drosophila* rDNA probe gave a total of 254 positive clones. Of these, one clone was positive with probe 'c' (left-hand end of the 7.2 Kb EcoRI rDNA fragment), but not probe 'b' (the right-hand end of the 7.2 Kb EcoRI rDNA fragment). This clone, designated λAav1 was picked for further analysis. One other clone showed hybridisation signal with probe 'b' (the *Ae. aegypti* non-transcribed spacer region) but not with probes 'a' and 'c' (*Drosophila* rDNA and the right-hand

end of the 7.2 Kb EcoRI rDNA fragment). This clone, designated λAas1 (spacer-homologous) was also picked.

Summary of isolated rDNA variants

NAME	INSERT SIZE (Kb)	VARIANT CHARACTERISTICS
λAar29	8.5	rDNA size variant
λAar35	8.3	rDNA size variant
λAav1	8.6	rDNA positive, N.T.S. negative
λAaS1	5.1 + 3.5	rDNA negative, N.T.S. positive

3.5. Restriction mapping of *Ae. aegypti* rDNA variant clones

The rDNA variant clones isolated in Section 3.4 were re-plated to give single plaques. A single plaque from each plate was then used to inoculate 500 ml cultures and DNA was prepared as described in materials and methods. Insert DNA was isolated from each of the variant rDNA clones by digestion of whole clone DNA with EcoRI followed by extraction of the insert DNA band(s) from low-melting Agarose.

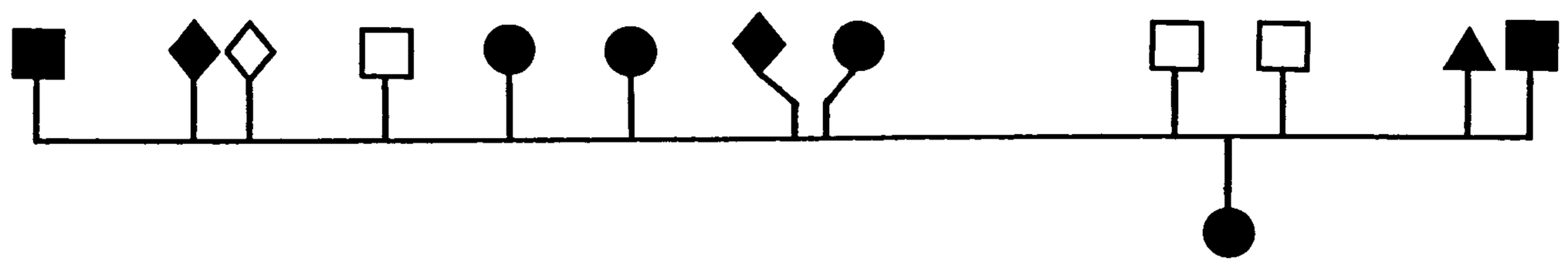
The insert DNA from the clones λAar29, λAar35 and λAav1 was restriction mapped using single and multiple restriction enzyme digestions. The restriction maps were found to show considerable similarity with the normal rDNA repeat and were checked by running appropriately digested DNA in parallel with previously mapped ribosomal clone DNA. The derived restriction maps of these rDNA

variants are shown in Figure 3.5 together with the corresponding region of the normal *Ae. aegypti* rDNA repeat.

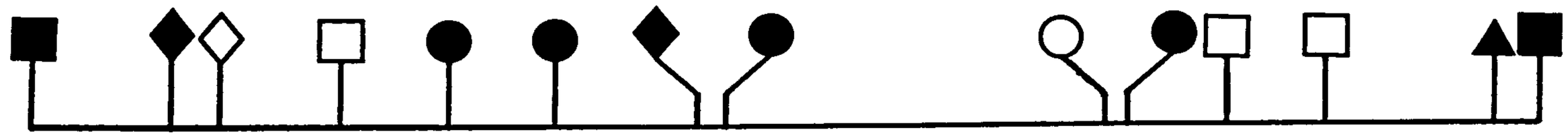
Figure 3.5 (overleaf); Restriction maps of the 7.2Kb EcoRI fragment of *Ae.aegypti* rDNA (present in clones λAar1,3,7 and 9) and of three variant rDNA clones: λAar29, λAar35 and λAav1. The extra Hinc II site present in clone λAar9 is shown as an inverted symbol.

Key to restriction enzyme sites;

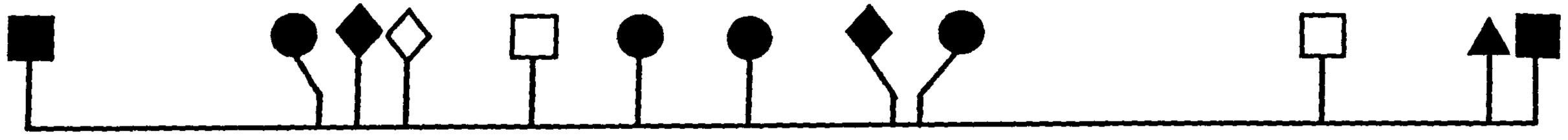
ʎAar 1,3,7,9



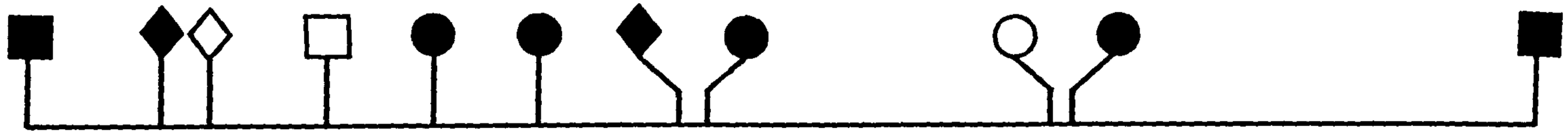
ʎAar 29



λAar35



λAav1



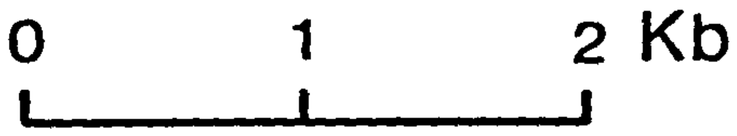


Fig. 3.5

Comments on rDNA variant clone restriction maps

The clones λAar29, λAar35 and λAav1 each contain a single EcoRI restriction fragment of 8.4, 8.2 and 8.6 Kb in length respectively. All three clones were positive with the *Drosophila* rDNA and *Ae. aegypti* I.T.S./5.8S rDNA probes used in Section 3.4 and clearly contain regions homologous to the *Ae. aegypti* normal rDNA repeat. The clone λAar29 is identical to the 7.2Kb EcoRI fragment of clones λAar1, 3 and 7 with the exception of the 1.7 Kb HincII/KpnI fragment of the normal rDNA repeat. This is 1.2 Kb longer (2.9 Kb) and contains two extra restriction sites (one PvuII and one HincII site) which are not present in the non-variant rDNA clones. Clone λAar29 therefore contains 1.2Kb of extra sequence within the rDNA unit. It is not possible from the restriction map to judge whether this extra DNA is due to an insertion event, duplication of rDNA, or the presence of a longer non-transcribed spacer region in this clone.

The restriction map of clone λAar35 differs from that of the normal rDNA repeat 7.2 Kb EcoRI fragment in only one region. The 0.8 Kb EcoRI/PvuII fragment of the clones λAar 1, 3, 7 and 9 is replaced by a 1.8 Kb long restriction fragment in λAar35. In addition, this region (corresponding to internal transcribed spacer and 5.8S coding sequences in the proposed transcription unit shown in Figure 2.5b) contains a HincII site adjacent to the PvuII site which is not present in the normal rDNA repeat. λAar35 may therefore contain an extra 1.1 Kb of DNA within this region (insertion or duplication) or the region found in the normal rDNA repeat may have been largely replaced by a DNA segment containing the EcoRI site which forms the left-hand end of this clone (as shown in Figure 3.5).

The clone λAav1 was isolated on the basis that hybridisation signal was not observed for this clone with the non-transcribed spacer probe used in the differential screening experiment (Section 3.4) and therefore contains only part of the 7.2 Kb EcoRI restricion fragment found in the normal ribosomal clones. The restriction map of λAav1 shown in Figure 3.5 confirms the absence of the non-transcribed spacer in this clone. This region is replaced by a longer DNA segment, which contains a PvuII site linked closely with a HincII site, as observed in the altered region of clone λAar29. This clone may represent the end of a rDNA tandem repeat (terminating at the 3' end of the 28S rRNA gene) or a rDNA repeat containing an insertion, which in turn contains the EcoRI site which forms the right-hand terminus of this clone as shown in Figure 3.5.

None of the three variant clones mapped contain the extra HincII site present in λAar9 (see Figure 2.5a). λAar35 does however show a restriction polymorphism distinct from the alteration at the left-hand end of this clone. The KpnI site 1.7 Kb from the right-hand end of the other rDNA clone restriction maps shown in Figures 2.5a and 3.5 is absent from λAar35. This polymorphism is located near the 3' end of the 28S rRNA gene as shown in the proposed transcription map shown in Figure 2.5b. The accuracy of the transcription map is insufficient to resolve the exact location of this polymorphism with respect to the 28S rRNA gene/N.T.S. junction.

λAas1

This clone showed hybridisation to the N.T.S. probe, but not to the other rDNA probes used for the differential screening in Section

3.4. The insert DNA of this clone consists of two EcoRI fragments (5.1 Kb and 3.5 Kb). When EcoRI digested λAas1 whole clone DNA was probed using *Ae. aegypti* rDNA, the 5.1 Kb EcoRI fragment showed strong hyridisation signal and the 3.5 Kb fragment showed zero hybridisation signal. (Results not shown).

No restriction map of the 5.1 Kb restriction fragment of λAas1 was obtained due to problems encountered in digesting the DNA extracted from low-melting Agarose. Partial digestion data indicates that this fragment contains three HincII sites, several KpnI sites, one HindIII site 0.64 Kb from one end, and a single SacI site 1.8 Kb from one end. No sites were detected for the enzyme PstI.

3.6; Further Analysis of the Variant Ribosomal Clone λAav1

The clone λAav1 was selected for further analysis because of the ease with which the variant segment of this clone could be isolated (see below). The aim of the further analysis was to determine whether this variant represents the end of a tandem rDNA repeat or is the product of the insertion of non-ribosomal DNA (e.g. a transposable genetic element) into the rDNA repeat.

The gel used for the restriction mapping of λAav1 clone DNA was blotted onto nitrocellulose and probed using *Ae. aegypti* rDNA (the total insert DNA from clone λAar7). No hybridisation was detected to the 2.7 Kb EcoRI/PvuII fragment of this clone. This restriction fragment replaces the non-transcribed spacer region found in the other ribosomal clones. A quantity of this non-ribosomal DNA fragment was isolated for more detailed analysis as follows; 200 μg of λAav1 total clone DNA was digested to completion with the restriction endonucleases EcoRI and PvuII. After size fractionation through 1% low-melting Agarose, the 2.7 Kb and 5.9 Kb EcoRI/PvuII fragments which comprise the insert of λAav1 were clearly resolved from vector DNA. Extraction of the DNA from these two bands (see Materials and Methods) yielded 7.5 μg and 14 μg respectively of the purified fragments (approximately 50% recovery).

The clones λAav1 and λAar29 both contain closely linked PvuII and HincII sites which are not found in the normal *Ae. aegypti* rDNA repeat (see Figure 3.5). These sites are found at different distances from the next adjacent HincII site in the two clones. To determine whether any homology existed between the variant regions of

the two clones, the 2.7 Kb EcoRI/PvuII fragment from λAav1 was nick-translated and used to probe a Southern transfer of DNA from the other variant ribosomal DNA clones (λAar29, λAar35 and λAas1). No homology was detected between λAav1 non-rDNA and any of the other variant rDNA clones under the hybridisation conditions employed (see Materials and Methods). (Results not shown).

3.7; Genomic organisation and species specificity of sequences homologous to the non-rDNA segment of λAav1

Genomic DNA from a range of dipteran species was digested with the restriction endonuclease EcoRI, size fractionated through 0.8% Agarose and transferred to nitrocellulose. This Southern transfer was then hybridised with nick-translated λAav1 non-rDNA. The result of this experiment is shown in Figure 3.7a. Zero hybridisation signal was observed for all species with the exception of the species of origin of λAav1 (*Ae. aegypti*, Bangkok strain). A list of the species which did not show hybridisation with this probe is given in the legend to Figure 3.7a.

The observed pattern of hybridisation (to the entire genomic digest of *Ae. aegypti* DNA) is indicative of a repeated sequence which is interspersed with other sequences. No discrete bands of hybridisation signal were evident when a number of different restriction endonucleases were used to digest *Ae. aegypti* genomic DNA. (Results not shown).

To determine whether this type of hybridisation pattern was a consequence of the DNA used having been prepared from several hundred

pooled larvae, v1 non-rDNA was used to probe a Southern transfer of EcoRI digested DNA prepared from individual *Ae. aegypti* (Bangkok strain) adults. The hybridisation pattern obtained was again similar to that shown in Figure 3.7a. (Results not shown).

Of the species tested in Figure 3.7a, only *Ae. aegypti* was found to contain sequences homologous to the λAav1 non-rDNA fragment. To determine whether this specificity extended to the strain of *Ae. aegypti* used (Bangkok), v1 non-rDNA was used to probe a Southern transfer of genomic DNA from different *Ae. aegypti* strains. Genomic DNA (1μg) from each strain was digested with EcoRI and size fractionated through 0.8% Agarose on two minigels (see Materials and Methods). The gels were blotted onto nitrocellulose and the filter hybridised with nick-translated v1 probe DNA. The result is shown in Figure 3.7b. Details of the different *Ae. aegypti* strains tested are given in the legend to Figure 3.7. As for the other experiments, the form of the hybridisation signal is a smear to genomic fragments of all sizes. On this autoradiograph however, some faint bands are visible superimposed over the genomic smear. These bands may represent a sub-fraction of the λAav1 non-rDNA homologues which are present in a tandemly repeated arrangement. These are probably visible in this instance due to the fact that minigels were used (and may give better resolution of some bands) and intensifying screens were not used (this usually gives sharper bands of hybridisation signal). The poor resolution of these bands precludes the evaluation of intraspecific differences in hybridisation patterns.

Figure 3.7a; Result of probing a Southern transfer of EcoRI digested genomic DNA from a wide range of diptera with the non-rDNA restriction fragment from λAav1. Shown are *Culex pipiens quinquefasciatus* (TDK5 strain), *Aedes aegypti* (Bangkok strain) and *Aedes albopictus* (Warwick cell-line). No hybridisation was observed to genomic DNA from *Drosophila melanogaster* (Canton S. strain), *Ae.polynesiensis (Taveuni)*, *Ae. psuedoscutellaris* (cell-line), *Ae.horrescens*, *Ae.togoi*, *Ae scutellaris*, *An.gambiae* (Warwick cell-line), *An.stephensi* or *Simulium sanctipauli* (blackfly). (Results not shown).

Figure 3.7b; Result of probing a Southern transfer of different *Ae.aegypti* strains with λAav1 non-rDNA. DNA (1µg) was digested with EcoRI and size fractionaterd through 0.8% Agarose on two minigels prior to blotting onto nitrocellulose and hybridisation. Lanes 1-3; Bangkok strain adults, pupae and larvae, lane 4; London strain larvae, lane 5; Mos20A cell-line, lane 6; Porto-Novo larvae, lane 7; Cayenne larvae, lane 8; Liverpool larvae, lane 9; Warwick cell-line, lane 10; American ss. larvae.

Fig 3.7a

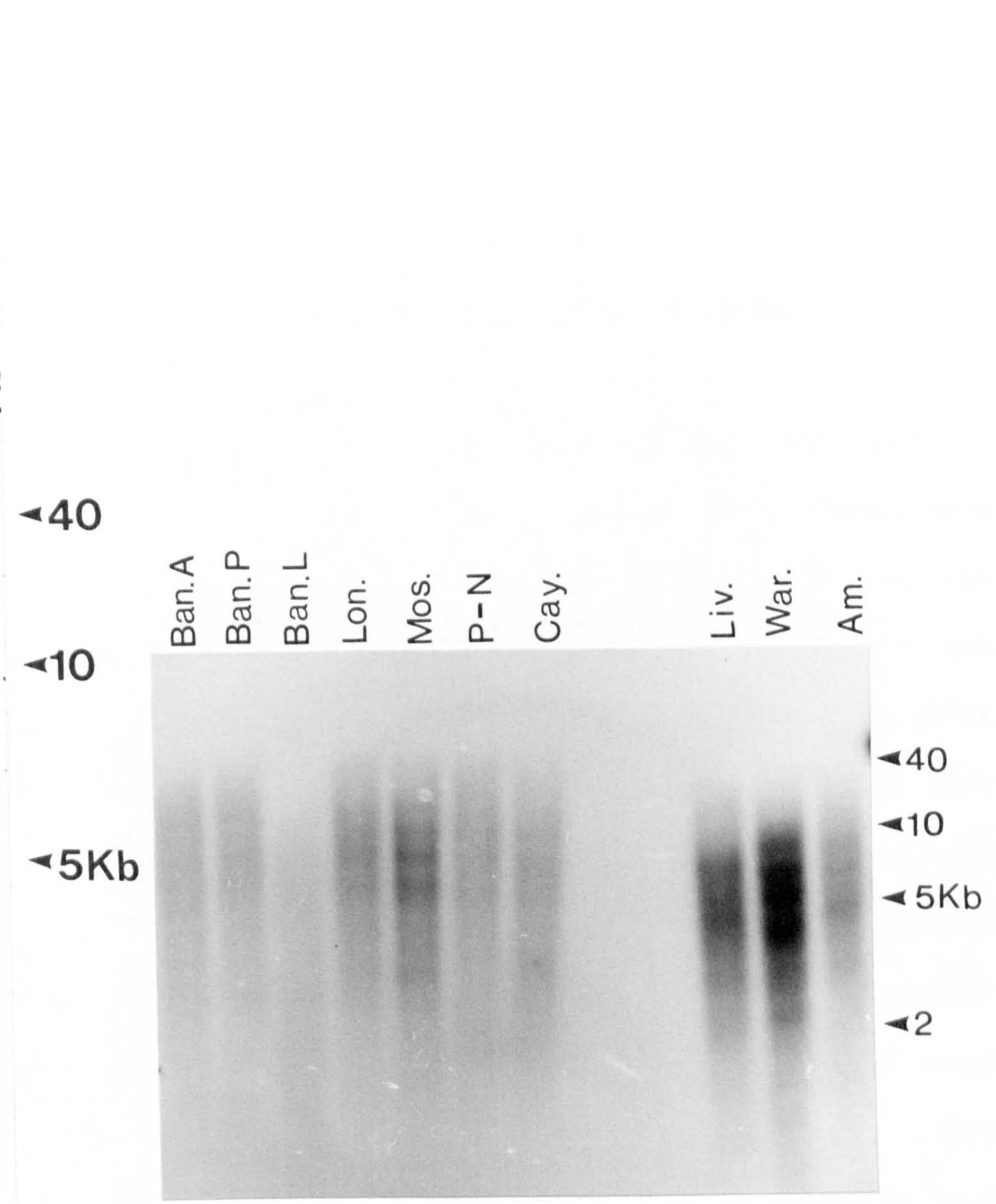


Fig 3.7b

3.8; Differential screening of genomic clones with λAav1 non-ribosomal DNA and λAar7 ribosomal DNA hybridisation probes

To investigate the distribution of sequences homologous to λAav1 non-rDNA in the *Ae. aegypti* genomic library, 1800 clones were screened in duplicate using v1 non-rDNA and λAar7 insert DNA (an intact Ae. aegypti rDNA) repeat as hybridisation probes. Results are shown in Figure 3.8. 170 positives were obtained with the rDNA probe and approximately 16 clear positives were obtained using the v1 non-rDNA probe (10% and 1% of the total clones screened respectively). No correllation was evident between the clones which showed hybridisation signal with the two probes. One clone which did appear to show hybridisation signal with both probes was subsequently found to consist of two separate plaques when re-screened at low plaque density. The higher background of hybridisation to all plaques with the v1 non-rDNA probe is probably due to the short length of the DNA fragment (2.7 Kb) which was nick-translated, producing short probe fragments.

These results clearly show that sequences homologous to v1 non-rDNA are infrequently associated with rDNA in the Ae. aegypti genomic library.

The pattern of hybridisation obtained on the Southern transfers of *Ae. aegypti* genomic DNA and the significant proportion of clones showing hybridisation with the v1 non-rDNA probe suggests that sequences homologous to v1 non-rDNA are present in middle to high copy number in the *Aedes aegypti* genome. The proportion of clones (1%) which hybridise with the v1 non-rDNA allows a rough estimate of copy number to be made.

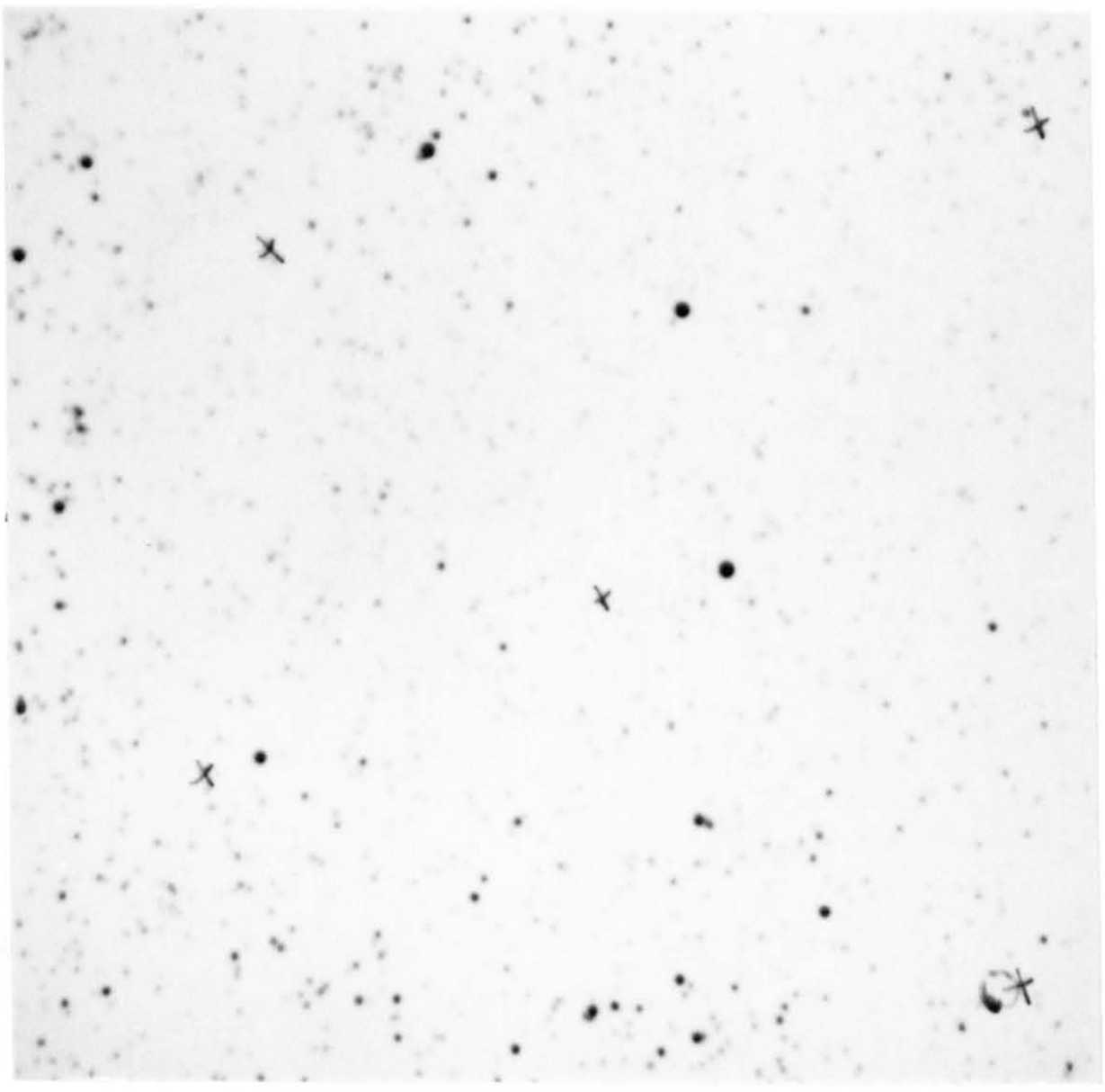


Fig.3.8a

b

Figure 3.8; Result of screening replica phage lifts of approximately 1800 *Ae.aegypti* genomic clones with a, λAav1 non-rDNA and b, *Ae.aegypti* rDNA (insert from clone λAar7) as hybridisation probes.

$$\text{Copy no.} = \frac{\text{genome size}}{\text{ave. insert size}} \times \text{proportion of positive clones}$$

$$\text{Copy no.} = \frac{8 \times 10^6}{10^3} \times 0.01 = 8000$$

This assumes the presence of one repeat only in λAav1 non-rDNA and homologous clones and that the proportion of clones which were positive with this probe is representative of the number of copies in the genome.

3.9; Detailed restriction mapping of the λAav1 non-rDNA EcoRI/PvuII restriction fragment

To obtain a detailed restriction map of the λAav1 non-rDNA segment, three tetracutter restriction enzymes (HpaII, HaeIII and TaqI) and HincII (a hexacutter with reduced specificity) were used singly and in combination to restrict gel-purified λAav1 2.7 Kb EcoRI/PvuII fragment DNA. Digested DNAs were size-fractionated through 2% low-melting Agarose to resolve the small DNA fragments produced. HaeIII digested pAT153 plasmid DNA was used as a low molecular weight marker (see Materials and Methods).

The derived restriction map is shown in Figure 3.9. The positions of the restriction sites in base pairs from the PvuII terminus are listed in the figure legend.

No internal repeat structure is evident within the non-rDNA restriction fragment of λAav1 from the restriction mapping data.

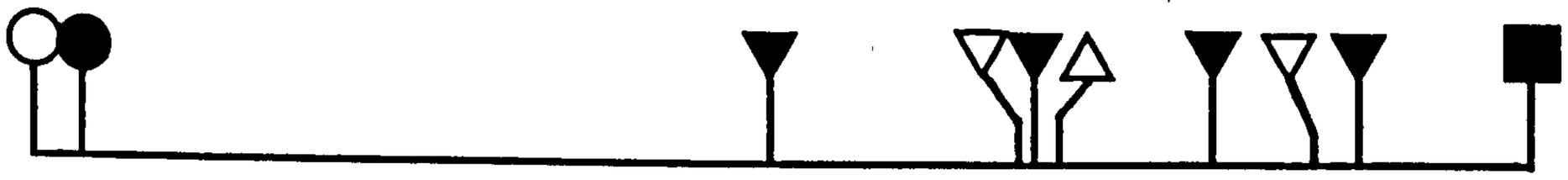


Fig. 3.9

1 Kb

Figure 3.9; Fine detail restriction map of the λAav1 non-rDNA restriction fragment. The distances of recognition sites from the PvuII terminus (in base pairs) are as follows; HincII; 100, TaqI; 1340, HaeIII; 1790, TaqI; 1340, HaeIII; 1790, TaqI; 1810, HpaII; 1860, TaqI; 2110, HaeIII; 2300, TaqI; 2380, EcoRI; 2700.

Key to symbols;

PvuII;

TaqI;

HpaII;

HincII;

HaeIII;

EcoRI;

3.10; Isolation and preliminary analysis of genomic clones showing homology to λAav1 non-ribosomal DNA

Eight of the clones which showed strong hybridisation signal with λAav1 non-rDNA in Section 3.8 were picked and re-plated to obtain single plaques. A single fresh plaque from each was used to inoculate 250 ml cultures. 1-2 mg of clone DNA was obtained from each lysate. 1μg of each clone DNA was digested with EcoRI and size fractionated through 0.8% Agarose. The sizes of the insert DNA of each clone (designated λAav11- λAav18) were as follows;

λAav11;	8.2 Kb
λAav12;	7.4 Kb
λAav13;	4.4 Kb
λAav14;	8.0 Kb
λAav15;	9.1 Kb
λAav16;	5.5 + 4.3 Kb
λAav17;	8.8 Kb
λAav18;	6.4 Kb

The gel used for insert size determination was blotted onto nitrocellulose and the Southern transfer probed usng λAav1 non-rDNA. All clones were confirmed as having v1-homologous inserts. Of the two EcoRI fragments contained in clone λAav16, only the 5.5 Kb fragment hybridised with the v1 non-rDNA probe. (Results not shown).

To compare the restriction enzyme digestion patterns of the eight λAav1-homologous clones with that of λAav1, total clone DNA was analysed as follows; 2μg of each clone DNA was double-digested with

EcoRI (to cleave the insert DNA from the vector DNA) and with one other enzyme (KpnI, XhoI, SacI or HindIII). The double-digested DNAs were then size fractionated through 1% Agarose. Photographs of the ethidium bromide stained gels are shown in Figure 3.10. The sizes of the λgtWESλB vector bands (common to all clones (are as follows.

EcoRI + KpnI; 17.43, 1.52, 2.76 and 13.84 Kb.

EcoRI + XhoI; 21.71, 1.81 and 12.03 Kb.

EcoRI + SacI; 21.71 and 13.4 Kb (no SacI sites in vector arms).

EcoRI + HindIII; 21.71, 5.24, 0.58, 4.46 and 4.14 Kb.

These vector DNA bands are clearly visible on the stained gels shown in Figure 3.10. Interestingly, no other restriction fragments produced are common between the nine different clones digested with the same combination of restriction enzymes.

These results, together with the finding that all clones contain inserts of different sizes, is evidence that the segment of DNA common to these clones is either short or quite variable in sequence between the different clones. The library screening results and Southern blotting data, together with the finding that *Ae. aegypti* has a short interspersed pattern of repeated sequences (see Section 1 of this thesis) provide strong evidence that the non-rDNA segment of the clone λAav1 contains one or more short interspersed repeated sequences.

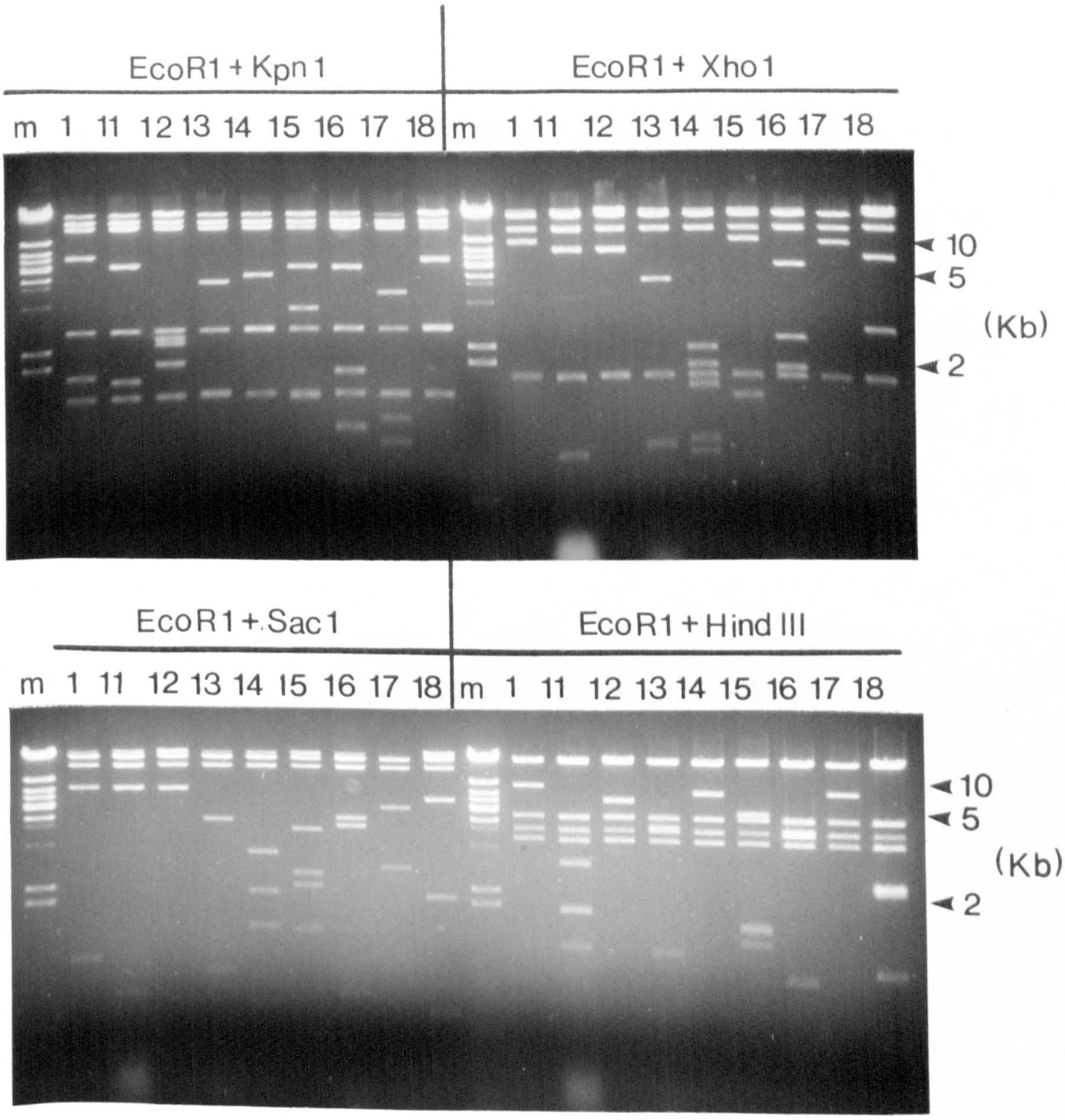


Fig.3.10

Figure 3.10; Restriction enzyme digestion of total clone DNA from clone λAav1 and clones showing homology to λAav1 non-ribosomal DNA. m; λ marker (see materials and methods) 1; λAav1, 11-18; λAav11-λAav18.

3.11; Hybridisation characteristics of the restriction fragment adjacent to the non-rDNA restriction fragment of λAav1

The 5.9 Kb PvuII/EcoRI restriction fragment which comprises the insert DNA of λAav1 minus the non-ribosomal DNA of this clone (see Section 3.6) was digested with the restriction enzyme HincII. After size fractionation through 1% low-melting Agarose, the 2.1 Kb PvuII/HincII fragment corresponding to the restriction fragment adjacent to the non-rDNA restriction fragment of λAav1 was extracted as described in Materials and Methods. This DNA was then nick-translated and used to re-probe the Southern transfer shown in Figure 3.3 of DNA from different *Ae.aegypti* strains digested with EcoRI. The result is shown in Figure 3.11.

This shows that the restriction fragment adjacent to the non-rDNA restriction fragment of λAav1 has hybridisation characteristics similar to those of ribosomal DNA. No extra bands of hybridisation signal are evident with this probe compared to the intact rDNA repeat probe used for Figure 3.3. Some bands of hybridisation signal are evident with the intact rDNA repeat probe, but absent with the probe used for Figure 3.11. These include the 1.8 Kb EcoRI fragment containing 18S rRNA coding sequences which are not contained in the probe used for Figure 3.11. Also, some minor rDNA bands are absent from the autoradiograph shown in Figure 3.11 (the 3.4 Kb band on the Cayenne strain digest shown in Figure 3.3. for example). These may represent variant rDNA fragments containing only part of the *Ae. aegypti* 7.2 Kb EcoRI rDNA fragment.

No hybridisation signal was observed when nick-translated λAav1 2.1 kb PvuII/HincII fragment DNA was used to probe λAav1 non-rDNA

immobilised on nitrocellulose (result not shown). This indicates that there is no repeated segment common to the two restriction fragments.

These results show that the non-rDNA in the clone λAav1 does not extend throughout the adjacent restriction fragment , which appears to consist of rDNA-homologous sequences.

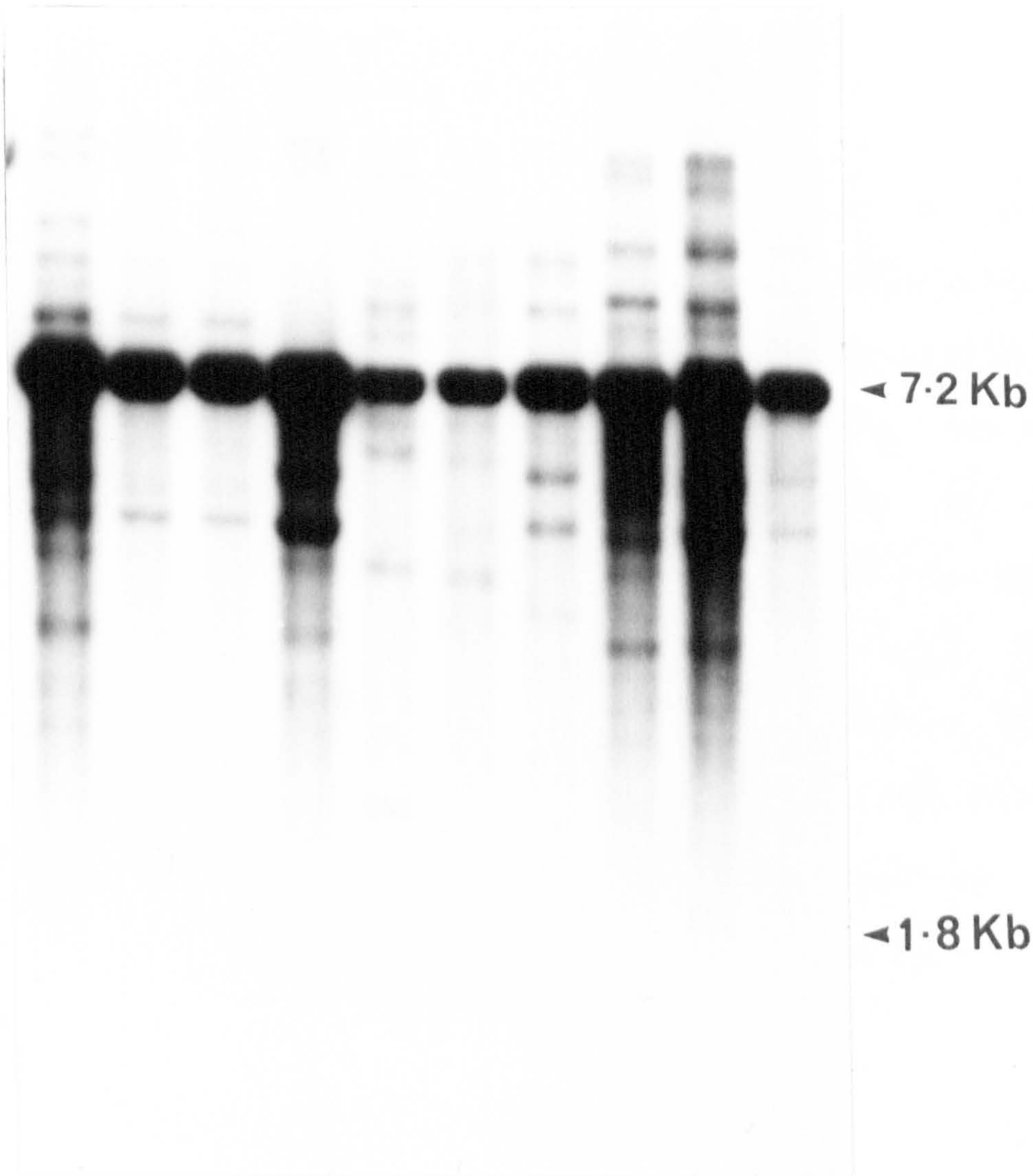


Fig. 3.11

Figure 3.11; Hybridisation characteristics of the restriction fragment adjacent to the non-rDNA fragment contained in clone λAav1. The Southern transfer used is the same as that used for figure 3.3a.

Further Analysis of the Variant Ribosomal Clone λAar35

The variant ribosomal clone λAar35 was selected as being a variant on the basis of the size of the insert DNA in this clone (see Section 3.4). The restriction map of this clone (Figure 3.5) differs from that of the typical *Aedes* rDNA repeat in one region. The 0.8 Kb EcoRI/PstI fragment at the left-hand end of the normal rDNA repeat (as shown in Figure 3.5) is 1.8 Kb in length in λAar35 and contains an extra HincII recognition site adjacent to the PstI site. It was decided that this 1.8 Kb restriction fragment represented a defined region of DNA amenable to more detailed analysis. The following sections describe the subcloning, hybridisation properties, fine-scale restriction mapping and partial DNA sequencing of this restriction fragment from λAar35.

3.12 Subcloning of the 1.8 Kb EcoRI/PstI restriction fragment from λAar35 into pUC12 and pUC13

The 8.3 Kb insert DNA fragment from λAar35 was isolated from low-melting Agarose as described in Section 3.5. 3 μg of insert DNA was digested with the restriction endonuclease PstI and then size fractionated through 1% low melting Agarose. Recovery of the 1.8 Kb EcoRI/PstI restriction fragment (see Materials and Methods) yielded 0.35 μg of DNA. This corresponds to a recovery of aproximately 50%.

The vectors chosen for sub-cloning this restriction fragment were pUC12 and pUC13 [103]. These plasmid vectors are almost identical, but contain a multiple cloning site (MCS) in opposite orientations with respect to the adjacent *E. coli* β-galactosidase promoter. Ligation of insert DNA into the MCS usually disrupts the promoter function (although in some cases, read-through of the inserted DNA may give rise to an altered, but still functional polypeptide) and thus allows the direct selection of recombinants using the 'X-gal' substrate. In this assay, non-recombinants generate blue colonies whereas recombinants usually give colourless colonies. The pUC plasmids also contain suitably located M13 sequencing and reverse sequencing priming sites which allow the ends (e.g. 300 bp) of the insert DNA to be sequenced by the method of Sanger *et al* [104].

Plasmid DNA of pUC12 and pUC13 was double-digested with PstI and EcoRI (which both cut once within the MCS) and dephosphorylated using calf intestine alkaline phosphatase (see Materials and Methods). For each of the two vectors, 70 ng of insert DNA and 200 ng of vector DNA (1:2 molar ratio) were ligated overnight at room temperature using T4 DNA ligase (0.5 units, BCL) in a 5 µl reaction volume. This ligation mixture was then used to transform competent *E. coli* MC1060 as described in Materials and Methods. Transformants were selected using ampicillin (200 µg/ml). Four transformants from each transformation were picked and plasmid DNA prepared by the medium-size, quick plasmid preparation method described in Materials and Methods. In both cases, three transformants were found to contain 1.9 Kb inserts and one to contain no insert DNA when subsequently digested with PstI and EcoRI. Digestion of one pUC12 and one pUC13 recombinant with TaqI and HincII restriction endonucleases confirmed

that these two clones contained the desired insert DNA, cloned in opposite orientations. These were designated pUC12-r35 and pUC13-r35. Large plasmid preparations of each clone (see Materials and Methods) yielded approximately 1 mg of each clone DNA.

3.13. Hybridisation analysis of λAar35 1.8 Kb EcoRI/PstI restriction fragment DNA

To compare the hybridisation properties of λAar35 variant rDNA and normal rDNA. DNA from the clone pUC12-r35 was nick-translated and used to probe a Southern transfer of EcoRI-digested *Ae. aegypti* genomic DNA (Bangkok strain adult and pupae, and Mos20A cell-line) and also to screen 8000 *Ae. aegypti* (Bangkok strain) genomic clones. The 7.2 Kb EcoRI rDNA fragment from clone λAar7 was nick-translated and used to probe a replica Southern transfer and phage lift, thereby allowing a direct comparison of the results obtained with the variant rDNA and normal rDNA probes.

The results of screening the 8000 *Ae. aegypti* genomic clones using the two different hybridisation probes was almost identical. All clones which were positive with the pUC12-r35 probe were also positive with the λAar7 insert DNA probe. One clone was positive with the λAar7 probe only. This clone, which may contain variant rDNA under the criteria discussed in Section 3.4 was picked and designated λAav2. No further analysis of this clone has been performed.

The result of probing the two replica Southern transfers with the two different hybridisation probes was similar (see Figure 3.13). No additional bands of hybridisation signal are evident with the pUC12-

r35 probe compared with the λAar7 probe. Some minor bands are more prominent with the λAar7 probe compared to the pUC12-r35 probe (on the Mos20A DNA track for example). The additional hybridisation using the λAar7 probe here and for the genomic library screening is probably due to the presence of rDNA not represented by the pUC12-r35 clone.

These studies indicate that the 1.8 Kb EcoRI/PstI fragment from λAar35 contains at least some region homologous to rDNA, but no other DNA sequences represented at high frequency in the *Ae. aegypti* genome. To obtain more detailed information regarding this region of λAar35, it was decided to further sub-clone regions of the 1.8 Kb EcoRI/PstI restriction fragment for DNA sequencing and further hybridisation studies.

Fig. 3.13

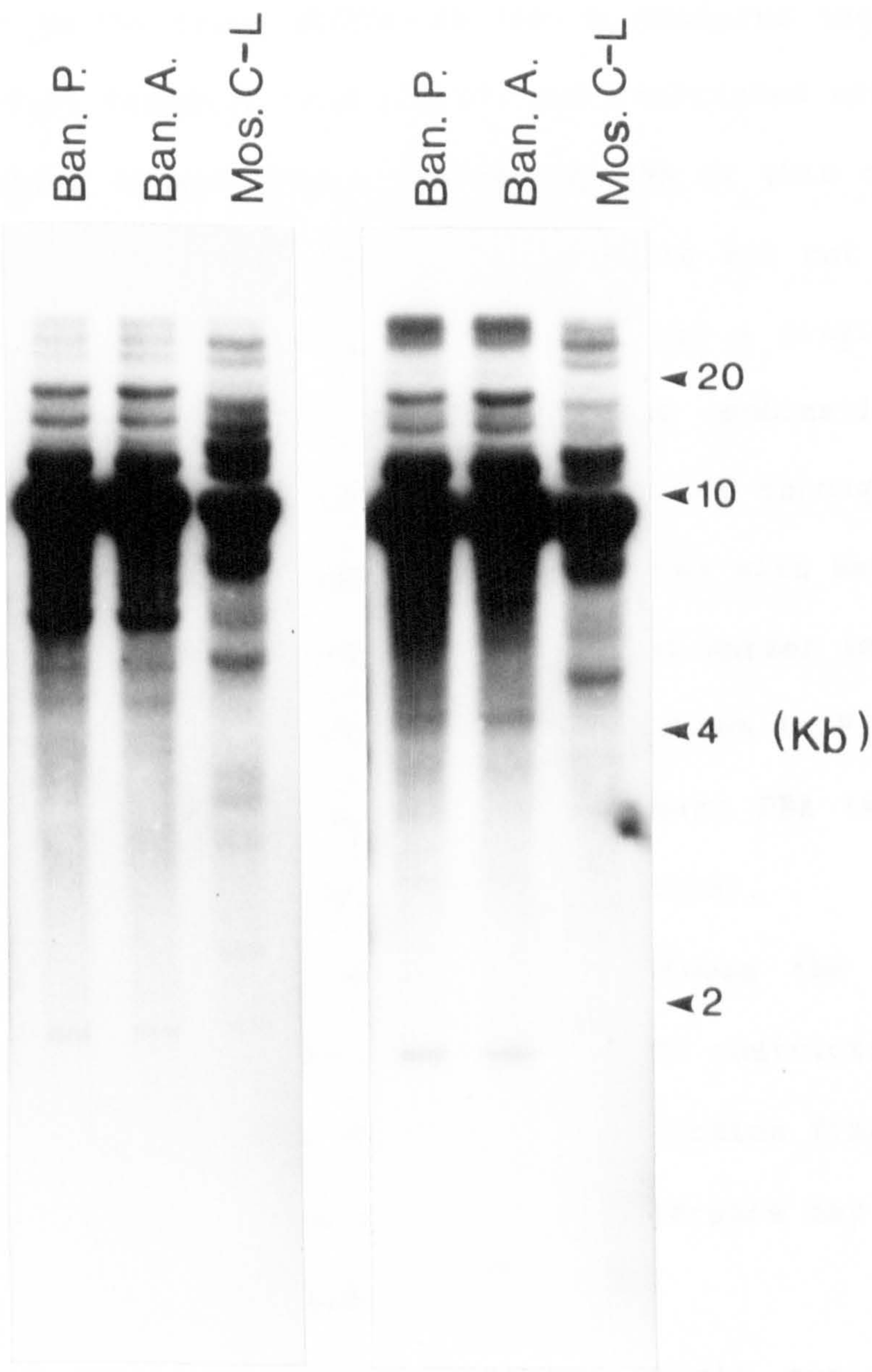


Figure 3.13; Comparison of the hybridisation properties of λAar35 variant rDNA and a normal rDNA segment. Two replica Southern transfers of *Ae.aegypti* (Bangkok strain) pupae and adults and Mos20A cell-line DNAs digested with EcoRI were prepared. Figure 3.13a shows the result of probing one replica with the 7.2 Kb EcoRI fragment from λAar7 (normal rDNA probe) and figure 3.13b shows the result of probing the second replica with DNA from clone pUC12-r35, which contains the variant region of λAar35.

3.14. Detailed restriction mapping of the 1.8 Kb EcoRI/PstI restriction fragment from λAar35

DNA from the clone pUC13-r35 (which contains the sub-cloned 1.8 Kb EcoRI/PstI fragment from λAar35) was restricted with six different enzymes which do not cleave the vector DNA of this clone. Three of the enzymes used (SacI, XbaI and BamHI) did not cut the plasmid DNA and three (AccI, SmaI and HincII) generated a single DNA band. The latter enzymes were then used in various combinations to restrict pUC13-r35 DNA which was then size fractionated through 2% low-melting Agarose. DNA of the plasmid pAT153 digested with HaeIII was used as a low molecular weight restriction fragment marker (see Materials and Methods). The derived restriction map is shown in Figure 3.14. This shows that HincII cuts the pUC13-r35 insert DNA twice (the 160 bp Hinc/EcoRI fragment was previously undetected).

A discrepancy of 0.48 Kb occurs between the molecular weight estimate obtained for the intact EcoRI/PstI restriction fragment from λAar35 (1.8 Kb) and the sum of the restriction fragments sub-cloned into M13 (1.32 Kb). This significant difference may be due to one or more of the following factors.

1, The generation of low molecular weight restriction fragments which were not detected on the restriction mapping gel used.

2, The presence of doublet or triplet bands on the restriction mapping gel. (These may be expected if the region under investigation contains internally duplicated sequences).

3, The difference in gels and molecular weight markers used for the intact fragment and sub-fragment estimates. (1% normal Agarose,

λwt/EcoRI,HindIII marker and 2% low melting Agarose, pBR328/HaeIII marker respectively.)

To resolve the exact source of this discrepancy, the restriction mapping experiment should be re-examined. Other restriction enzymes should be used in addition to previously used enzymes to restrict isolated insert DNA from pUC13-r35 with the aim of generating an unambiguous restriction map. Low molecular weight restriction fragments may be detected by end-labelling the products of the digestions and size fractionating these through non-denaturing polyacrylamide (e.g. 4%) in conjunction with end-labelled molecular weight markers. The presence of doublet or tiplet bands on the restriction mapping gels could be resolved by examining the products of partial digestions.

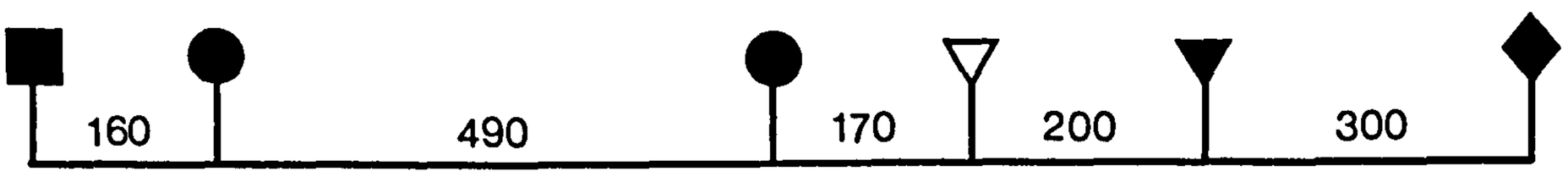


Fig. 3.14

Figure 3.14; Fine detail restriction map of the insert DNA from pUC12-r35 containing the variant region of λAar35. The distances between the restriction sites are shown in base pairs. Restriction enzyme symbols are as shown in figure 3.9.

3.15. Subcloning and DNA sequencing of λAar35 1.8 Kb EcoRI/PstI restriction fragments

DNA of the clone pUC12-r35 (12 µg) was digested with the appropriate restriction enzymes, size fractionated through 2% low melting Agarose and the desired restriction fragments (see Figure 3.14 restriction map) extracted from low-melting Agarose as described in Materials and Methods. M13 replicative form (R.F) DNAs were also digested using appropriate combinations of restriction enzymes and dephosphorylated using calf intestinal alkaline phosphatase. A summary of these cloning experiments is shown below.

Restriction fragment	Method of sub-cloning
1, Pst/AccI (500bp)	PstI + AccI cleaved mp8 and mp9
2. EcoRI/HincII (160bp)	EcoRI + HincII cleaved mp8 and mp9
3. Hinc II (490bp)	HincII cleaved mp9
4. AccI/EcoRI (820bp)	AccI + EcoRI cleaved mp8 and mp9
1a. Pst/AccI (500bp)	Made blunt, into SmaI cleaved mp9

The nomenclature of sequencing clones used is in the form; r35 mp-restriction fragment. e.g. r35 mp8 AccI/EcoRI. The identity of the different orientations obtained in '1a' above was possible from the sequence of the first bases of the insert DNA ('G' for the PstI terminus, 'ATAC' for the AccI terminus).

Fragments for '2' and '3' above were obtained from the same digestion. The blunt-ended HincII fragment (490bp) was cloned only into mp9. Different orientations of this fragment were obtained by

randomly sequencing three clones from '3' above. '1a' above was performed due to difficulty in obtaining clones in '1'.

Ligation of vector and insert DNA was performed overnight using T4 DNA ligase at room temperature. A 2:1 molar ratio of vector (7.3 Kb) to insert DNA was used for each reaction.

Competent *E. coli* JM103 host cells were transformed with the ligation mixtures as described in Materials and Methods. Transformation reactions were plated onto X-gal indicator plates to distinguish recombinants (clear plaques) from cells containing religated vector (blue plaques) as described in Materials and Methods. The M13 bacteriophage minipreparation was used to prepare small amounts of single stranded recombinant DNA (e.g. 2-5 µg) suitable for DNA sequencing. DNA sequencing was performed by the method of Sanger *et al* [104] as described in Materials and Methods. Gels were run either until the bromophenol blue marker reached the bottom of the gel or until the xylene cyanol marker reached the bottom of the gel (short and long runs). Part of an autoradiograph of a sequencing gel (short run) showing the sequence of the simple repeated sequences found downstream of 3' end of the *Ae. aegypti* 18S rRNA gene is shown in Figure 3.15.

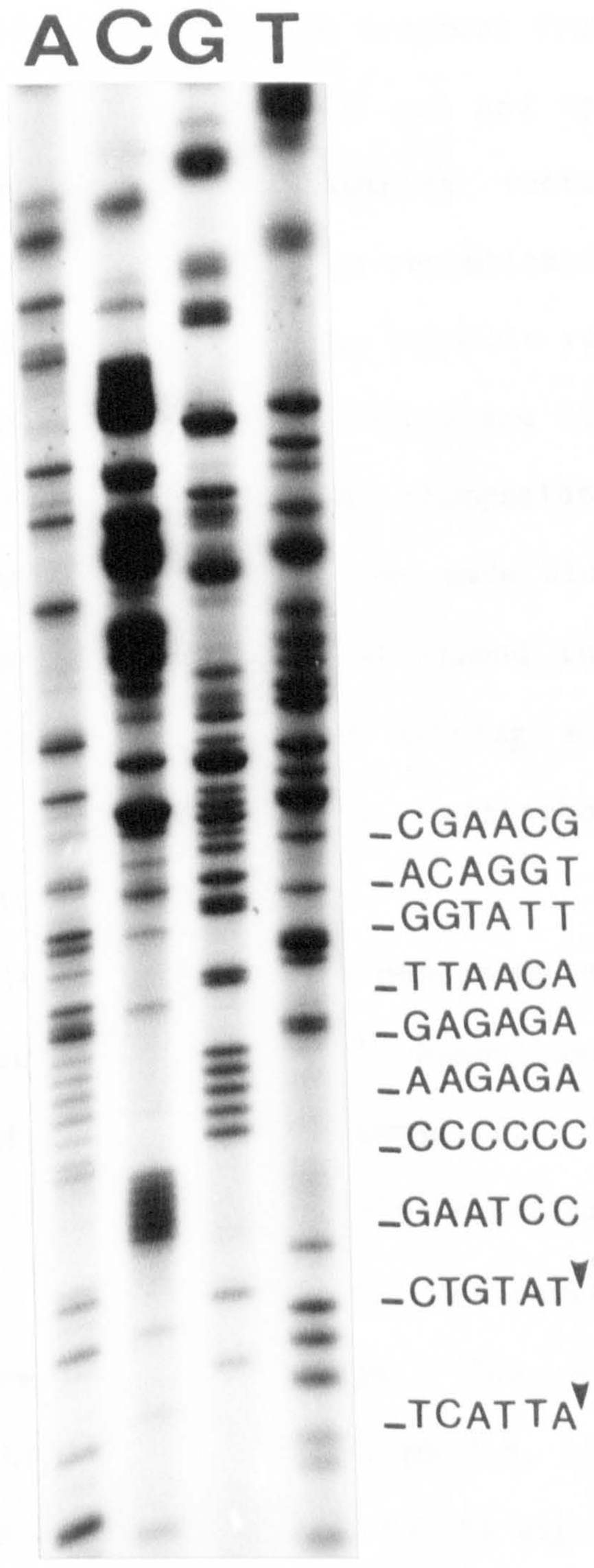


Figure 3.16; DNA sequencing of clone r35 mp9-HincII 'A' showing the sequence of the end of the *Ae. aegypti* 18S rRNA gene and the adjacent regions. ↓ marks the end of the 18S rRNA gene in the corresponding region of *Drosophila melanogaster* rDNA. ↓' marks the position of the end of sequence homology of this region between *Ae. aegypti* and *D. melanogaster* (see Figure 3.17).

Making restriction fragments blunt-ended using T4 DNA polymerase

The 500bp PstI/AccI restriction fragment from pUC12-r35 could not be cloned unidirectionally into M13 mp8 and mp9. This was due to difficulties encountered in preparing vector DNAs which gave sufficiently low backgrounds of non-recombinants when religated and transformed into *E. coli* JM103. One possible reason for this is that PstI generates recessed 5' termini which are less susceptible to the action of calf intestinal alkalinal phosphatase. To overcome this problem, this restriction fragment was made blunt-ended using T4 DNA polymerase (as described below) and cloned into M13 mp9 which had been cleaved within the multiple cloning site with SmaI (which generates blunt ends) prior to dephosphorylation using calf intestinal alkaline phosphatase.

The enzyme T4 DNA polymerase possesses a 5' to 3' polymerase activity and a substantial 3' to 5' exonuclease activity (more than 200 times that of *E. coli* DNA polymerase I). The net result of these activities is that under suitable reaction conditions, T4 DNA polymerase may be used to fill-in recessed 5' termini and simultaneously remove 3' overhangs. The reaction conditions used were: 100 mM NaCl, 100 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1mM each of dCTP, dTTP, dATP and dGTP, at 20°C for 30 minutes.

Examples:

1. AccI terminus (recessed 3' end)

```
- G T              →  dNTPs + T4 DNA pol. →        - G T A T
- C A T A                                          - C A T A
```

2. PstI terminus (recessed 5' end)

```
- C T G C A     →  dNTPs + T4 DNA pol. →        - C
- G                                             - G
```

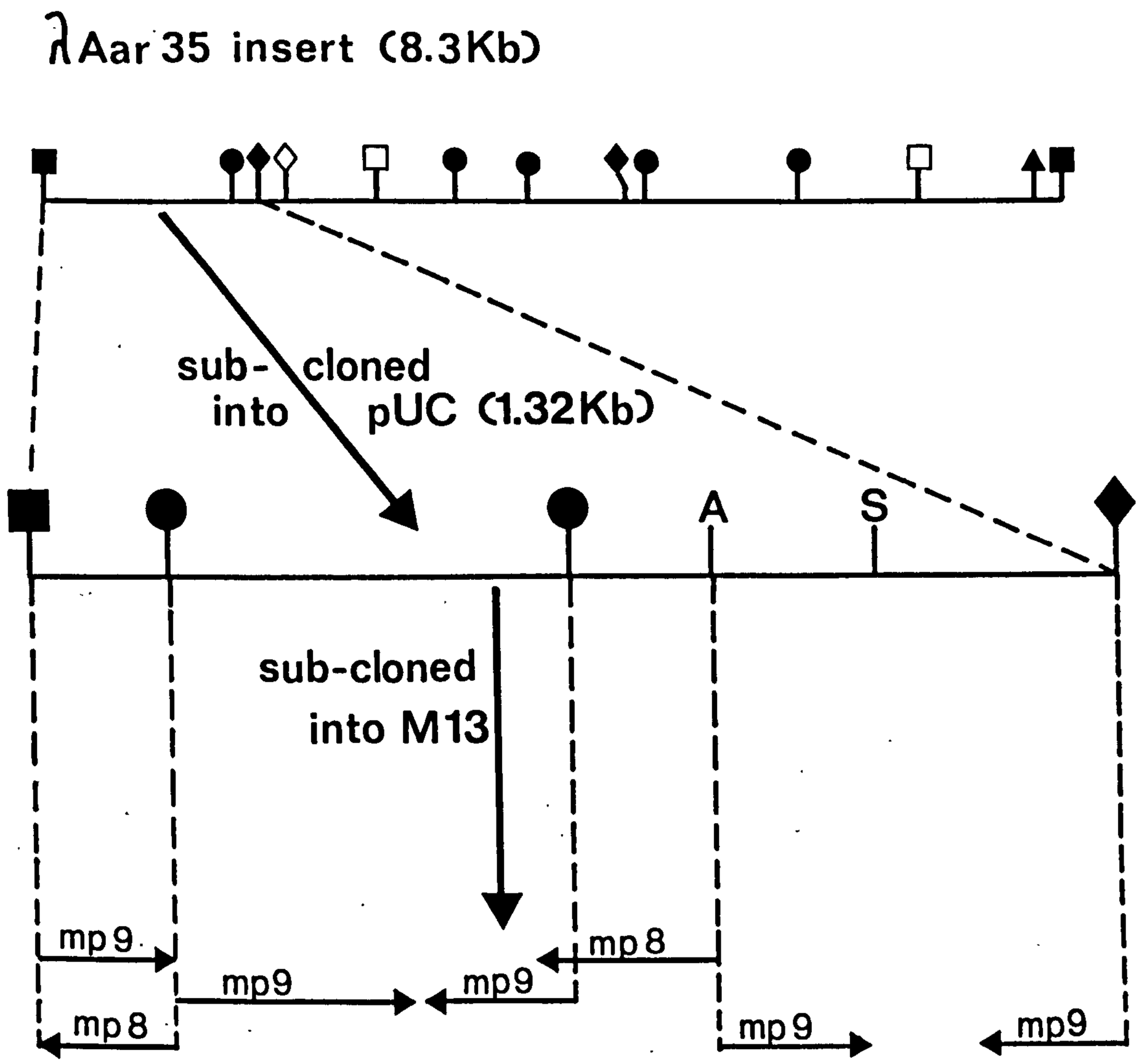


Figure 3.16; Summary of the sub-cloning and partial DNA sequencing of the ribosomal DNA variant λAar35. A = AccI, S = SacI, other restriction enzyme symbols are as for Figure 3.5. → mp8/mp9 → indicate regions sequenced.

3.16 List of sequence data

1. r35 mp9 - EcoRI/HincII (EcoRI → HincII) complete sequence;

```
0                   10                  20                  30
C C A G T A A G T G   C G T G G T C A T T   A T G C T A G C G T

                   40                  50                  60
T G A T T A A C G T   C C C T G C C C T T   T G T A C A C A T C

                   70                  80                  90
G C C C G T C G C T   A C T A C C G A T T   G G A T T A T T T A

                  100                 110                 120
T G A G G T T C C T   G G A G G T G A A C   A T A T T T T G C T

                  130                 140
A G T C C C T G G G   A T T T G A C A T T   T G A A T C
```

(146bp total)

2. r35 mp8 - EcoRI/HincII (HincII → EcoRI). This was found to be the reverse complement of r35 mp9 -EcoRI/HincII

3. r35 mp9-HincII 'A';

```
0                    10                   20                   30
 G A C C G A A C T T   G A T G A T T T A G   A G G A G T A A A G

                    40                   50                   60
T C G T A A C A A G   G T T T C C G T A G   T G A A C C T G C G

                    70                   80                   90
G A A G G A T C A T   T A C T G T A T G A   A T C C C C C C C C

                   100                  110                  120
A A G A G A G A G A   G A T T A A C A G G   T A T T A C A G G T

                   130                  140                  150
C G A A C G T G C G   C C G T T G A T G G   C T G A A G C G T G

                   160                  170                  180
T C G T C C T C C T   G A C G G T C C C G   C C T A G C C T A T

                   190                  200                  210
G G C C T C C C C C   A A G G C A C G G T   T A C C A G G G T T

                   220                  230                  240
G T T T G T T A T T   C T G A G G A C A T   G C T G C A C C G C
```

250 260

T C A C C G A G C A C G G C A C A C G C T T T G C A (266 bp total)

4. r35 mp8 - AccI/EcoRI and r35 - mp9 HincII 'B' (joined);

0 10 20 30

G T C T C A T T C A A T A T A T A T C A G G A C G G T T G C

40 50 60

A T C A A C T G A T T G T T G T A T T C G G T T A A T C G C

70 80 90

T T A A T A T A T A T C A G T T C C T G C T T A A T A T A T

100 110 120

A T C A G G T T C C A T A G A G C T A G G A T G C T A C T G

130 140 150

G T T C A C T G C T G T G A C A G A C C T C C T C C A T T A

160 170 180

G T A T C C A G A C T C A T C G A T A T T G C T A T A C C A

↓ 190 200 210

C C A G *C T G T G C A T G T G G C G G C T* C A T G T G G A T

220 230 240

T G A A G G A T T A G A A A A A G A T T T G T G G T T A A G

250 260 270

G A A G G T G G T G G G T G C C A T T G A C T C G A G C C G

280 290 300

G T A A C T G A T A T G G A C A G C A T T A A C C G C G A C

310 320

T A T T T A C G G A T C T C G T C G A G T A C G (324 bp total)

↓ Marks the beginning of the mp9-HincIIB sequence. The overlap of mp8-AccI/EcoRI sequence with mp9-HincIIB sequence (17 bases) is indicated by bold italics. The complement of the mp9-HincII B sequence is listed here.

5. r35 mp9-PstI/AccI 'A' (AccI → PstI);

0 10 20 30

A T A C G T T T G C A T T G G C G C T C C C A C T A C G G C

40 50 60

A A C G T A G C G C C T T T G T A A T A T G G C A A C A T G

70 80 90

A C A T C C T T T T C T T G C G G C A G A A G C C A A C A A G

```
                100                 110                 120
A G A C T A T C G G   C A A G A G T T G C   T C T G C C C T G T

                130                 140                 150
T T G A C A G T C A   C A T A G C C A T G   G A G  (143 bp total)
```

6. r35 mp9-PstI/AccI'B' (PstI → AccI);

```
0                10                  20                  30
 G C G G T T C T C C   C G C C G G A C C G   G T C C G C G A C A

                40                  50                  60
 C C T G G A C C G T   C C G A G C C T G A   C C G T G C G G G G

                70                  80                  90
 G C A C G C G T C T   G A G G C A A A C C   G C T C G G A C C G

                100                 110                 120
 C T C G G C T C C C   C G T G G C C A A G   C G C C G C G C G A

                130                 140                 150
 A C G G A C C A G C   C G A C G G G A G T   G C A A C C G G A G

 T A T A G C C  (157 bp total)
```

3.17; Comparison of *Ae. aegypti* sequencing data with rDNA sequences from other species

Sequence data was compiled into a DNA sequencing computer text file. The DNA sequence analysis software (University of Minnesota version 2.1) was then used to compare this data with data from other published DNA sequences.

The dot matrix function of the computer program was used to search for homology between the sequence data obtained from λAar35 and the following:

1. The entire 18S and 28S rRNA coding sequences of *Xenopus laevis* [106,107].
2. The 5.8S and 2S coding regions of *D. melanogaster* rDNA [108].
3. The terminal 227bp of the *D. melanogaster* 18S rRNA gene together with 50 bp of adjoining internal transcribed spacer sequence [109].
4. The terminal 227bp of the *Bombyx mori* (silkworm) 18S rRNA gene and 18bp of the adjoining internal transcribed spacer region [110].

Homology of sequence data with 18S rRNA from other species

Strong sequence homology was found between regions within the sequence of clones r35 mp 9- EcoRI/HincII and r35 mp9 - Hinc IIA and the terminal 220-230 bases of the 18S rRNA of other higher eukaryotes. Figure 3.17 shows a comparison of these homologous regions between *Ae. aegypti* (from λAar35 sequencing described here), another Dipteran; *Drosophila melanogaster* and the Lepidopteran *Bombyx mori*. In this figure, only bases which differ from the *Ae. aegypti* sequence are shown on the *D. melanogaster* and *B. mori* sequences. An asterisk is used to denote the absence of a base.

Figure 3.17; Comparison of Sequenced *Ae.aegypti* rDNA with that of *D. melanogaster* and *B. mori* rDNA

35-mp9 EcoRI/HincII;

```
                          0                   10                  20
Ae. aegypti               C C A G T A A G T G   C G T G G T C A T T
D. melanogaster           G                     T   * A
B. mori                                 C

                                     30                  40
Ae. aegypti               A T G C T A G C G T   T G A T T A C G T C
D. melanogaster             A *     C
B. mori                     *       C

                                     50                  60
Ae. aegypti               C C T G C C C T T T   G T A * * C A C A T C G
D. melanogaster                                                 C
B. mori                                 G             A C       C

                                     70                  80
Ae. aegypti               C C C G T C G C T A   C T A C C G A T T G
D. melanogaster
B. mori

                                     90                  100
Ae. aegypti               G A T T A T T T A G   T G A G G T T C C T
D. melanogaster           A                                 C T   C
B. mori                   A     G                           C T T C

                                    110                  120
Ae. aegypti               G G A G G T G A A C   A T A T T T T G C T
D. melanogaster                 C         T     C T G   G A C G C
B. mori                         C C G A C     G * * C G G     G C

                                    130                  140
Ae. aegypti               A G T C C C T G G G   A T T T G A C A T T
D. melanogaster           C T   G G G     T T     C G G T T G T
B. mori                   T T *   A   G   C C   G   C * * * G G C G
```

r35-mp9 HincII 'A';

```
                          1                   10                  20
Ae. aegypti               * * * * G A C C G A A C T T   G A T G A T T T A G
D. melanogaster           A G T T                             T
B. mori                   A G T T         A                   C

                                     30                  40
Ae. aegypti                 A G G A A G T A A A   * G T C G T A A C A A
D. melanogaster                                   A
B. mori                                           A
```

	50										60
Ae. aegypti	G	G	T	T	T	C	C	T	A	G	* T G A A C C T G C G
D. melanogaster											G
B. mori											G *

	70										↓				80
Ae. aegypti	G	A	A	G	G	A	T	C	A	T	T	A	C	T	G T A T G A
D. melanogaster													T		A
B. mori													A	C	G G

	90										100
Ae. aegypti	A	T	C	C	C	C	C	C	C	C	A A G A G A G A G A
D. melanogaster	T	A	T			T	T	A			G T T A T A A T
B. mori	T	G	G	G	A	A	G	A	A	A	(END)

	110										120
Ae. aegypti	G	A	T	T	A	A	C	A	G	G	T A T T A C A G G T
D. melanogaster	A	T			G	T	A		T	T	A T A C A T A A

	123		
Ae. aegypti	C	G	
D. melanogaster	A	A	(END)

KEY;
↓ marks the position of the end of the *Drosophila melanogaster* 18S gene [109]
* denotes the absence of a nucleotide at this position.
Only bases which do not match the *Ae. aegypti* sequence are shown on the *Drosophila* and *Bombyx* sequences.

The two regions of strong homology between *Aedes aegypti*, *D. melanogaster* and *B. mori* are:

A. The sequence from nucleotide 1 to nucleotide 96 of r35 mp9- EcoRI/HincII. This shows strong homology with the region 223 to 130 bp from the 3' end of the *Drosophila* 18S rRNA and the corresponding region in *B. mori*.

B. The sequence from nucleotide 1 to nucleotide 73 of r35 mp9-HincII A. This shows strong homology with the last 75bp of *Drosophila* and *Bombyx* 18S rRNA.

The region of r35 mp9- EcoRI/HincII from position 97-146 shows little homology with the corresponding regions in *Drosophila* and *Bombyx*. Also, the corresponding regions in *Drosophila* and *Bombyx* show little homology over this region between these two species. This region, therefore represents an evolutionarily non-conserved region within the 18S rRNA gene. A comparison of the results obtained here with those of a detailed study of evolutionarily conserved and non- conserved regions within rRNA [111] shows that the two conserved regions described above correspond to two domains within 18S rRNA which are conserved in a wide variety of different species. These conserved domains in *Xenopus laevis* for example, extend from nucleotide 1586 to nucleotide 1696 and from nucleotide 1753 to nucleotide 1821 of the 18S rRNA of this species (1825 bp total). The sequence of the internal transcribed spacer segment of *Drosophila* does not show significant homology with the corresponding region of *Ae. aegypti*. This result is in agreement with the

observation that in the species studied, spacer sequences are not evolutionarily conserved [67].

Quantitation of levels of homology

The levels of homology of the different regions described above were calculated as follows:

$$\% \text{ homology} = \frac{\text{no. of matching bases}}{\text{total no. of bases}} \times 100$$

Results

Regions compared	% homology
1. Conserved region described in 'A' above	
a. *Ae. aegypti/D. melanogaster*	91
b. *Ae. aegypti/B. mori*	91
c. *D. melanogaster/B. mori*	89
2. Conserved region described in 'B' above	
a. *Ae. aegypti/D. melanogaster*	96
b. *Ae. aegypti/B. mori*	93
c. *D. melanogaster/B. mori*	96
3. Non-conserved region of r35 mp9-EcoRI/HincII	
a. *Ae. aegypti/D. melanogaster*	38
b. *Ae. aegypti/B. mori*	24
c. *D. melanogaster*/B. mori	34
4. Internal transcribed spacer region	
Ae. aegypti/D. melanogaster	33

No other regions of homology were detected between the sequenced region of λAar35 and the published rRNA sequences used for comparison.

Other features of sequencing data

1. Repeated sequences downstream of the 3' end of the 18S rRNA coding region

A short region of simple repeated sequences occurs after the 3' end of the 18S rRNA coding region in *Ae. aegypti*. This sequence (C C C C C C C A A G A G A G A G A G A) from positions 83 to 101 in the r35 mp9-HincII A sequence) probably corresponds to part of the internal transcribed spacer region of this species. Also, adjacent to this repeated sequence, the sequence A T T A A C A G G T is almost perfectly repeated to give two adjacent copies from positions 101 to 119 of r35 mp9-HincIIA.

2. A C C G C T C G G A C C G C T C G G

This 9 base-pair direct repeat occurs from positions 78 to 95 in the sequence of r35 mp9-PstI/AccI.

3. A A T A T A T A T C A G

This sequence is directly repeated three times in r35 mp8-AccI/EcoRI within the region from positions 10 to 102 of this clone. Two of these repeats have the sequence C A T near their 3' end. Some sequences adjacent to this repeat are also common between the different repeats, but are not perfectly repeated.

3.18; Further studies of the A A T A T A T A T C A G repeat

This sequence shows some similarity with the 'TATA-box' of eukaryote RNA polymerase II promoters (reviewed in [112]). Also, some features of this repeat show similarity with the consensus sequence of the *E. coli* RNA polymerase promoter. This is shown below for one of the repeats.

E. coli promoter: A-T rich ----- T T G A C A ----- T A T A A T--- C A T

E. coli bases from start: -43 -35 -10 +1

Ae. aegypti repeat: A A T A T A T A T C A G G A C G G T T G C A T

Regions of particular interest within the repeat sequence are underlined. The repetition of this sequence does generate A-T rich regions upstream from the 2nd and 3rd repeats and also places the sequence T T G C A (similar to the T T G A C A of the *E. coli* promoter) 35 bp upstream from the T A T A region of one of the repeats.

To determine whether this isolated sequence did show cross-hybridisation with *E. coli* DNA or whether this sequence was in fact part of the normal rDNA repeat, an oligonucleotide probe was constructed.

Synthesis of an oligonucleotide probe of the λAar35 TATATA repeat

An oligonucleotide probe 31 residues in length was constructed using an automated DNA synthesiser (Du Pont Coder 300). This machine uses a chemistry system based on the phosphite coupling method using β-cyanoethyl phosphoramidites. The 3' terminal nucleotide comes attached to the controlled pore glass resin column. A cyclical process of deprotection, coupling, and protection (capping and oxidation) steps yields the desired oligonucleotide, still attached to the column support. The high efficiency of each cycle (approximately 95% efficiency of coupling) together with the capping function at each step enables highly specific oligonucleotides (e.g. 15-40 residues in length) to be synthesised in large amounts. The sequence of the oligonucleotide synthesised was;

```
                10                  20                  30    bp
A A T A T A T A T C  A G G A C G G T T G  C A T A A T A T A T  A
```

The scale of synthesis used was 1 micromole. The terminal capping group was removed by the machine at the end of the synthesis.

Cleavage of the oligonucleotide from the column support, deprotection of the 3' hydroxyl group and cleavage of the protecting groups on the 'A', 'G' and 'C' residues was accomplished by incubating the column contents in 1.2 ml of concentrated ammonia in an airtight vessel for 2 days at 50°C. (Protecting group on 3' hydroxyl = cyanoethyl, protecting group on 'A' and 'C' residues = benzyoyl, protecting group on 'G' = isobutyrl). The insoluble column

support material was removed by filtration through nylon filter wool packed into the bottom of a Pasteur pipette.

This solution was then freeze-dried to remove most of the ammonia. The lyophilisate was then redissolved in 2 ml of 70% ethanol and freeze dried to remove any remaining ammonia. The lyophilisate was then redissolved in 200 µl of water and precipitated by the addition of 6 volumes of ethanol. After incubation at -20°C overnight, the precipitated oligonucleotide was pelletted (30 minutes, Eppendorf microfuge) and redissolved in 50 µl of T.E. buffer. The yield was estimated by optical density to be 1.4 mg (1 O.D. unit at 260 nm = 33 µg/ml oligonucleotide). Using an estimate for the molecular weight of the 31-mer oligonucleotide of 2×10^4 Daltons, this corresponds to 0.07 micromoles of crude oligonucleotide.

Gel purification of oligonucleotide

A minute amount of crude oligonucleotide (1.4 pmoles/28 ng) was kinase labelled in 10 µl of Kinase buffer containing 10 µCi of α^{32}P ATP and 4 units of T4 polynucleotide kinase for 1 hour at 37°C (see Materials and Methods). This reaction mixture was then fractionated through a Sephadex G100 column (2.5 ml column volume, 200 µl fractions) and the first four fractions which contained radioactivity (fractions 5-8) were separately ethanol precipitated and each redissolved in 10 µl of formamide. The counts per minute of these fractions were 5.7, 5.2, 10.7 and 9.2×10^5. 10^5 counts from each fraction were run in 15% acrylamide (see Materials and Methods) Also, 4×10^5 counts from fraction 5 were added to 8 O.D. units (264

µg) of crude oligonucleotide and this was electrophoresed over four wells of the same 15% Acrylamide gel.

All samples were dissolved in formamide and boiled (100°C) for two minutes prior to loading. Xylene cyanol (which migrates with the mobility of a 25-mer) and bromophenol blue (which migrates with the mobility of an 18-mer) were used as markers. After electrophoresis, the gel was wrapped in cling-film and autoradiographed overnight at room temperature without an intensifying screen. The result is shown in Figure 3.18. This shows that most oligonucleotide is in the form of a 31-mer, although bands corresponding to shorter, terminated oligonucleotides are clearly visible. The separation of the 31-mer from the shorter oligonucleotides by Sephadex gel filtration was not efficient.

The tracer labelled 31-mer oligonucleotide (lanes 5-8 of the autoradiograph) was excised, cut into small strips and incubated in 1 ml of water at 50°C for 12 hours to allow the oligonucleotide to diffuse out of the gel. The aqueous solution was then aspirated and passed through a long Sephadex G-25 column to separate the oligonucleotide from dissolved Acrylamide and urea. (The column used was 7 ml in volume, made in a 10 ml pipette). The fractions containing radioactivity were pooled and freeze dried. The lyophilisate was finally redissolved in 200 µl of T.E. buffer. The yield of purified 31-mer oligonucleotide obtained was 2 O.D. units (66 µg).

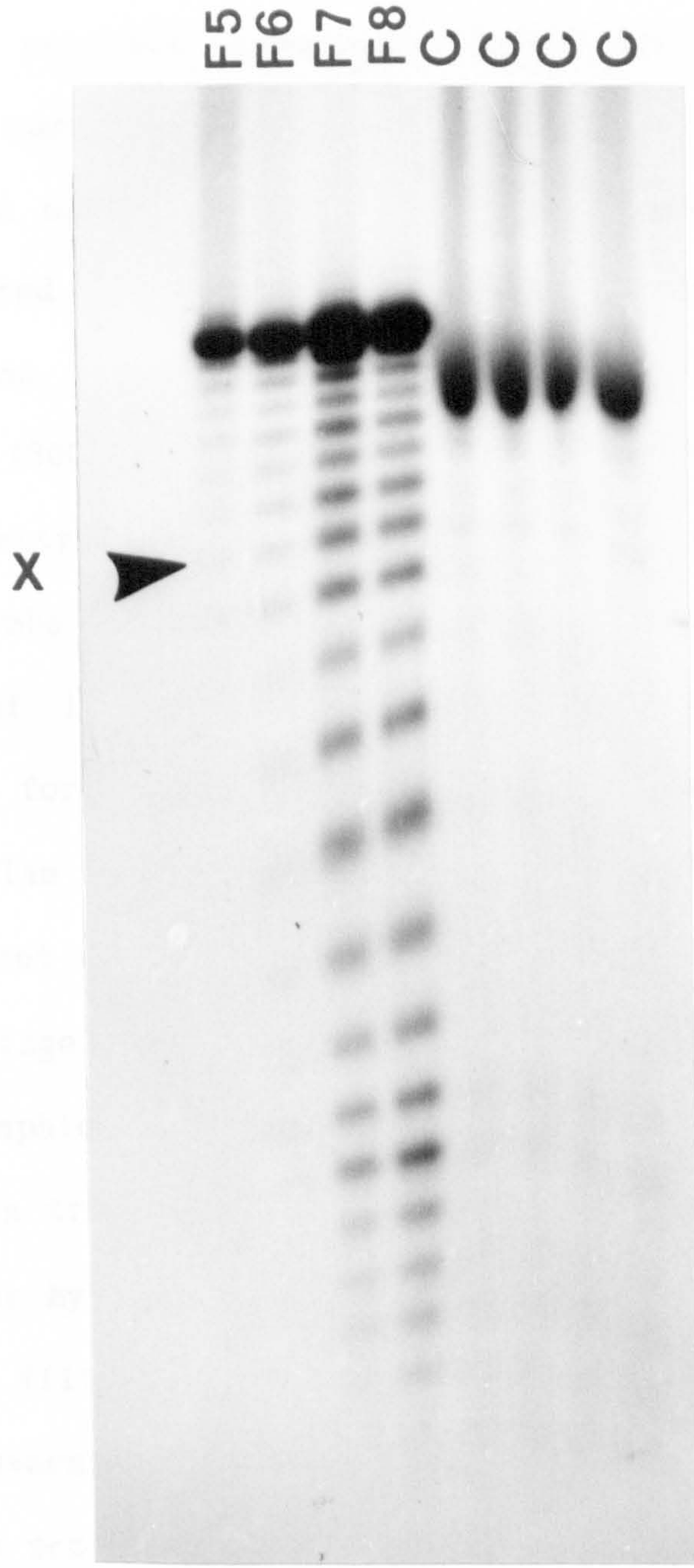


Figure 3.18 Gel purification of TATA oligonucleotide. Autoadiograph of a 15% Acrylamide gel. F5-F8; Sephadex G100 fractionated, kinase-labelled, crude oligonucleotide. C; crude oligonucleotide (2 O.D. units per lane) with added kinase-labelled tracer. 'X' = xylene cyanol marker position (25-mer equivalent).

Use of the TATATA oligonucleotide to probe *Ae. aegypti* and *E. coli* genomic DNAs

Aedes aegypti (Bangkok strain) and *E. coli* (MC1060 strain) genomic DNAs (4 µg) were digested overnight at 37°C with a 2-fold excess of EcoRI restriction endonuclease. The MC1060 DNA was a gift from A.M. Warren of this department. The digested DNAs were then size fractionated through 0.8% Agarose and transferred to nitrocellulose as described in Materials and Methods. Purified oligonucleotide (300 ng) was kinase labelled as described previously. A specific activity of 6 x 10^7 dpm/µg was obtained. This probe was then used to probe the Southern transfer described above. A probe concentration of 10 ng/ml was used in a hybridisation of high stringency (50% formamide, 3 x SSC, 42°C, as described in Materials and Methods). The hybridisation time used was 18 hours. The result of this experiment was that no hybridisation signal could be detected after a low stringency post-hybridisation wash (20°C, 2 x SSC) and 2 day autoradiographic exposure at -75°C with intensifying screen.

The Southern transfer was therefore re-hybridised using the same probe at a lower hybridisation stringency (in 50% formamide, 6 x SSC, at 37°C). The filter was washed at low stringency, (2 x SSC, 20°C) and exposed overnight at -75°C, with intensifying screen. The oligonucleotide probe showed hybridisation signal for *Ae. aegypti* DNA only. This was in the form of a single 7.2 Kb band. This corresponds to the size of the major EcoRI fragment of *Ae. aegypti* rDNA and therefore confirms the oligonucleotide sequence as being derived from this part of the ribosomal DNA repeat.

The failure to obtain hybridisation signal under the stringent hybridisation conditions is probably due to the low G-C content of the oligonucleotide (26%). The melting temperature (Tm) of this sequence is theoretically 56°C under these conditions (see Materials and Methods), just 14°C above the temperature of hybridisation (42°C). Under the low stringency hybridisation conditions used for re-hybridisation, the Tm is increased to 60°C, which is 23°C below the hybridisation temperature (37°C).

3.19 Section 3: SUMMARY

Analysis of rDNA variation within *Ae. aegypti* and between closely releated species

The rDNA 'fingerprints' of a number of different mosquito genomic DNAs digested with the restriction endonuclease EcoRI were compared. The results obtained show a significant level of interspecies differences, even between the very closely related members of the *Aedes scutellaris* group. In general, the degree of variation between different species was consistent with the relatedness (as judged by classification) of the mosquito species compared. The rDNA of different *Ae. aegypti* strains and cell-lines is much less variable. No difference in the length or EcoRI restriction pattern of the major rDNA bands was detected. Some of the minor rDNA bands do show variation between different strains however. This indicates a low level of intraspecies rDNA polymorphism. One of the possible sources of such polymorphism has been identified at the molecular level: sequence heterogeneity within the non-transcribed spacer region of the rDNA repeat (as indicated by restriction site polymorphism within this region). No variation in rDNA 'fingerprint' was evident between genomic DNA prepared from *Ae. aegypti* larvae, pupae and adults.

Isolation of variant rDNA clones

The *Aedes aegypti* genomic library described in Section 1 was screened to isolate variant rDNA clones. Two methods were employed.

A number of randomly picked rDNA clones were analysed by EcoRI digestion to identify size variants. Also, differential screening using the two ends of the major rDNA EcoRI fragment was used to identify variants in which the two ends had become separated. Two rDNA variants were isolated using each of these methods. The two size selected variants λAar29 and λAar35 contained insert DNA fragments of 8.4 Kb and 8.3 Kb respectively. Restriction maps of these clones indicate that these contain the 7.2 Kb normal EcoRI rDNA fragment with a variant region (different in the two clones) caused by an insertion or duplication of DNA at this site. The two rDNA variants obtained by differential screening (λAav1 and λAas1) contain insert DNAs of 8.6 Kb and (5.1 + 3.5 Kb) respectively. The restriction maps and hybridisation characteristics of these clones indicate that λAav1 represents a 7.2Kb normal EcoRI rDNA fragment in which the non-transcribed spacer region has been replaced by a longer segment of non-rDNA, whereas λAas1 contains non-transcribed spacer homologous sequences which are not associated with rRNA coding sequences.

Further analysis of λAav1

The hybridisation characteristics of the non-rDNA segment of this clone were studied. When this non-rDNA segment was radiolabelled and used to probe a Southern transfer of genomic DNAs from a wide variety of different dipteran species, hybridisation was detected only with *Ae. aegypti* DNA. The form of this hybridisation (to DNA fragments of all sizes) was suggestive of an interspersed repeated sequence, which

might also be present in tandemly repeated arrays. Sequences homologous to λAav1 non-rDNA were detected in all *Ae. aegypti* strains tested.

The λAav1 non-rDNA segment was used to screen the *Aedes aegypti* genomic library in parallel with a rDNA probe. Approximately 20 strong positives were obtained using the v1 non-rDNA probe (0.1% of the total clones screened). None of these clones showed hybridisation signal with the rDNA probe. This demonstrated that sequences homologous to the repeats within λAav1 non-rDNA are not normally associated with rDNA.

Eight of the *Ae. aegypti* genomic DNA clones which showed strong hybridisation signal with the λAav1 non-rDNA probe were picked and the clone DNA prepared. All were found to contain insert DNAs of different sizes. Preliminary restriction enzyme analysis of these clones indicates that the insert DNAs do not show any similarity with respect to the number and position of restriction enzyme sites within the insert DNA.

The restriction fragment adjacent to the non-rDNA restriction fragment contained in λAav1 was found to contain rDNA. This region did not generate bands of hybridisation signal in addition to the rDNA bands when used as a hybridisation probe against *Ae. aegypti* genomic DNA.

Further analysis of λAar35

The restriction fragment of λAar35 which showed a 1.1 Kb size increase in this clone relative to the normal rDNA repeat was sub-cloned into pUC12 and pUC13. The hybridisation characteristics of

this segment from λAar35 were found to be very similar to those of the 7.2Kb normal EcoRI fragment of *Ae. aegypti* rDNA. When these two DNA fragments were used to probe replica Southern transfers of EcoRI digested *Ae. aegypti* rDNA, no additional bands were visible with the λAar35 probe. Some bands of hybridistion signal were more prominent with the normal rDNA probe however. These probably correspond to rDNA sequences not present in the λAar35 probe (e.g. non-transcribed spacer sequences and 28S rRNA coding sequences).

When genomic clones were screened using these two hybridistion probes, all clones which were positive with the λAar35 probe were also positive with the normal rDNA probe. One clone (λAav2) showed hybridisation signal with the normal rDNA probe only. A fine detail restriction map of the pUC12 insert DNA was prepared and a number of different restriction fragments were sub-cloned from pUC into the M13 bacteriophage vectors mp8 and mp9. Sequencing of single-stranded DNA from these vectors was successful in obtaining sequence data representing approximately 80% of the 1.32 Kb region of interest from λAar35. (Only one DNA strand has been sequenced for 85% of the sequenced region).

Analysis of the sequence data shows that the left-hand terminus of the variant fragment from λAar35 (as shown in Figure 3.14) contains two regions of DNA sequence which show very strong homology with the 3' end of 18S rRNA. It was deduced that this corresponds to part of the 18S rRNA gene of *Ae. aegypti*. This therefore extends approximately 220 bp into the EcoRI/PstI fragment of λAar35, a finding which is in agreement with the proposed transcription unit of *Ae. aegypti* rDNA shown in Figure 2.5.

The HincII recognition site 143 bp from the left-hand terminus of the EcoRI/PstI fragment of λAar35 was not detected in the initial mapping of the total insert DNA of this clone. (DNA fragments of this size are not usually visualised in 1% Agarose gels). This HincII site may also be present in the other rDNA clones (including the normal rDNA clones) due to the conserved sequence of 18S rDNA at this position.

Comparison of the region of 18S rRNA of *Ae. aegypti* which has been sequenced with the corresponding regions in *Drosophila melanogaster* and *Bombyx mori* 18S rRNA sequenced by other workers [109,110], demonstrates the very high degree of sequence conservation of some regions of rRNA. Two regions showing very strong homology (over 90%) with *Drosophila* 18S rRNA were identified within the *Ae. aegypti* rDNA region sequenced. These two regions (96 and 73 bp in length) are separated by a 50bp region which shows little or no homology between the different species. This finding is in agreement with the finding of other workers [111] that only certain domains within rRNA are conserved (approximately 10% of 28S rRNA coding sequences and 20% of 18S rRNA coding sequences) and that the position of these conserved regions does not vary between different species.

Some repeated sequences have been identified within the remaining sequence of the λAar35 region of interest. Hybridisation analyses using both pUC sub-cloned DNA and a synthetic oligonucleotide of one of these repeated sequences as hybridisation probes has failed to identify any region of DNA within λAar35 with hybridisation characteristics which differ from those of rDNA.

Sections 2 and 3; CONCLUSIONS AND FUTURE PERSPECTIVES

Preliminary studies of *Ae. aegypti* rDNA by Southern blotting analysis have shown that the rDNA repeat of this species is relatively short (9.0Kb) and homogeneous in length. The homogeneity of the rDNA repeats has allowed the construction of a restriction map of *Ae. aegypti* rDNA from Southern blotting data. The copy number of *Ae. aegypti* rDNA has been estimated as approximately 500 per haploid genome (8 x 10^8bp). Ribosomal DNA therefore comprises 0.6% of the total genome of this species (c.f. 2% of the *Drosophila* genome consists of rDNA [23]. As for all other higher eukaryotes studied, the rDNA of *Ae. aegypti* is tandemly repeated in a head-to-tail manner.

Restriction mapping of cloned rDNA repeats has largely confirmed the homogeneity of rDNA in *Ae. aegypti*. The exception discovered is a polymorphic HincII recognition site located in the non-transcribed spacer region.

Comparison of the rDNA of different mosquito species has shown that one EcoRI fragment (1.8Kb in length, containing most of the 18S rRNA coding sequences) is conserved between different species of mosquitoes. This indicates that the two EcoRI sites within *Ae. aegypti* rDNA are present in evolutionarily conserved regions within the rDNA repeat. One of these is the conserved EcoRI site near the 3' end of the 18S rRNA gene, which is also present in this position in *D. melanogaster* and *B. mori* rDNA [109,110]. The other EcoRI site is located near the 5' end of the 18S rRNA gene. The large (7.2 Kb)

rDNA fragment produced by EcoRI digestion of *Ae. aegypti* DNA shows variation in length between different species. This occurs even between the closely related and morphologically very similar species of the *Aedes scutellaris* group. This restriction fragment contains the non-transcribed spacer of *Ae. aegypti*. This region is the least evolutionarily conserved region of rDNA [67].

Some minor rDNA bands which are evident when an homologous rDNA probe is used to probe Southern transfers are invariant between different stages of development, but show variation between different strains of *Ae. aegypti*. One source of these minor bands of hybridisation signal is a limited degree of intra-species polymorphism of rDNA. A study of the rDNA of individual mosquitoes indicates that this is probably due to polymorphism of a fraction of the rDNA repeats within each individual. The possible causes of these minor rDNA bands are; restriction polymorphisms within the rDNA, variation in length of the rDNA repeat (including insertions into, and duplications of, rDNA) and the presence of sequences homologous to rDNA outside the ribosomal repeat. The restriction fragments produced at the ends of tandemly repeated arrays of ribosomal genes may also generate minor rDNA bands of variant length. Restriction mapping of rDNA clones and Southern blotting analysis of genomic rDNA has confirmed that restriction polymorphisms do contribute to variation of *Ae. aegypti* rDNA (the HincII polymorphism in λAar9 and the KpnI polymorphism in λAar35).

Ribosomal variant clones have been isolated by their variant lengths (λAar29 and λAar35) and hybridisation characteristics (λAav1 and λAas1). λAas1 contains sequences homologous to the non-transcribed spacer region of *Ae. aegypti* rDNA, but does not contain

any conserved rRNA coding region. This clone may therefore contain rDNA-homologous sequences which are not located in the rDNA repeat within the *Aedes* genome. Another possibility is that this variant was produced by the insertion of non-rDNA containing an EcoRI site into the NTS of a rDNA repeat. The length of the EcoRI fragment contained in λAas1 which shows homology to a rDNA probe is 5.1 Kb. A minor band of 5.0 Kb is visible on Southern transfers of *Ae. aegypti* DNA digested with EcoRI and probed with rDNA. (Figure 3.3).

The variant λAav1 contains most or all of the rRNA coding regions of the 7.2 Kb EcoRI fragment of *Ae. aegypti* rDNA, but no sequences homologous to the NTS region. In place of the NTS, this clone contains a 2.7 Kb restriction fragment which is not homologous to rDNA. This non-rDNA is specific to the species *Ae. aegypti* and shows the hybridisation characteristics of an interspersed repeated sequence. Preliminary analysis of eight λAav1 non-rDNA homologous clones indicates that the non-rDNA sequences within this clone are not normally associated with rDNA. The different λAav1 - homologous clones show no similarity in restriction enzyme digestion patterns. The conclusion made from these results is that the non-rDNA segment of λAav1 contains one or more short interspersed repeats (see Section I of this thesis). λAav1 may represent part of a rDNA unit at the 3' end of an array of tandemly repeated ribosomal genes. The non-rDNA region of λAav1 would therefore be essentially unique DNA (containing one or more short interspersed repeats) adjoining the rDNA tandem repeat. This would imply that the rDNA tandem repeat ends with the 3' end of a 28S rRNA gene and not a non-transcribed spacer. Circumstantial evidence in support of this is that the NTS of eukaryote rDNA serves an important regulatory function for the rRNA

transcription unit downstream of the NTS [69,70]. Although the restriction fragment adjacent to the non-rDNA restriction fragment of λAav1 contains rDNA, S1 nuclease mapping [113] or DNA sequencing of the rDNA/non-rDNA interphase of λAav1 is required to identify precisely the 3' end of the rDNA sequences within this clone. Further analysis of the hybridisation properties of the restriction fragments within the non-rDNA portion of λAav1 is required to define the length(s) and number of repeated sequences in this clone. Identification a unique sequence region of non-rDNA in this clone would allow the flanking regions to be isolated from the *Aedes aegypti* genomic library. This may resolve whether λAav1 does correspond to the end of a tandem repeat.

The variant clone λAar35 contains part of a rDNA repeat with an extra 1.0 Kb of DNA in the region of the internal transcribed spacer. Detailed hybridisation studies have failed to identify any non-rDNA sequences within the variant region of this clone. DNA sequencing studies have shown that the left-hand terminus of this clone (as shown in Figure 3.5) is located 220bp from the end of the *Ae. aegypti* 18S rRNA gene. This region of 18S rRNA is apparently normal, as judged by comparison with the corresponding regions of 18S rRNA from the distantly related dipteran *Drosophila melanogaster* and the lepidopteran *Bombyx mori*. The region of 18S rRNA sequenced contains two highly conserved regions (over 90% homology to *Drosophila* rRNA) separated by one non-conserved region which shows no significant homology with the corresponding region in *Drosophila*. These findings are consistent with the observed distribution of evolutionarily conserved regions within the rRNA of all organisms studied [67,111]. The conserved regions may be considered as sequence domains essential

to the function of the ribosome. Models of the secondary structure of rRNA have been proposed from the sequence of these domains [62]. The non-conserved regions are also essential (their length and position is quite invariant between different species), but these are unlikely to serve a sequence-specific function.

The region of strong homology between *Ae. aegypti* and *D. melanogaster* rDNA extends beyond the 3' end of the *Drosophila* 18S rRNA gene (as proposed by Jordan *et al* [109]) by six base-pairs. The proposed end of the *Drosophila* 18S rRNA gene corresponds to the end of the sequence homology between *Drosophila* and *Bombyx* rDNA. The extended homology between *Aedes* and *Drosophila* rDNA may indicate that the 18S rRNA gene of both species extends at least a further 6 bp than the reported terminus in *Drosophila*.

Some repetitive DNA sequences have been identified downstream of the 18S rRNA gene of *Aedes aegypti*. This region probably corresponds to part of the internal transcribed spacer of this species. One of these repeated sequences is extremely A-T rich. Pavlakis *et al* [108] have proposed that A-T rich regions within the internal transcribed spacer (in this case between the 5S and 2S coding regions of *Drosophila* rDNA) may represent 'processing sites' essential for the maturation of pre-rRNA into the constituent rRNA chains of the ribosome. One feature of the 'processing site' which these authors describe is the potential to form a stem and loop structure. The repeated sequence within *Ae. aegypti* rDNA also has this potential. One of the possible stem and loop structures which could be formed is shown below.

```
Nucleotides in loop:          C A G T T C C T G C T T A
                                        T-A
                                        A-T
                                        T-A
Nucleotides in stem:                    A-T
                                        T-A
                                        A-T
                                        T-A
Nucleotide position in                -TAA TCA-
r35-mp8 Acc1/EcoR1 sequence:          62     93
```

The aim of the further studies of the ribosomal variant λAar35 was to determine the source of the length increase observed in the restriction fragment of this clone (1.8 Kb) relative to that of the normal rDNA repeat (0.8 Kb). Of particular interest in this study was the possible relationship between the extra DNA contained in λAar35 and a eukaryote TGE. The hybridisation properties of the variant region of λAar35 were examined to determine whether this region contains any non-rDNA. A comparison of the results obtained using the variant restriction fragment of λAar35 and the corresponding restriction fragment from a normal rDNA repeat to probe Southern transfers of *Ae.aegypti* genomic DNA and to screen the *Ae.aegypti* genomic library, show that the variant region of λAar35 does not contain a moderately repeated, chromosomally interspersed segment of non-rDNA.

The source of the length increase in clone λAar35 cannot be resolved from the data presented here. Fine detail restriction mapping and DNA sequencing of the corresponding region of a normal rDNA repeat is a prerequisite to the detailed analysis of the variant segment contained in λAar35. This omission also restricts other conclusions concerning the sequence data obtained from the variant rDNA. The construction and use of the oligonucleotide probe described in Section 3.18 was of limited value in view of the unknown nature or

source of the DNA for which the probe was constructed. Also, it should be noted that the sequence data of *Ae.aegypti* 18S rDNA used for comparison with *D.melanogaster* and *B.mori* rDNA (Section 3.16) is derived from a variant rDNA repeat and not a normal rDNA transcription unit.

EcoRI restricted *Ae. aegypti* rDNA gives a minor band of hybridisation signal on Southern transfers of 9.0 Kb. This may correspond to rDNA fragments produced from copies of λAar29, λAar35 or λAav1 cloned rDNA (8.3, 8.2 and 8.6 Kb respectively) in the *Aedes* genome.

Analysis of the hybridisation properties of the variant region of λAar29 (which contains an extra 1.2 Kb of DNA in the 28S rRNA coding region) is required to determine whether this variant was produced by an insertion or duplication event. Similarly, further studies on the structure of λAas1 and the distribution of sequences homologous to the non-rDNA contained in this clone are required to establish the nature of this variant.

This study of a particular repeated gene family is the first undertaken for a species of mosquito. Ribosomal RNA comprises a large proportion (e.g. 80-90%) of the total RNA of a eukaryotic cell. It is evident from the studies performed in Section I of this thesis, that it is of great importance to be able to distinguish rDNA clones from other highly expressed sequences. Section I describes a means of constructing an *Ae. aegypti* genomic library largely devoid of rDNA using the information obtained concerning the structure and organisation of the rDNA of this species. Such a library would facilitate differential screening for clones containing highly expressed, non-ribosomal genes at high plaque density.

Ribosomal gene copy number may be used as an internal standard by which to compare the repetition frequency of other sequences (see Section 4 of this thesis for example). When screening genomic libraries using complex probes (e.g. total genomic DNA), identification of rDNA clones allows an estimate of signal intensity expected for a given repetition frequency (see Section I).

The *Aedes aegypti* rDNA clone λAar7 is currently being used by R. Bonet of this Department as a marker to detect the region of the *Ae. aegypti* chromosme 3 containing the nucleolus organiser region in pulsed-field gel electrophoresis mapping experiments.

Ribosomal DNA represents 0.6% of the total genome of *Ae. aegypti*. Thus rDNA represents 0.6% of the potential 'target DNA' for a transposition event in this organism. The highly ordered and homogeneous nature of the rDNA tandem repeat makes this an excellent marker of a transposition event, as indicated by the generation of a variant rDNA segment. The other sources of rDNA polymorphism which are discussed in this section are not so frequent that they would mask such an event completely, although an understanding of the other sources of rDNA polymorphism is essential if rDNA variants are to be analysed in an effort to isolate transposable genetic elements (TGEs). A number of TGEs have been identified inserted into the rDNA of other species [92-94]. The two *Ae. aegypti* rDNA variants studied in detail here (λAav1 and λAar35) do not represent the products of an insertion of a TGE into rDNA based on the evidence available. The two other variants (λAas1 and λAar29) have yet to be studied to this extent.

Re-screening genomic clones for further rDNA variants by the methods described here will identify other interesting clones. These

may be distinguished from variants already analysed by running restriction enzyme digested DNA of new and previously mapped rDNA clones in parallel. Restriction mapping of new rDNA variants will be much simplified by the availability of the detailed restriction map of the normal rDNA repeat.

Identification of a sequence which is thought to represent a TGE (isolated using a different method, e.g. [15]) inserted into a ribosomal repeat will provide strong evidence that the sequence is transposable in origin. This could be accomplished by screening a large number of genomic clones in parallel with rDNA and the putative TGE as hybridisation probes. The identification of a TGE inserted into rDNA would greatly simplify the analysis of the TGE structure (e.g. size determination and definition of the ends of the element) due to the defined nature of the flanking rDNA regions compared with analysis of the TGE inserted into unique DNA.

The absence of sequence conservation between different species of a large proportion (e.g. 80%) of the rDNA repeat is of significance to the isolation of species-specific DNA probes (see Section 4). The extent of variation in sequence of some regions of rDNA, particularly the spacer regions, may be sufficient to facilitate the differentiation of very closely related and morphologically similar species. The repeated nature of such a species-specific probe (500 copies in the *Ae. aegypti* genome) would give good probe sensitivity compared to a unique sequence probe of the same length.

SECTION 4; DNA Probes for Species Identification of Mosquitoes in the *Anopheles gambiae* Complex

4.1. Introduction

A species complex is the term used to describe a group of distinct species which are morphologically very similar or identical. Many species complexes have been recognised within the insect order Diptera. Often the members of a species complex are only distinguished when a detailed study is made of what was formerly thought of as a single species. The scientific interest in Anopheline mosquitoes due to the medical importance of some species has led to the recognition of a number of *Anopheles* species complexes. Of these, the best studied is the *Anopheles gambiae* species complex. Other notable species complexes include the *An. culicifacies* complex (the major vectors of malaria in India), the *An. maculipennis* complex, formerly responsible for malaria transmission in Europe, the *An. punctulatus* complex of Australasia and the oriental *An. maculatus* and *An. balabacensis* complexes. Ecological and behavioural differences observed between different populations of mosquitoes currently called a single species may prove to be the result of as yet unrecognised species complexes [115].

Species with different ecology and behaviour frequently exhibit differing capacities to transmit disease. Species which readily feed on humans (are anthropophilic) and which readily rest and feed indoors (are endophilic) are much more efficient vectors of human disease than physiologically similar mosquitoes which are more

zoophilic and exophilic. Mosquito ecology and behaviour also determines the efficacy of various control measures against a particular species. For example, a species which often rests inside buildings after taking a blood-meal is much more susceptible to control using residual spraying of insecticides in houses than a physiologically similar species which may feed indoors, but then goes to rest outside.

Identification of species within a species complex is therefore a pre-requisite to understanding fully the ecological, behavioural and hence vectorial capacity variation within a species complex and the relation of these to the effective planning and assessment of vector control programmes. This section describes the *An. gambiae* species complex as it is known at present and surveys the methods currently available for species identification. The development of a novel method for species identification using DNA probes is then described and the possible application of this technique is discussed.

4.2. The *Anopheles gambiae* species complex

Mosquitoes of the *Anopheles gambiae* species complex include the major vectors of human malaria in tropical Africa and secondary vectors of bancroftian filariasis and some arboviruses. Although it is impossible to assess properly the morbidity caused by malaria in tropical Africa, one million infant deaths per year are thought to result directly from malaria. In addition to this is the debilitating effect of the disease on the more resistant adult population and the resulting loss in economic output.

Six different species within the *An. gambiae* complex have been identified on the basis of mating incompatibility [116-118]. Four are freshwater breeding species: *An. gambiae* sensu-stricto, *An. arabiensis*, *An. quadriannulatus* and *An. bwambae* (formerly known as species A, B, C and D respectively). Two species are usually restricted to saltwater breeding sites: *An. melas* and *An. merus*. The distribution and vectorial capacities of the six different species are summarised in Figure 4.2.

The first recognition of differences within what was then called *An. gambiae* came in 1944 when *An. melas* was differentiated from *An. gambiae* on the basis of small morphological differences [119]. *An. melas* was defined as a separate species by crossing experiments in 1948 [120]. Differences within freshwater breeding *An. gambiae* were discovered by chance in 1956 when male hybrid sterility was observed in a cross between an insecticide resistant *An. gambiae* isolate and an insecticide sensitive laboratory colony [121]. Extensive crossing experiments subsequently led to the identification of the six different *An. gambiae* species (see [122] for review).

Mosquitoes of the Anopheles gambiae Complex

SPECIES	DISTRIBUTION	MALARIAL VECTORIAL CAPACITY
An. gambiae sensu-stricto	Throughout Tropical Africa	Most efficient vector
An. arabiensis	Throughout Tropical Africa	Good vector
An. quadriannulatus	East and South Africa, and Ethiopia	Thought to be zoophilic not a vector
An. bwambae (formerly species 'D')	Semliki Forest of Uganda only	Local vector
An. melas	Coastal areas, W. Africa. (saltwater associated)	Vector in coastal regions
An. merus	Coastal areas and along some rivers, E. Africa (saltwater assoc.)	Inefficient vector

Fig. 4.2

An. gambiae s.s. and *An. arabiensis* are the most abundant and widespread of the *An. gambiae* species and are responsible for the majority of malaria transmission. Both species are found throughout tropical Africa with the exceptions that *An. gambiae* s.s. is not found in the arid regions of S. Arabia and the horn of Africa (where *An. arabiensis* is common) and *An. arabiensis* is absent from the more humid rain forests of W. Africa and the Congo basin where *An. gambiae* s.s. predominates. Where the two species do co-exhist, their relative abundance may vary considerably over time and by locality [115]. This is probably due to climatic variations, but may also involve differences in ecology within the two species. It is therefore important to assess mosquito populations and their ecology on a local basis, as generalisations concerning species distribution and behaviour may be misleading.

The infectivity of *An. gambiae* s.s. and *An. arabiensis* for malaria sporozoites and *Wuchereria bancrofti* microfilariae is usually several percent of the specimens examined. *An. gambiae* s.s. has been observed to develop particularly high infectivity rates in some cases. One report for an area in Cameroon describes a 60% infectivity rate for malaria sporozoites [123]. It is not clear whether *An. gambiae* s.s. is a physiologically more efficient vector than *An. arabiensis* or whether the areas where *An. gambiae* s.s. predominates (more humid in general) promote higher survival and thus infectivity of this species relative to *An. arabiensis* [124]. *An. gambiae* s.s. is generally considered to be more endophilic than *An. arabiensis* [125].

The use of residual insecticides in houses is more effective against *An. gambiae* s.s. than *An. arabiensis*. The effect of spraying

houses with the insecticide fenitrothion in the Kisumu area of Kenya resulted in an increase in the prevalence of *An. arabiensis* relative to *An. gambiae* s.s. and an overall increase in exophily of the mosquito population [126]. In the Garki area of northern Nigeria, outdoor resting by both *An. gambiae* s.s. and *An. arabiensis* may have contributed to the failure of this efficient house spraying campaign to significantly reduce malaria transmission [127]. Coluzzi *et al* [128] have reported significant correlation between intra-species inversion polymorphisms and the degree of exophily exhibited by *An. gambiae* s.s. and *An. arabiensis*.

An. arabiensis was the first of the *Anopheles gambiae* complex for which insecticide resistance was reported (to dieldrin in 1956 [129]). This species now shows very widespread resistance in W. Africa and some resistance in E. Africa and Madagascar. In addition, DDT resistance occurs in Sudan and Senegal. Resistance of *An. gambiae* s.s. was first observed in 1962 to dieldrin [130] and this is now widespread in many parts of W. Africa, Transvaal and Kenya. DDT resistance occurs in Burkino Faso and Togo.

An. quadriannulatus is limited in distribution to Zanzibar, the Ethiopian highlands and South Africa (the Zambesi, Sabi, Lundi and Limpopo valley systems especially). *An. quadriannulatus* is very zoophilic and exophilic and therefore not thought to be a vector of human malaria. In an area of Ethiopia where *An. arabiensis* and *An. quadriannulatus* are sympatric, only *An. arabiensis* was identified resting in houses with no resident livestock, whereas *An. quadriannulatus* made up a significant proportion of resting mosquitoes outdoors and in animal shelters. While the human blood indices of *An. arabiensis* was over 50%, that of *An. quadriannulatus*

was less than 1% [131]. No sporozoites were identified in *An. quadriannulatus* females. *An. quadriannulatus* appears to exhibit less intra-species inversion polymorphism and is correspondingly less ecologically versatile than the two more polymorphic freshwater species [122]. Although medically unimportant, correct identification of *An. quadriannulatus* when present is highly pertinent.

In several control programmes, the failure to erradicate *An. gambiae* by house-spraying actually represented simply a failure to erradicate the exophilic and zoophilic *An. quadriannnulatus* sibling species [132,133] and spraying was continued longer than was necessary. *An. quadriannulatus* thus represents an obstacle to malaria vector surveillance if not identified precisely.

An. bwambae is restricted in location to an area of the Semliki Forest in the county Bwamba, Uganda, where it breeds in mineral water swamps. *An. bwambae* is anthropophilic and endophilic and is therefore a local vector of malaria amongst the Bwamba pigmies [134]. Although ecologically interesting, the limited distribution of this species makes it a very minor vector.

An. melas is sympatric with *An. gambiae* s.s. and *An. arabiensis* on the West coast of Africa between 16°N and 13°S of the equator, where it breeds prolifically in saltwater mangrove swamps. Similarly, *An. merus* is sympatric with the two common freshwater species on the East coast of Africa between 1°S to 29°S of the equator. *An. merus* may also spread in distribution inland along rivers during the dry season and has been reported up to 120 Km inland in Swaziland [135]. *An. merus* is also present on the East coast islands of Pemba, Madigascar and Maritius.

Both saltwater breeding species are physiologically resistant to salinities of up to two times that of seawater and are also capable of breeding in freshwater. *An. melas* and *An. merus* have been shown to have a physiological capacity to sustain *Plasmodium falciparum* similar to that of *An. gambiae* s.s. and *An. arabiensis* [136]. In natural populations however, they do not develop such high infectivity rates. Studies on *An. merus* on the Kenyan coast show that in this area at the time of study, *An. merus* was a vector of secondary importance to *An. gambiae* s.s. and *An. arabiensis* [136]. Additionally, no inversion polymorphism was observed in *An. merus* (774 specimens examined). Similar studies at a location on the Tanzanian coast reported that although *An. merus* was more exophilic than *An. gambiae* s.s., *An. merus* was an important vector of *Wucheria bancrofti* [137]. Bryan [138,139] has demonstrated that in the Gambia, *An. melas* has a consistantly much lower sporozoite rate than *An. gambiae* s.s. (approximately one tenth). Possible reasons for this may be the shorter lifespan of *An. melas* relative to *An. gambiae* s.s. (as judged by parous rates of biting females) and a higher degree of zoophily exhibited by *An. melas*. It is suggested that *An. melas* alone may be incapable of maintaining malaria transmission in the area studied.

Crosses between different *An. gambiae* species may be performed easily in laboratory colonies. The male FI offspring of such crosses are sterile and the extent of sterility varies between different combinations of species [117]. In natural populations, sterile hybrid males are extremely rare [140], even in areas where two or more *An. gambiae* species are sympatric (in S. Mozambique, three freshwater breeding species and *An. merus* co-exhist). Hybrids

between *An. gambiae* species have been identified on the basis of polytene chromosome banding patterns at a very low frequency. In two studies of mosquitoes in the Gambia where *An. melas* and *An. gambiae* s.s. are very common, Bryan [138,139] identified hybrids representing less than 0.01% of the specimens examined. White [141] reported a higher percentage (0.15%) of hybrids between *An. gambiae* s.s. and *An. arabiensis* in Tanzania. These results infer that in natural populations which breed eurygamously, pre-copulatory mating barriers exhist between species (i.e. there is assortive mating) and that these barriers break down readily in laboratory breeding (stenogamous) colonies. There is evidence against genetic exchange between *An. gambiae* s.s. and *An. melas* in areas where *An. gambiae* s.s. has become resistant to insecticides, but *An. melas* remains insecticide susceptible [122].

Comparisons of vectorial capacity of freshwater *An. gambiae* and *An. melas* indicate that the different species have similar physiological susceptibilities to infection by *P. falciparum* [142]. However, *An. gambiae* exhibits a higher human biting rate and develops a larger number of oocysts than *An. melas* and may survive longer after infection [143]. *An. gambiae* is therefore a more efficient vector than *An. melas* and usually exhibits a higher sporozoite rate [144].

4.3. Species identification

The definitive method for species identification of mosquitoes in the *Anopheles gambiae* complex is to perform crossing experiments of the offspring of the unknown specimen against laboratory colonies of known species identity [117]. Although this method has been of great value in the elucidation of species within the *An. gambiae* complex, it is far too laborious and time consuming to use for routine identification of individuals.

Many external morphological characteristics of *An. gambiae* species have been examined in an attempt to identify species-specific features. Although some characteristics of *An. melas* are fairly reliable for separation of this species from freshwater *An. gambiae* [119,120], no reliable features have been found to allow identification of the other species.

The most reliable and widely used method for species identification is that of the polytene chromosome banding technique [145-149]. This allows the separation of the three freshwater species *An. gambiae* s.s., *An. arabiensis* and *An. quadriannulatus* by diagnostic 'X' chromosome inversions and identification of the other species using more subtle autosomal features. The main disadvantages of this method are that identification is limited to fourth instar larvae or semi-gravid blood-fed females and that the technique reqiures skill to perform. The number of specimens which may be processed easily is also limited.

Isoenzyme typing [150-152] is also used routinely in some laboratories for species identification of *An. gambiae*. This method relies on different species having differing electromorphs of a

number of enzymes. Starch gel electrophoresis is usually used and in some cases, each gel is split to allow the assay of two different enzymes. This technique is time consuming and relatively expensive in terms of equipment and materials. The number of specimens which may be processed individually is also limited and some specimens give anomalous results. Specimens must be stored frozen in liquid nitrogen prior to identification due to the lability of the enzymes tested for.

The use of cuticular and internal hydrocarbons to identify *An. gambiae* species has been investigated [153-155]. In this technique hydrocarbons are extracted using an organic solvent and analysed by gas chromatography. This method has not been fully evaluated, but the requirement for sophisticated equipment to obtain and handle data, together with the time consuming nature of identification by this procedure make this an unlikely technique for routine species identification.

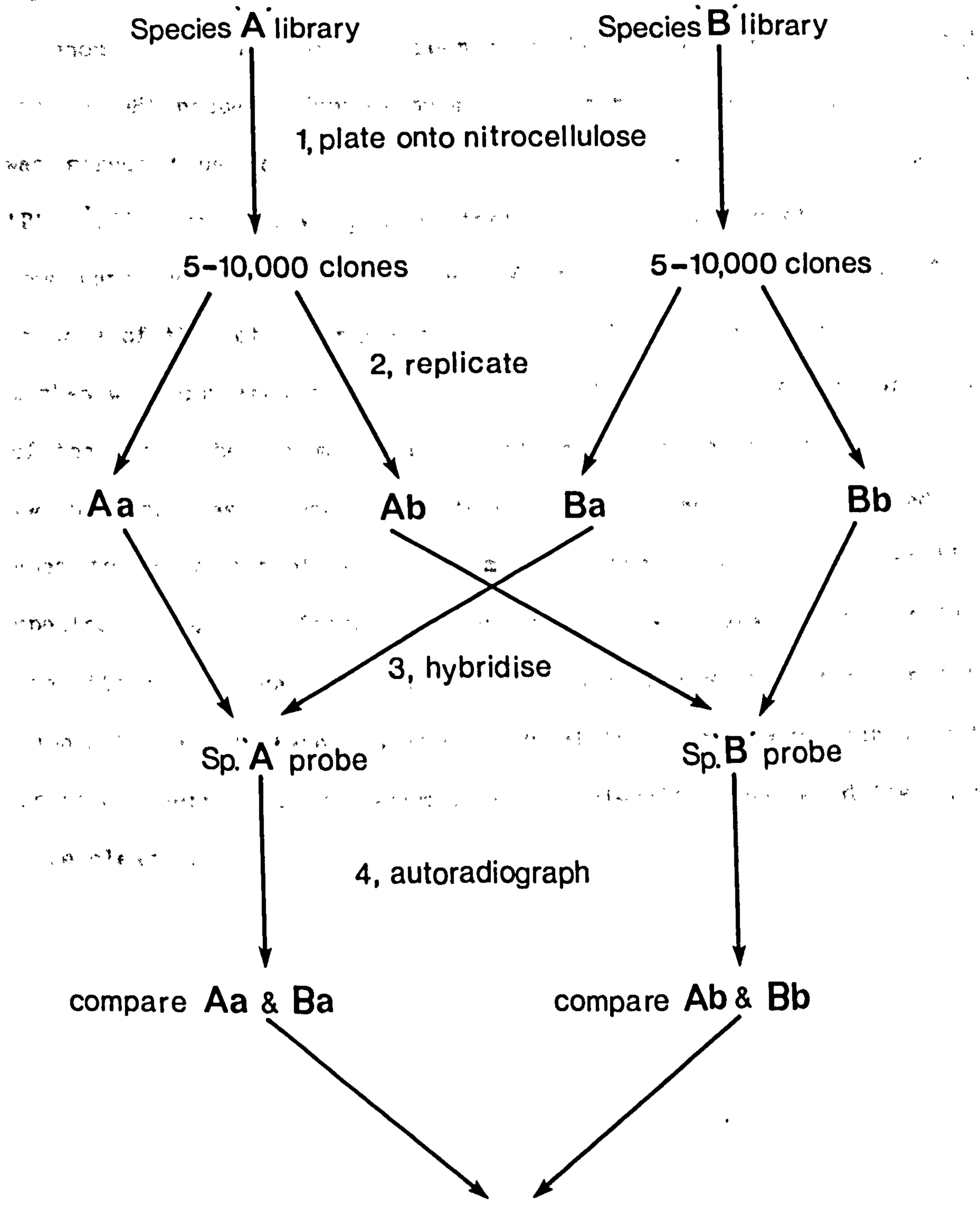
Some success in differentiating *Anopheles gambiae* sibling species using the fluorescent stain Hoechst 33258 has been described [156,157]. This stain preferentially stains A-T rich regions in DNA and so allows the position of heterochromatic blocks of DNA to be visualised on mitotic chromosomes. Unfortunately, in addition to major differences in banding patterns between species, a significant level of intraspecies variation is also evident. This greatly reduces the potential of this method for routine species identification.

4.4. Strategy for the isolation of species-specific DNA sequences (see Figure 4.4)

DNA sequences encoding functional proteins or structural RNAs are usually very highly conserved throughout evolution. Non-coding DNA sequences exhibit a much larger degree of variation however. Regions within ribosomal RNA coding sequences are highly conserved whereas non-coding spacer sequences are not evolutioarily conserved [67]. This reflects the evolutionary selection against most changes which occur in functional proteins and RNA species. A particularly variable component of eukaryotic genomes with respect to DNA sequence is the highly repeated DNA fraction [158]. It was therefore decided that a comparison of highly repeated DNA sequences between the different members of the *An. gambiae* complex would be the most efficient strategy for the isolation of species-specific DNA sequences. This approach also has the advantage that any species-specific DNA sequences which are isolated will be highly reiterated in at least some of the species studied and would therefore provide sensitive hybridisation probes. An alternative strategy to that described here would involve the study of a particular DNA species to identify a variable portion suitable for use as a species-specific DNA probe. This approach has been used to isolate species-specific sequences from the kinetoplast DNA of *Leishmania* species [159].

Genomic libraries of the four species available (*An. gambiae* s.s., *An. arabiensis*, *An. melas* and *An. merus*) were constructed in the plasmid vector pBR328. Restriction endonuclease EcoRI limit-digested genomic DNA was used as insert. This gave a wide range of insert sizes, but most of the inserts were less than 5 Kb in length.

Fig. 4.4 Strategy used to isolate species-specific DNA probes



Short inserts are advantageous in the strategy employed here because large genomic fragments may contain several short repeated sequences which may mask the hybridisation characteristics of other sequences (see Section 1).

Genomic libraries were screened in duplicate using two different genomic DNA probes. For example, the genomic library of species 'A' was screened using total genomic DNA from species 'A' and species 'B'. All clones showing hybridisation signal must contain at least some repeated DNA sequence, (see Section 1) which comprises 0.1-0.5% or more of the total genomic probe [48]. The autoradiographs for two probes were compared to identify clones which hybridise with only one of the two probes or show large differences in intensity between the two probes. These clones were picked and the plasmid DNA prepared and used to probe dot-blots of genomic DNA from the different sibling species. This confirmed whether the clone contained a species-specific repetitive DNA sequence or a sequence which varied in copy number between different species. Promising clones were then used to probe Southern blots of genomic DNAs to visualise species differences more clearly.

4.5. List of laboratory stocks and wild-caught materials

Material from a number of laboratory stocks was obtained from the London School of Tropical Medicine and Hygiene. These were supplied by C. Malcolm, R.G. Bonnington and B. Sawyer, as adults from laboratory colonies which had been checked by cytotyping.

An. gambiae s.s. laboratory stocks

NAME	DATE COLONISED	ORIGIN
1. PALA	1963	Pala, Burkino Faso
2. G3	1975	McCarthy Island, The Gambia
3. 16CSS	1951	Lagos, Nigeria
4. KWA	1975	Kwale, N. of Tanga, Tanzania
5. BAD	1980	Badagry, Lagos, Nigeria
6. REFMA	1979	From cross of LD (Sokoto, Nigeria) to 16CSS
7. Z11	1984	DDT resistant line selected from ZANU
8. KIL	1975	Marangu, Kilimanjaro, Tanzania
9. ZANU	1982	Zanzibar, Tanzania
10. IAN P20	1975	Iwaro, Nigeria
11. TOGO	1974	Togo

An. arabiensis laboratory stocks

NAME	DATE COLONISED	ORIGIN
1. SENN	Unknown	Sennar, Gezira, Sudan
2. POP	1985	Popo, Zanzibar, Tanzania
3. G-MAL	1978	Gezira, Sudan

An. melas laboratory stock

BREFET	1982	Brefet, Gambia

An. merus laboratory stock

ZULU	Supplied from colony held at South.

Wild caught material

Wild caught material consisted of cytotyped adults or larvae reared from cytotyped adults, preserved in isopropanol. Material was collected from the following areas in Zambia by M.W. Service of this department; Ndola (March 1985), Mushili (March 1985), Kafue (April 1985) and Kampumba (March 1985). All material was identified as *An. arabiensis*. A number of female *An. arabiensis* were also collected from the Mwea irrigation scheme (May 1985), Mgita Point (August 1985) and Kisumu (August 1985) in Kenya.

4.6. Construction and screening of *An. gambiae* s.s. and *An. arabiensis* genomic libraries

Genomic libraries of *An. gambiae* s.s. (PALA) and *An. arabiensis* (SENN) were constructed as follows. For each library, 2 µg of EcoRI digested genomic DNA was ligated with 2 µg of EcoRI-cleaved, dephosphorylated pBR328 DNA. Ligations were performed overnight at 37°C using two units of T4 DNA ligase in a 10 µl reaction volume. One quarter of this reaction was then used to transform competent *E. coli* MC1060 (see Materials and Methods). Approximately 5000 colonies were obtained for each library when the transformations were plated onto tetracycline selective agar (12.5 µg/ml). The colonies obtained were washed off into L-broth. The cells were then pelleted by centrifugation and resuspended in 1.2 ml of 50% glycerol/50% L-broth. These glycerol stocks were then stored at -20°C. The titre of these stocks were as follows (11/1/85).

Species	Titre on tetracycline	Titre on chloramphenicol
An. gambiae s.s.	3.45×10^{10} per ml	1×10^{8} per ml
An. arabiensis	3.30×10^{10} per ml	7×10^{7} per ml

Insertion of genomic DNA at the EcoRI site of pBR328 inactivates the chloramphenicol resistance gene. Therefore the background of non-recombinants in the two libraries is 0.3% and 0.2% respectively. These values are approximate due to the possibilities of small, in-frame inserts allowing read-through and thus (altered) chloramphenicol acetyl transferase (CAT) production. Also, re-

ligation of vector DNA with damaged ends (e.g. due to nuclease activity) may cause inactivation of the CAT gene in some plasmids which do not contain inserts.

The remaining ligation reactions were stored at -20°C and represent a potential further 15,000 recombinants.

Total genomic DNA from each of the two species was nick-translated to a specific activity of 1-2 x 10^8 dpm/µg. Each of these probes was then used to probe replica filters of 500-1000 clones from each library (see Materials and Methods) on 9 cm diameter nitrocellulose discs. The autoradiographs from each set of replica filters were compared and any clones showing differential hybridisation were picked. The autoradiographs from one differential screening experiment are shown in Figure 4.6. Typically, six of the clones showing the most obvious differences were picked from each plate. These clones were then streaked out to obtain single colonies and one colony from each was used to inoculate 20 ml cultures for plasmid isolation (see Materials and Methods).

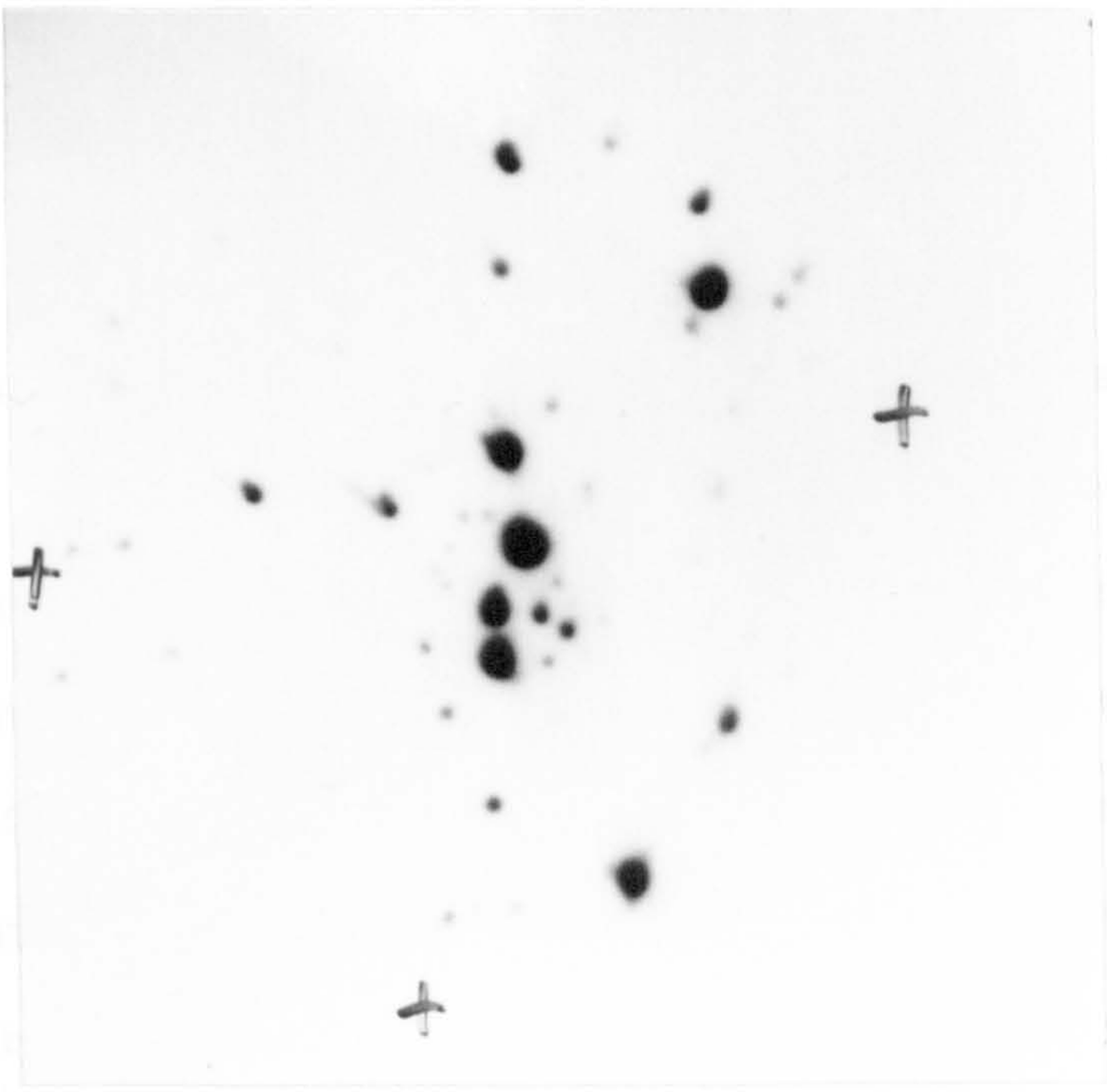
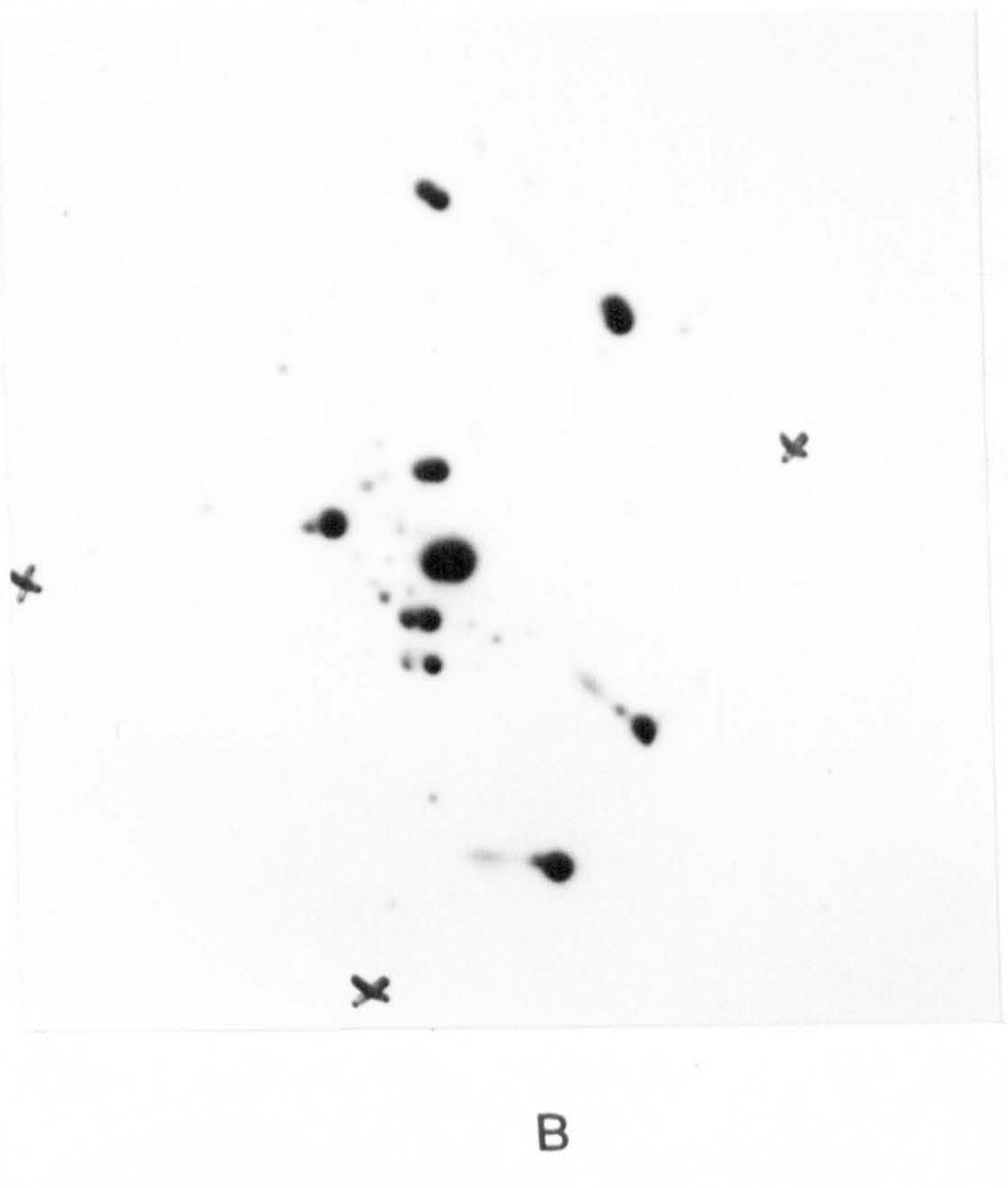


Fig. 4.6 A



B

Figure 4.6; Differetial screening of approximately 500 clones from the *An.arabiensis* genomic library using A; *An.gambiae* s.s. and B; *An.arabiensis* total genomic DNAs as hybridisation probes.

4.7. Dot-blots: testing of putative species specific sequences

Plasmid DNAs from clones containing putative species-specific DNA sequences were nick-translated and used to probe dot-blots of genomic DNA from various *An. gambiae* species. Dot-blots were prepared using a vacuum operated apparatus (Sleicher and Schuell "Minifold"). For each sample, 100 ng of DNA in 20 µl of T.E. buffer was denatured by boiling for three minutes, combined with 0.5 ml of 6 x SSC, applied to the filter and rinsed in with 1 ml of 6 x SSC. Filters were baked (2 hours, 80°C) prior to hybridisation. Controls were performed using pBR328 vector DNA only to ensure that no hybridisation occurred with this probe.

Of the clones originally picked in the differential screening experiments, less than half were confirmed as showing species-specific distribution. A total of five clones were isolated from the *An. arabiensis* genomic library which showed strong hybridisation to *An. arabiensis* genomic DNA, but no hybridisation to *An. gambiae* s.s. genomic DNA.

4.8. Hybridisation characteristics of *An. arabiensis* species-specific sequences

Plasmid DNA from the five clones which hybridised to *An. arabiensis* genomic DNA, but not *An. gambiae* s.s. were nick-translated and used to probe dot-blots of all available laboratory stocks and wild-caught material.

The results for one clone (designated pAnaFI) are summarised in a simplified form in Figure 4.8. The other clones (pAna F3, A6, G2 and E3) showed identical hybridisation characteristics. No hybridisation was detected to any of the eleven different laboratory stocks of *An. gambiae* s.s. and strong hybridisation was observed to the *An. merus* laboratory stock, whereas only weak hybridisation signal was observed with *An. melas* genomic DNA. An important finding was that when DNA from *An. gambiae* s.s. and *An. arabiensis* individual males and females was probed (all laboratory stock DNA is usually prepared from several hundred pooled larvae), hybridisation signal was only observed to DNA derived from male *An. arabiensis*. Material was not available to test for sex-specificity of this sequence in *An. merus* and *An. melas* due to the loss of these laboratory stocks. Figure 4.8 also shows the result of screening a replica filter with mosquito ribosomal DNA (λAar7 insert DNA, representing the entire rDNA repeat from *Aedes aegypti*, see Section 2 of this thesis). This serves as a control for normalising DNA loadings. No hybridisation signal was evident for any of the DNA samples when a replica filter was probed using pBR328 vector DNA only.

The DNA probes pAna FI, F3, A6, G2 and E3 clearly distinguish male *An. arabiensis* and male *An. gambiae* s.s. individuals. *An.*

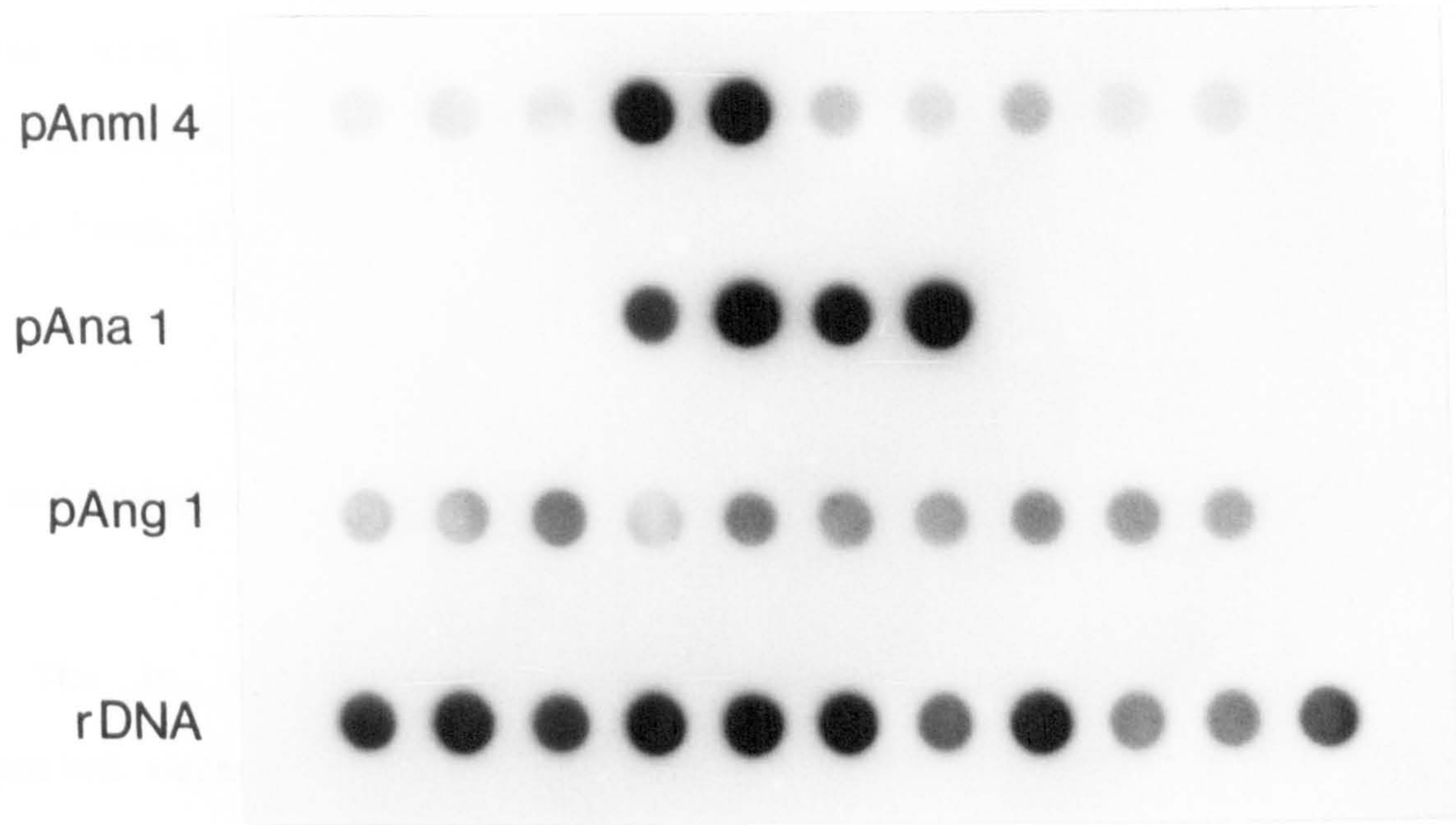


Fig.4.8

Figure 4.8; Hybridisation characteristics of clones containing species specific DNA sequences. 100ng of genomic DNA from each laboratory stock or from individual adults (where sex is indicated) was used to prepare replica dot-blots. One replica was then hybridised with each of the probes shown. The rDNA probe used for normalising DNA quantities was the insert DNA from clone λAar7 which contains the intact rDNA repeat from *Aedes aegypti* (see section II). G1-G4; *An.gambiae* s.s. IAN P20, KWA, PALA and G3 laboratory stocks respectively, Ml; *An.melas* (BREFET), Mr; *An.merus* (ZULU), A1; *An.arabiensis* (POP), A2; *An.arabiensis* (SENN), S; *An.stephensi*.

gambiae s.s. is the only species that shows zero hybridisation signal with these probes and so males of this species may therefore be distinguished from males of the other three species tested. Similarly, *An. melas* is the only species to show weak hybridisation signal with these probes and so may be distinguished from the other species tested. *An. arabiensis* and *An. merus* cannot be distinguished using these DNA probes.

Re-screening *An. gambiae* s.s. and *An. arabiensis* genomic libraries

The *An. gambiae* s.s./*An. arabiensis* differential screening was repeated using genomic DNA probes prepared from female adults only. This excluded the male-specific repetitive sequences from *An. arabiensis* genomic DNA. Although a further 2000 recombinants were screened from each genomic library, no species-specific probes to distinguish the females of these two species were obtained.

An. gambiae specific sequences

The majority of repetitive sequence clones from each genomic library did not show significant differences in hybridisation signal when screened with genomic DNA from different *An. gambiae* species. Five of these clones were picked. Figure 4.8 shows the hybridisation characteristics of one of these clones (designated pAng1). Hybridisation signals are evident for all *An. gambiae* species and no differences are apparent between sexes. No hybridisation signal is

observed for *An. stephensi* DNA. This type of DNA probe may therefore serve as a means of checking that all mosquitoes analysed are *An. gambiae*.

4.9. Construction and screening of *An. melas* and *An. merus* genomic libraries

Aim

To isolate DNA probes to distinguish *An. arabiensis* from *An. merus* (not possible using probes described in Section 4.8) and to improve the resolution of *An. arabiensis* and *An. melas*.

An. melas and *An. merus* genomic libraries were constructed as follows. 1 µg of EcoRI digested genomic DNA was ligated with 2 µg of EcoRI cut, dephosphorylated pBR328 vector DNA in a reaction volume of 10 µl. Ligations were performed overnight at room temperature. The reaction mixture was then used to transform competent *E. coli* MC1060. Approximately 2000 transformants were obtained for each species when the transformation reactions were plated onto tetracycline agar. Glycerol stocks were prepared as described in Section 4.6. These gave the following titres (23/11/85).

Species	Titre on tetracycline	Titre on chloramphenicol
An. melas	10^{10} per ml	3×10^{6} per ml
An. merus	8.5×10^{9} per ml	3×10^{6} per ml

Backgrounds of non-recombinants were therefore approximately 0.03% (*An. melas*) and 0.035% (*An. merus*).

Approximately 1000 recombinants from the *An. melas* genomic library were screened in duplicate using 'a', *An. melas* genomic DNA and 'b', *An. arabiensis* genomic DNA. Similarly, the *An. merus* genomic library was screened using *An. merus* and *An. arabiensis* genomic DNAs.

Two clones were isolated from the *An. melas* genomic library which showed lower hybridisation signal with *An. arabiensis* DNA than *An. melas* DNA (designated pAnml 2 and pAnml 4). Also, two clones showed weak hybridisation signal with *An. melas* genomic DNA and strong hybridisation signal with *An. arabiensis* DNA. These were designated pAnml 1 and pAnml 5.

Three clones from the *An. merus* genomic library were isolated which gave a strong hybridisation signal with *An. merus* genomic DNA and weak hybridisation signal with *An. arabiensis* genomic DNA. These were designated pAnmr 1, 2 and 3.

4.10. Hybridisation characteristics of clones isolated from *An. melas* and *An. merus* differential screening

The clones pAnml 2, pAnml 4 and pAnmr 1, pAnmr 2 and pAnmr 3 all showed identical hybridisation characteristics when nick-translated and used to probe dot-blots of genomic DNA from the four *An. gambiae* species. The results for one of these clones (pAnml 4) are shown in Figure 4.8. This clone shows weak hybridisation signal for all *An. arabiensis* and *An. gambiae* s.s. laboratory stocks (in total, 3 different *An. arabiensis* and 11 different *An. gambiae* s.s. stocks were tested) and strong hybridisation to *An. melas* and *An. merus* genomic DNA. No difference in hybridisation signal was observed between the male and female *An. arabiensis* and *An. gambiae* s.s. genomic DNAs. Material was not available to test for sex-specificity in *An. melas* and *An. merus* due to the loss of these laboratory stocks.

The two clones pAnml 1 and pAnml 5 showed hybridisation characteristics identical to those of pAnaF1 and related probes (shown in Figure 4.8 for pAnaFI).

The DNA probes pAnml 2, pAnml 4, and pAnmr 1, 2 and 3 clearly distinguish the two saltwater species *An. melas* and *An. merus* from the two freshwater species *An. arabiensis* and *An. gambiae* s.s. The distinction of *An. arabiensis* and *An. melas* is better using these probes than with the pAnaFI - type probes described in Section 4.8.

4.11. Probe Sensitivities

To determine the relative sensitivities of one of each type of species-specific DNA probe, each was compared with the sensitivity of a ribosomal DNA (rDNA) probe. The ribosomal DNA used was a 7.2 Kb EcoRI fragment from the clone λAar1 (see Section 2 of this thesis) which corresponds to the entire 28S rRNA gene and non-transcribed spacer region of *Aedes aegypti*. Although no information is available concerning the genome complexity and rDNA copy number in *Anopheles gambiae*, the rDNA probe used serves as a standard against which the other probes may be compared. In addition, the evolutionarily conserved 28S gene (4 Kb in length) represents a known middle-repetitive sequence class against which other sequences may be compared.

Genomic DNA was serially diluted and used to prepare replica dot-blots as described previously. Cloned DNA of a species-specific sequence and the rDNA fragment were nick-translated to the same specific activity. This was achieved by using identical probe DNA concentrations in identical nick-translation reactions. Reaction products were monitored to confirm equal specific activities. A replica dot-blot was then hybridised with each of the probes at equal probe concentrations (50 ng per ml) as described in Materials and Methods). Filters were washed in 0.1 x SSC at room temperature. Autoradiographs of three experiments using pAnaF1 to detect *An. arabiensis* genomic DNA (mixed sexes), pAnml 4 to detect *An. melas* genomic DNA and pAng 1 and pAnml 4 to detect *An. arabiensis* DNA are shown in Figure 4.11'a'.

Fig.4.11a

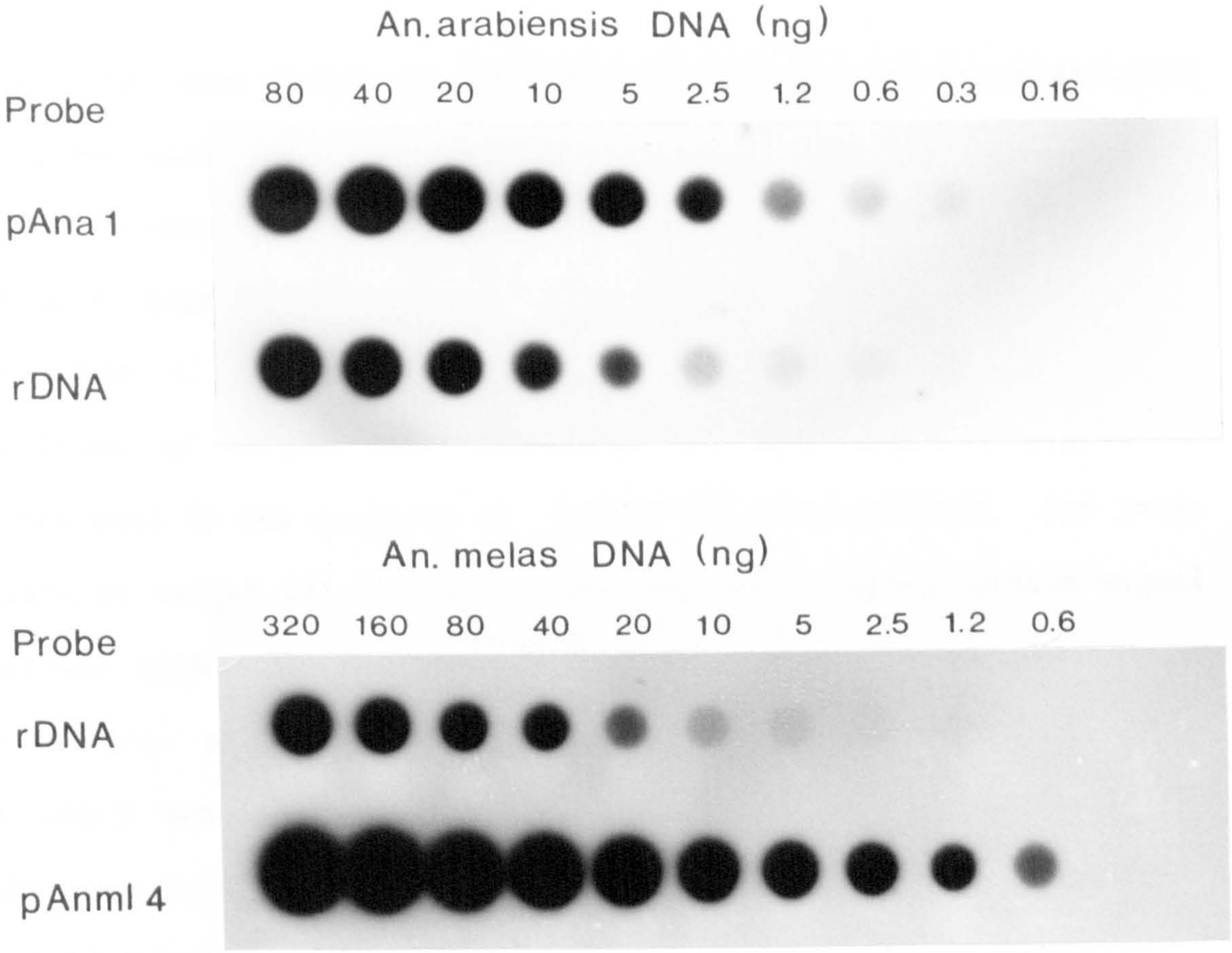


Figure 4.11a; Comparison of the sensitivities of species specific DNA probes. Replicates of serially diluted *An. arabiensis* or *An. melas* genomic DNAs were probed using species specific clone DNA and ribosomal DNA of equal specific activities.

To quantitate levels of hybridisation signal, the dot-blots were cut up and each dot counted in a scintillation counter. A background reading corresponding to the signal from a dot containing no DNA was subtracted from each value. These results are presented graphically in Figures. 4.11 'b', 'c' and 'd'. These graphs show that for low quantities of target DNA, the level of hybridisation signal is proportional to the quantity of target DNA on the filter. For large amounts of target DNA (80 ng per dot and above) hybridisation signal increases less with increasing target sequence. This may be due to two factors: probe depletion (the hybridisations were performed in stationary sealed plastic bags) and saturation of target DNA with probe (due to steric considerations not all target DNA is available for hybridisation at high concentrations of filter-bound DNA).

The gradients of the linear sections of each curve were calculated using linear regression computer analysis. This value is proportional to the sensitivity of the hybridisation probe. The relative sensitivities of the different clones as hybridisation probes (expressed as a proportion of the rDNA probe sensitivity) are shown below.

Fig.4 .11b; pAna1 probe sensitivity against An. arabiensis DNA

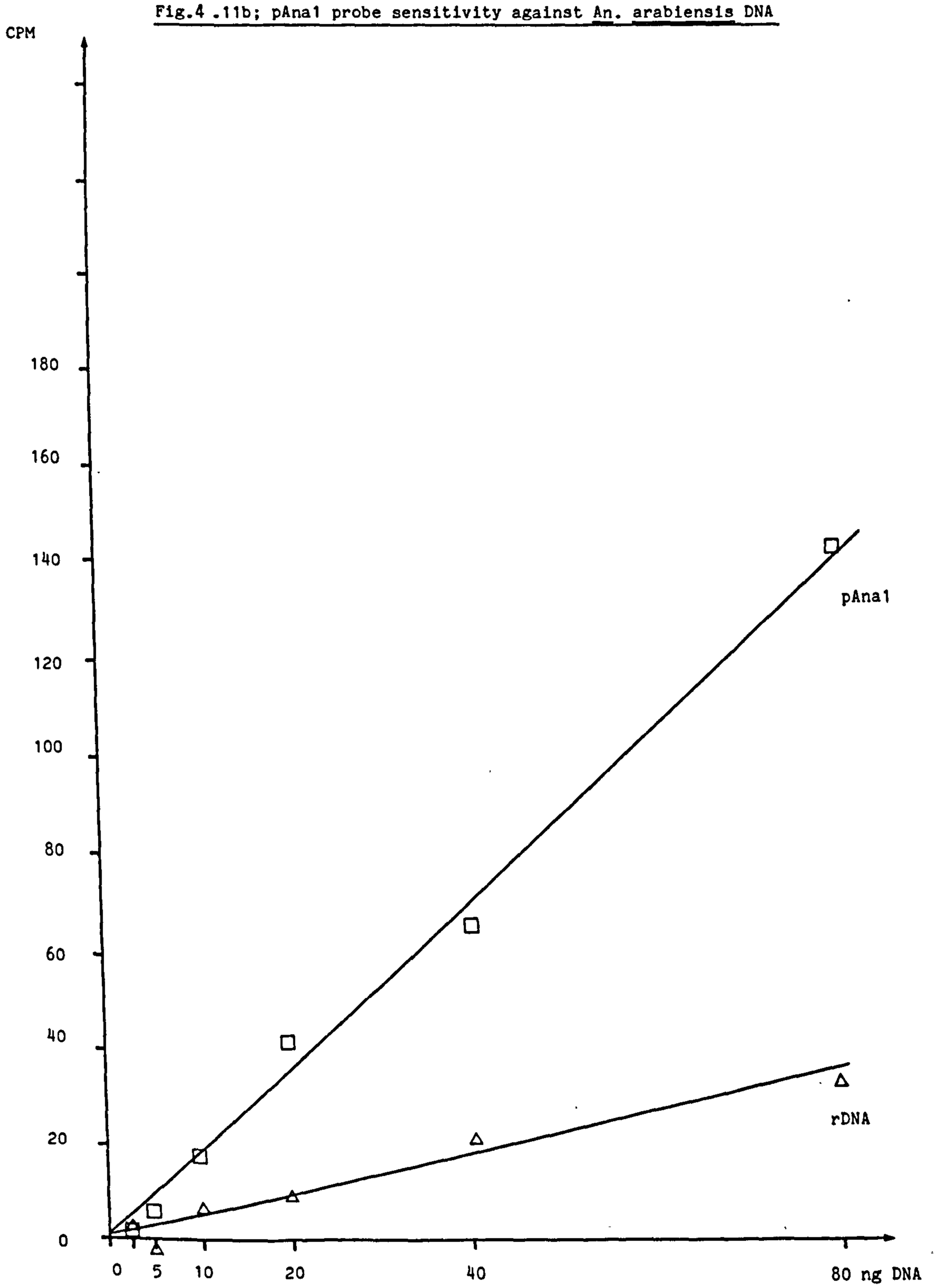


Fig.4 .11c; pAnml 4 probe sensitivity against An. melas DNA

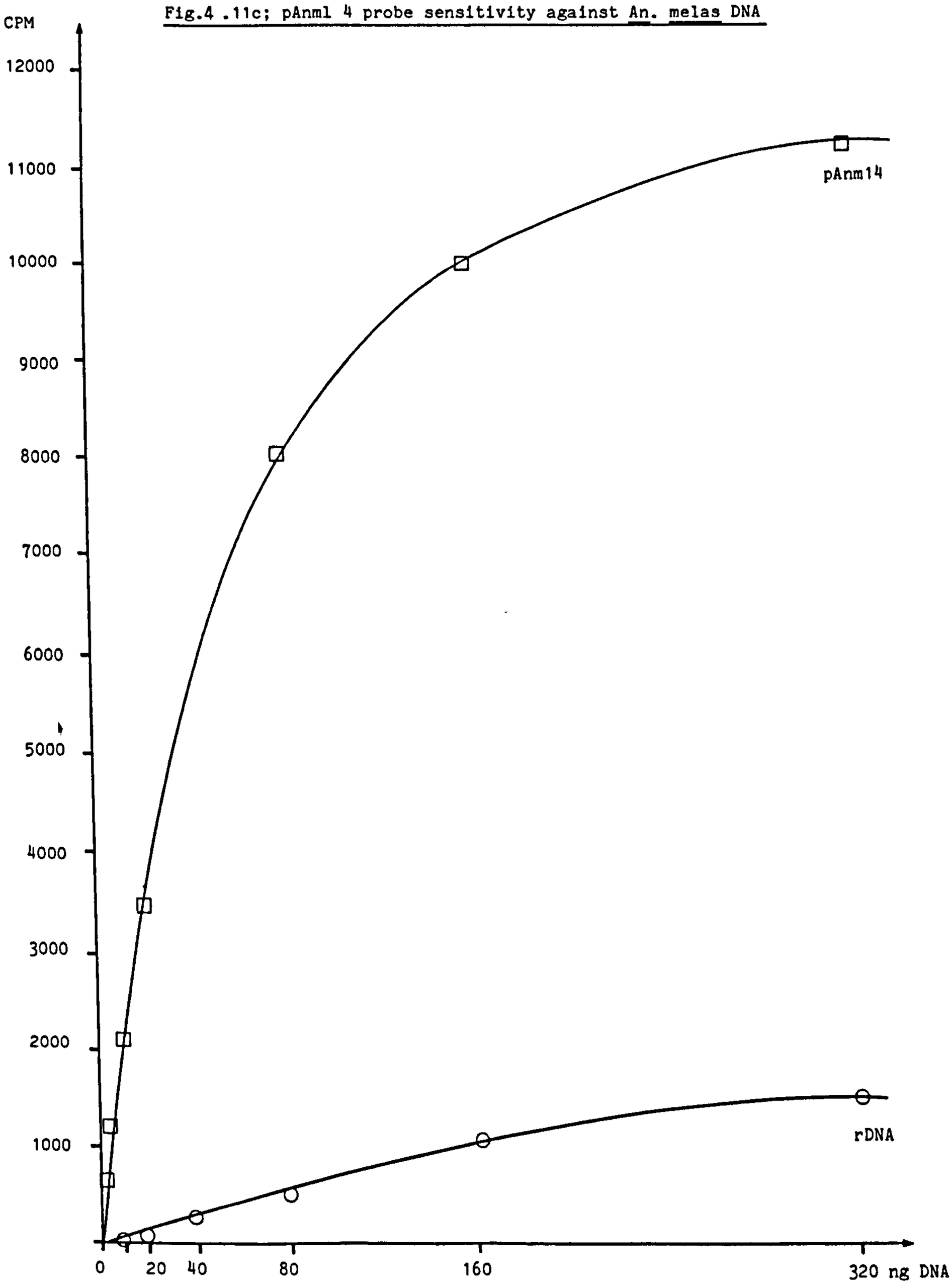
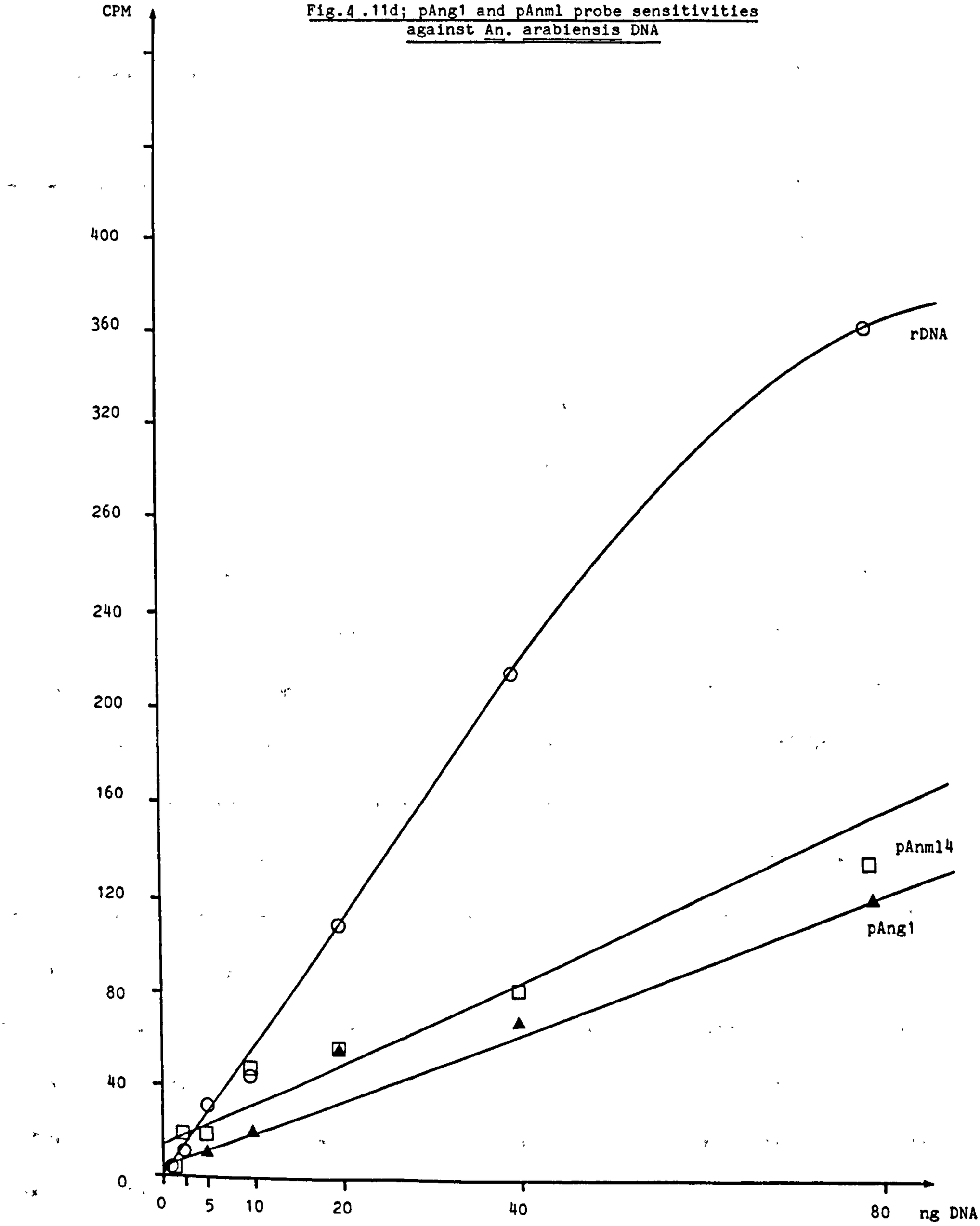


Fig.4 .11d; pAng1 and pAnm1 probe sensitivities against An. arabiensis DNA



Experiment Target / Probe	Initial gradient	Correlation coefficient	Relative sensitivity
An. arabiensis/pAnaFI	1.78	0.996	4.1
An. arabiensis/rDNA	0.43	0.971	1.0
An. melas/pAnml 4	174	0.993	27
An. melas/rDNA	6.4	0.999	1.0
An. arabiensis/pAng 1	1.59	0.9641	0.34
An. arabiensis/pAnml 4	1.39	0.9410	0.30
An. arabiensis/rDNA	4.60	0.9954	1.0

Clearly the most sensitive probe is pAnml 4 when used to detect *An. melas* genomic DNA. Sequences homologous to this probe are highly repeated in the genomes of *An. melas* and *An. merus* and present in low to middle copy number in *An. gambiae* s.s. and *An. arabiensis* (as determined by previous dot-blotting experiments). The data here show that pAnml 4 detects *An. melas* DNA approximately 90 times more sensitively than *An. arabiensis* DNA (i.e. pAnml 4 homologues differ in abundance by a factor of 90 between the two species). The probe pAnaFI detects mixed genomic DNA of *An. arabiensis* with a sensitivity of approximately four times that of rDNA. Hence the sensitivity of detection of male *An. arabiensis* DNA should be approximately 8 times that of rDNA.

The probe pAng 1 is only one third as sensitive as rDNA for the detection of *An. arabiensis* genomic DNA. This is therefore a low to middle abundance sequence.

4.12. Interrelation of cloned sequences

The clones isolated and characterised in Sections 4.6 - 4.10 were investigated for cross-hybridisation of insert DNAs. Clone DNA (0.5 µg) was digested with EcoRI and electrophoresed through 1.0% Agarose to separate insert and vector DNA. Three replica gels were prepared containing digested DNA from all clones. These were then blotted and probed with insert DNA from either pAnaF1, pAnml 4 or pAng 1. (Insert DNA from these clones was extracted from low-melting Agarose as described in Materials and Methods). This demonstrated that all clones with hybridisation characteristics similar to pAnaFI contained inserts homologous to insert DNA from this clone. Similarly, all pAnml 4 - type clones contained inserts homologous to pAnml 4 insert DNA. Of five clones selected which hybridised to all *An. gambiae* species, three showed homology to pAng 1 insert DNA (including pAng 1). No cross-hybridisation was evident between the three types of cloned sequence. (Results not shown). The insert sizes of the clones tested are listed below.

1. **pAnaFI-homologous clones;**

pAnaFI	0.7 Kb
pAnaF3	1.30 Kb
pAnaA6	0.45 + 0.55 Kb (2 inserts)
pAnaG2	2.50 Kb
pAnaE3	0.45 Kb

2. **pAnml 4-homologous clones;**

pAnml 2	0.4 Kb
pAnml 4	1.6 Kb
pAnmr 1	0.6 Kb
pAnmr 2	2.4 Kb
pAnmr 3	1.8 Kb

3. **pAng 1-homologous clones;**

pAng1	2.0 Kb
pAngE1	4.2 Kb
pAngG1	6.5 Kb

4.13. Genomic organisation of pAaFI-homologous sequences

To investigate the distribution of sequences homologous to pAnaFI in the *Anopheles* genome, clone DNA was used to probe a Southern transfer of various *An. gambiae* genomic DNAs (4 µg per lane) digested with the restriction endonuclease EcoRI. Figure 4.13a shows the ethidium bromide stained gel and Figure 4.13b shows the Southern transfer probed with pAnaF1 clone DNA. pAnaF1 hybridises strongly to the whole digest for the species *An. arabiensis* and *An. merus*, with some bands just visible superimposed on this. *An. melas* shows five bands of hybridisation signal with this probe (5.1, 4.1, 3.9, 2.9 and 2.7 Kb). This confirms that sequences homologous to pAnaFI are present in low copy number in the genome of this species. No hybridisation signal is observed to DNA from any of the *An. gambiae* s.s laboratory stocks.

The pattern of hybridisation to *An. arabiensis* genomic DNA when the experiment was repeated using different restriction endonucleases was similar to that obtained using EcoRI (results not shown). To determine whether this type of hybridisation pattern was due to the fact that the genomic DNA used for the Southern transfer had been prepared from several hundred pooled larvae, the experiment was repeated using individual adults. Figure 4.13d shows the result of probing a Southern transfer of EcoRI digested genomic DNA from six male and six female *An. arabiensis* with pAnaFI clone DNA. This clearly demonstrates that hybridisation occurs in the same form as for pooled genomic DNA and confirms the sex-specificity of pAnaFI-homologous sequences.

Figure 4.13 A-C; Organisation of sequences homologous to pAnaF1 and pAnml4 in the genomes of *An.gambiae* s.s., *An.melas*, *An.merus* and *An.arabiensis*. DNA (4µg) from each species was digested with EcoRI and size fractionated through 0.8% Agarose. Fig. 4.13'A' shows the ethidium bromide stained gel, 'B'; the corresponding Southern transfer probed with pAnaFI and 'C'; the corresponding Southern transfer re-probed with pAnml4.

Fig. 4.13

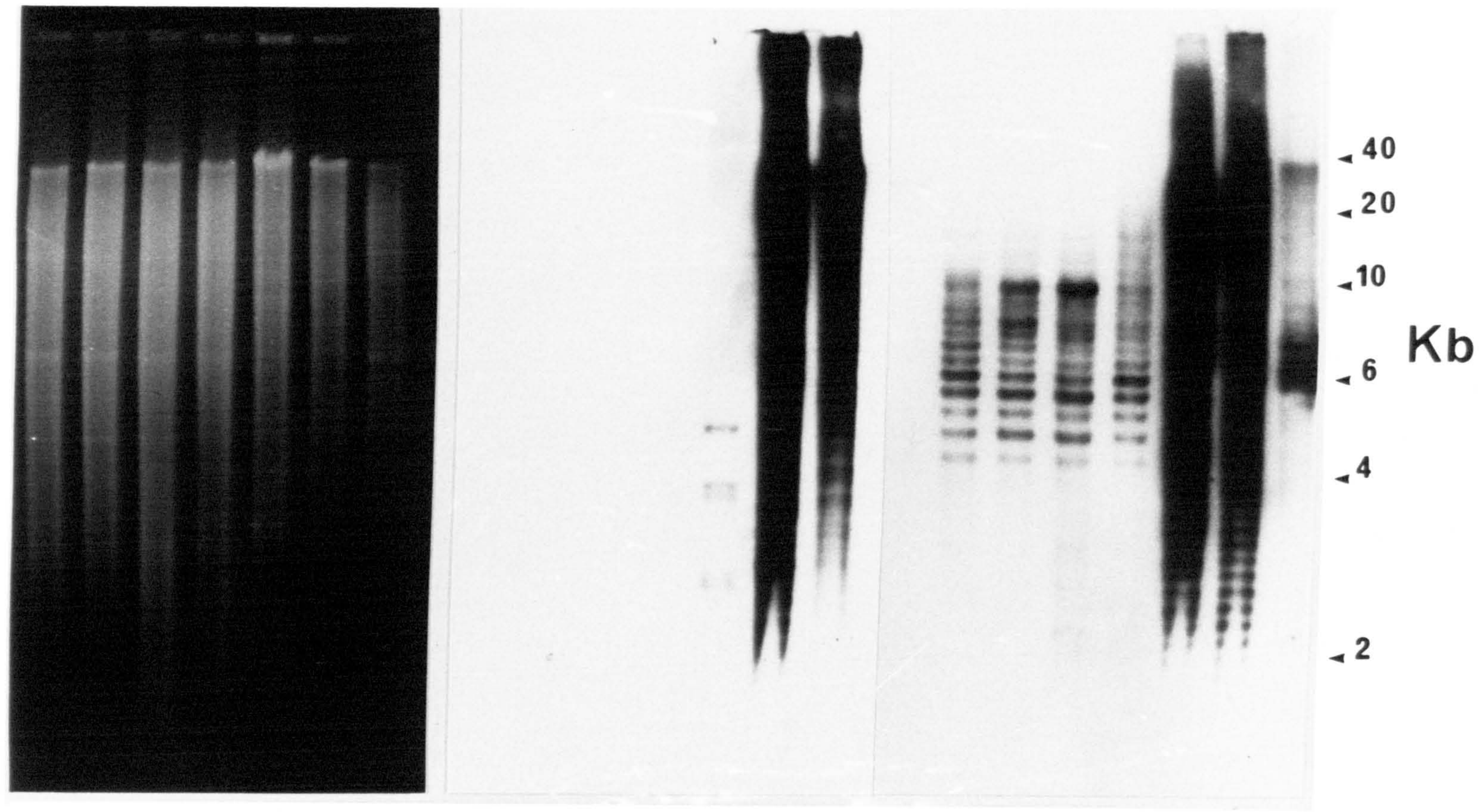

gambiae IAN
gambiae G3
gambiae PALA
gambiae KWA
melas
merus
arabiensis
40
20
10
6
4
2
Kb
A
B
C

Figure 4.13d,e; Result of probing a Southern transfer of EcoRI digested DNA from six female and six male *An arabiensis* (SENN) individual adults, using pAnaFI probe DNA. a; ethidium bromide stained gel and b; corresponding Southern transfer.

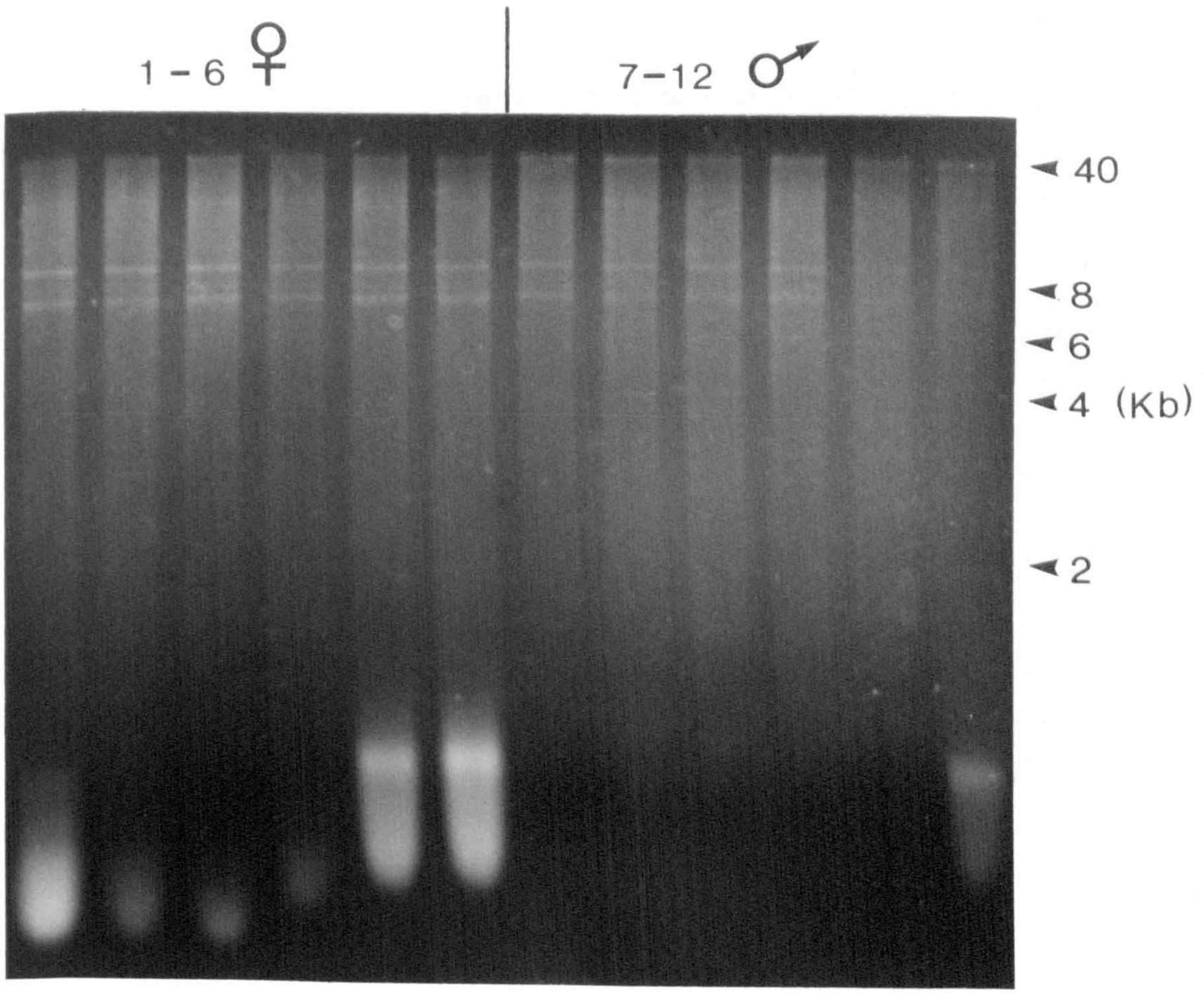
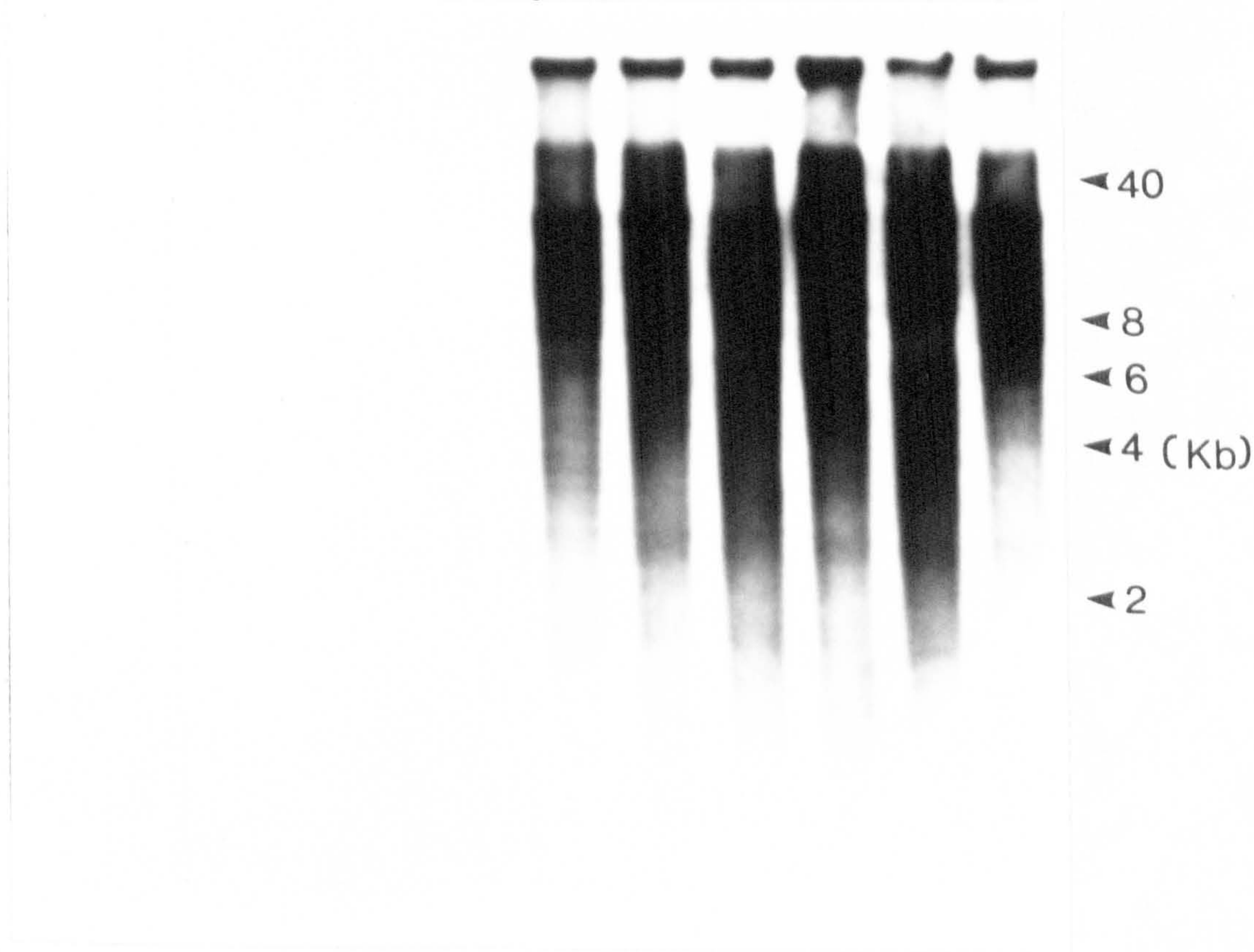


Fig. 4.13 d



e

4.14. Genomic organisation of pAnml 4-homologous sequences

To investigate the distribution of sequences homologous to pAnml 4 in the *Anopheles* genome, clone DNA was radiolabelled and used to re-probe the Southern transfer shown in Figure 4.13b (after the radioactive decay of the first probe). The results are shown in Figure 4.13c. This shows that hybridisation to the different *An. gambiae* s.s. laboratory stocks is qualitatively similar, consisting of approximately twelve distinct bands in the form of an oligomer series. The molecular weight difference between adjacent bands is 450 bp and the bands visible range from approximately 10 Kb to 1 Kb. Hybridisation to *An. melas* and *An. merus* also consists of an oligomer series, visible for the low molecular weight bands on the autoradiographic exposure shown in Figure 4.13c. The size range of bands visible is from very large (e.g. 50 Kb) down to approximately 1 Kb. No periodicity is observed for the hybridisation to *An. arabiensis* genomic DNA. This consists of five bands (7.5, 7.0, 6.3, 5.8 and 5.4 Kb in length).

4.15. Identification of wild-caught specimens

All wild-caught specimens were *An. arabiensis* (identified cytogenetically). These were collected prior to the discovery of the male-specificity of the *An. arabiensis*/*An. gambiae* s.s. distinguishing probe and were consequently mainly females. These are easier to identify cytogenetically than fourth instar larvae. Male adults cannot be identified cytogenetically. The number of males available for identification was therefore limited to five adults reared from a cytotyped female from Kafue, Zambia. These males were positive with the pAnaFI probe. All females tested were negative (results not shown).

4.16. Fly squashes; a simplified DNA extraction/immobilisation protocol

The method normally used to extract DNA from a single mosquito adult, pupae or larva (see Materials and Methods) yields 1-2 µg of DNA of fairly high purity and high molecular weight. This method is however very time-consuming and laborious, involving homogenisation of individual specimens followed by several precipitation steps. The number of specimens which may be processed by a single worker per day is therefore limited to about 40. This method also relies on the use of laboratory facilities such as pipettes, water-baths, a centrifuge and dot-blot apparatus. To enable larger numbers of samples to be processed with a minimum requirement for equipment, a method of immobilising crude DNA liberated directly from squashed material was needed. Tchen *et al* [160] have described such a technique for hybridisation to *Drosophila* individual flies. This section describes a similar technique developed for simplified identification of mosquitoes using DNA probes.

Frozen adults were cut into three portions; head, thorax and abdomen. These were then placed on nitrocellulose soaked in 2 x SSC and squashed using a glass rod. The nitrocellulose filter was then placed on a Whatmans filter paper soaked in 0.5 M NaOH, 1.5 M NaCl and left for 3 minutes to liberate and denature genomic DNA. The filter was then blotted dry and placed on Whatmans filter paper soaked in 0.5 M Tris, 3 M NaCl, pH 7.0 for 5 minutes (neutralisation). The filter was then baked (2 hrs, 80°C) to fix the DNA to the nitrocellulose. To remove unbound material, the filter

was soaked in 0.1 M NaOH for 5 minutes and the debris rubbed off with a gloved hand. After rinsing in 2 x SSC, the filter was prehybridised and hybridised as usual.

Figure 4.16a shows the result of this experiment using ten male *An. arabiensis* and ten male *An. gambiae* s.s. adults probed with pAnaFI clone DNA. It was concluded from these results that any segment of the mosquito was suitable for preparation of a squash, but that the head gave highest hybridisation signal in relation to the amount of material used. The head is also very easy to squash and is not required for other analyses (e.g. age or infectivity determination, or blood meal identification) which are routinely performed on adult mosquitoes.

To reduce the background hybridisation signal observed to the *An. gambiae* s.s. specimens observed in Figure 4.16a, the following modifications were derived from experiments using head squashes only.

1. Heads were squashed onto nitrocellulose soaked in 10% S.D.S. by covering filter with plastic and applying downward pressure.
2. Denaturing and neutralising steps were both carried out twice.
3. After denaturation, chitinous parts of heads were removed using forceps.

Results from this improved protocol are shown in Figure 4.16b. This also shows the effects of using a nylon ("Gene-screen") filter instead of nitrocellulose and of omitting the NaOH washing step. The background of hybridisation to *An. gambiae* s.s. in this experiment is

effectively zero (overnight exposure). No difference is observed between the results obtained using nitrocellulose and nylon filters. Slightly higher hybridisation signals were evident for the filters subjected to the NaOH washing prior to hybridisation.

This method may also be used on isopropanol preserved material subsequent to rehyrdration (soaking in T.E. buffer for five minutes for example). The ease of sample preparation using the technique described here would allow a single worker to prepare large numbers of fly squashes (e.g. several hundred) in a single day. These may then be baked, prehybridised and hybridised in a batch process.

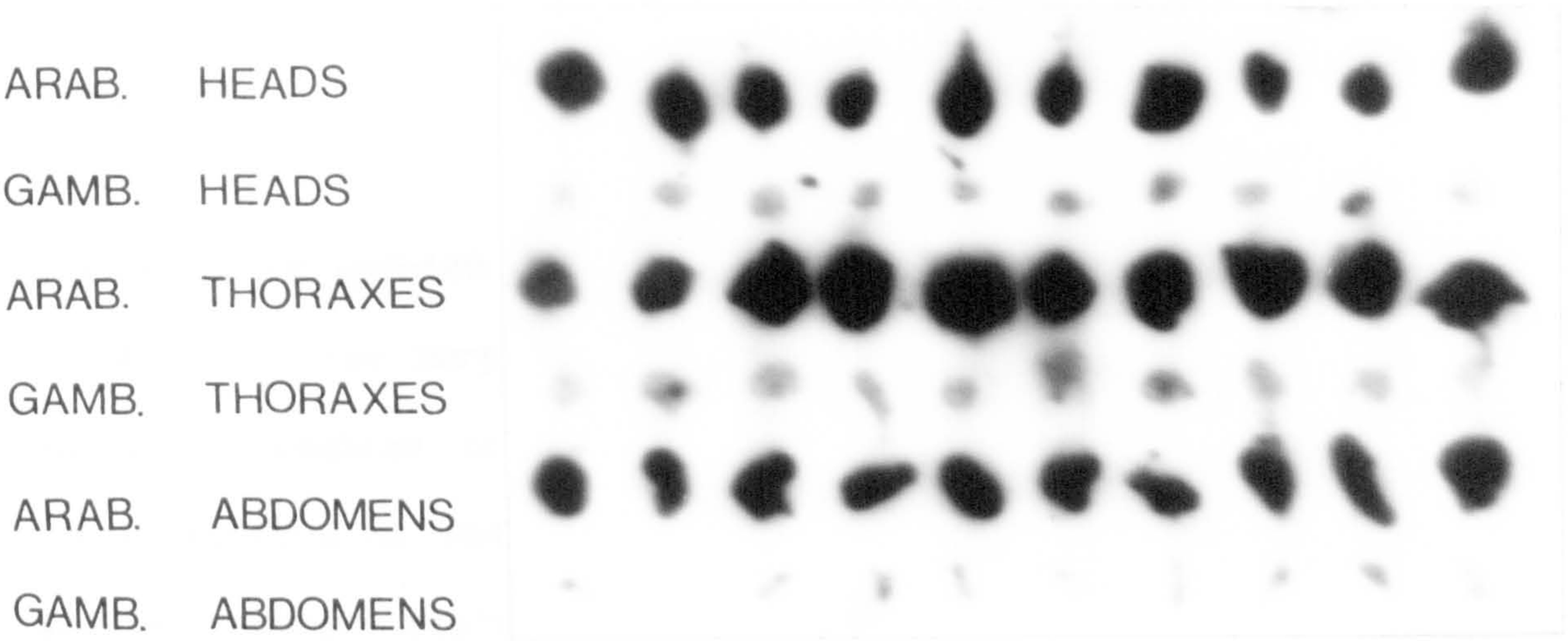


Fig. 4.16 a

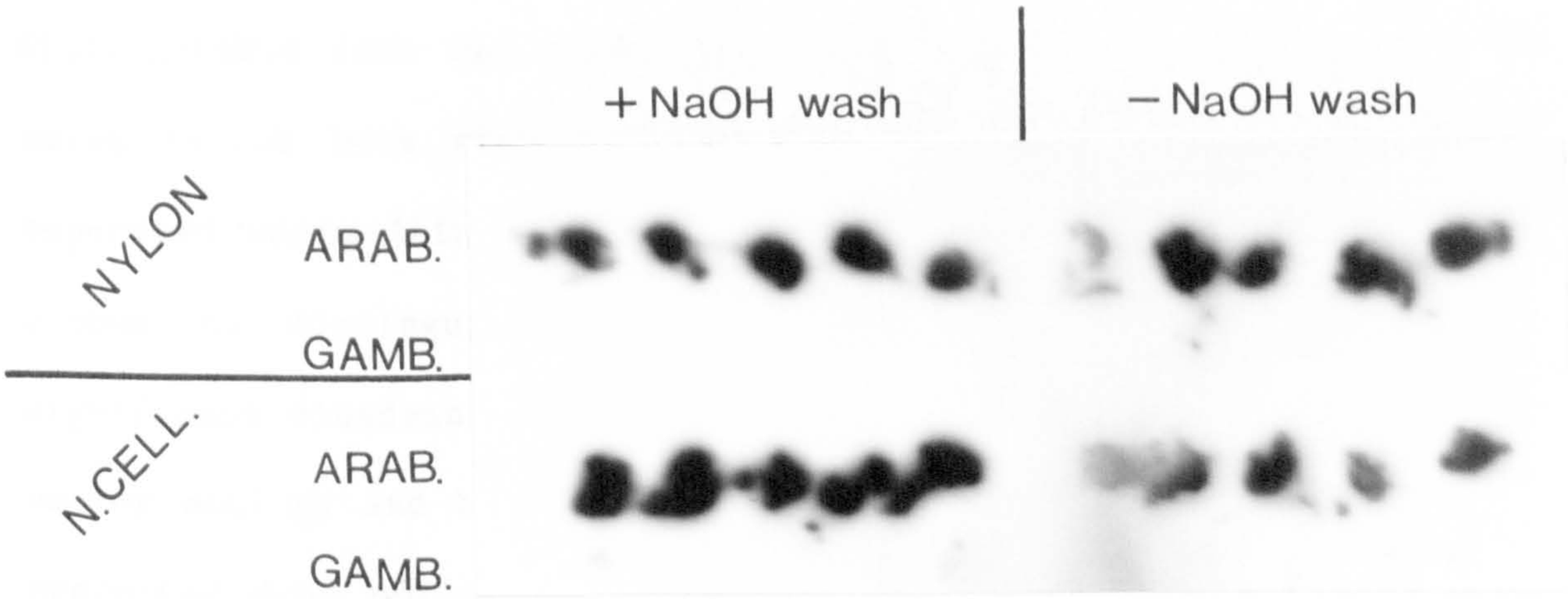


Fig. 4.16 b

Section 4; DISCUSSION

Separation of species

Although the morphological characteristics of mosquitoes in the *Anopheles gambiae* complex shows insufficient variation to allow identification of individual species, some repetitive DNA sequences show much more variation.

The most important species to distinguish in the *An. gambiae* complex are the two main vectors; *An. gambiae* s.s. and *An. arabiensis*. When used as hybridisation probes, pAnaFI and homologous clones clearly distinguish the males of these species. *An. gambiae* s.s. (males) is the only species to show zero hybridisation signal with these probes and can therefore be distinguished from all other species tested. Similarly, *An. melas* is the only species to show weak hybridisation signal with this probe and so may also be distinguished from the other three species. *An. arabiensis* and *An. merus* (which both show intense hybridiation to pAnaFI) cannot be separated using this probe. The limitation of pAnaFI - homologous probes to distinguishing only the males of the species is a significant disadvantage, as the biting females must be identified in vector ecology and behaviour studies. The male-specificity of these sequences does not prevent their use for the species identification of populations however, as both adults and pupae may be sexed easily. Populations of larvae (which may be sexed only by dissection) may also be identified if a 1:1 sex ratio is assumed. Probes homologous to pAnaFI therefore represent a potentially efficient means of

surveying vector populations before, during and after control programmes.

The pAnml 4 - homologous probes clearly distinguish the two saltwater species *An. melas* and *An. merus* (which give a very strong hybridisation signal with these probes) from the two freshwater species *An. gambiae* s.s. and *An. arabiensis* (which give a weak hybridisation signal). Sequences homologous to pAnml 4 are not sex-specific in the species available for testing (*An. gambiae* s.s. and *An. arabiensis*).

DNA probes which hybridise to all *An. gambiae* species, but not to species outside the *An. gambiae* complex (e.g. pAng 1) may be used to eliminate false negatives caused by misidentification of species. Further testing of such probes against a wide variety of species is required to verify their potential.

A combination of the three types of probe described here allows the identification of all four of the *An. gambiae* species available and gives a possible means of eliminating errors caused by misidentification of other *Anopheline* mosquitoes as *An. gambiae*. The high sensitivity of these probes enables identifications to be made using a small portion of an individual (the head gives excellent results) or a fraction (e.g. one tenth or less) of the total DNA extracted from an individual. Several tests involving a number of different DNA probes coupled with other analyses (e.g. infectivity determination, age determination or blood meal typing) may therefore be performed on the same mosquito.

Genomic organisation of pAnaFI - homologous sequences

The hybridisation of pAnaFI to *An. arabiensis* genomic DNA fragments of all sizes on Southern blots, irrespective of the restriction enzyme used to generate the fragments, infers that homologous sequences are interspersed with other sequences within the chromatin. It is also possible that a fraction of homologous sequences are tandemly repeated.

The karyotype of *An. gambiae* consists of two pairs of autosomes and one pair of sex chromosomes (XX : female, XY: male). The 'Y' chromosome is entirely heterochromatic in all *An. gambiae* species except *An. melas*, which may have 'Y' chromosome euchromatic regions [156]. The detection of pAnaFI homologous sequences in *An. arabiensis* males, but not females, and the high abundance of this sequence (or sequence family) in *An. arabiensis* males is compelling evidence for the supposition that pAnaFI homologous sequences are highly repeated on the heterochromatic 'Y' chromosome of this species. Material was not available to determine whether this is also the case for pAnaFI - homologous sequences in *An. merus* (in which homologous sequences are also present in high abundance) or *An. melas* (in which pAnaFI homologous sequences are present in low to middle copy number). *In-situ* hybridisation to mitotic chromosome spreads could be used to confirm the 'Y' chromosome location of pAnaFI homologous sequences in *An. arabiensis*.

It is interesting to note that heterochromatic differences between *An. gambiae* s.s. and *An. arabiensis* as visualised by use of the fluorescent stain Hoechst 33258 [156] are confined to the sex chromosomes of the two species. The 'Y' chromosome of *An. arabiensis* appears entirely fluorochrome dull, whereas that of *An. gambiae* s.s.

contains bright staining (A-T rich) regions. The distribution of bright-staining regions on the 'X' chromosome also differs between the two speices. In hybrid crosses between the two species, the fertility of the FI-backcross progeny is determined largely by the X : Y chromosome composition of the males. Approximately 50% of males with an 'X' and 'Y' chromosome from the same species are fertile, whereas those with 'X' and 'Y' chromosomes from different species are 100% sterile [161].

Differentiation of *An. gambiae* s.s. and *An. arabiensis* by the polytene chromosome banding technique (using aceticorcein staining, which allows the visualisation of positional or quantitative, but not qualitative differences in heterochromatin) relies on differences between the 'X' chromosomes of the two species. The 'Y' chromosome does not undergo polytenisation.

Genetic and cytogenetic differences between *An. gambiae* s.s. and *An. arabiensis* are therefore most evident between the sex chromosomes of the two species. It is not possible to say whether these differences have resulted in the speciation process or whether these differences have arisen after speciation as a result of genetic drift. Evidence for sequence-specific functions of heterochromatin (e.g. fertility factors in *Drosophila melanogaster* 'Y' chromosome heterochromatin [162]) support the former possibility.

Genomic organisation of sequences homologous to pAnml 4

Hybridisation of pAnml 4 to *An. gambiae* s.s., *An. merus* and *An. melas* EcoRI digested genomic DNAs is in the form of an oligomer series with periodicities of 450 bp, 170 bp and 170 bp respectively. This type of hybridisation pattern has been observed for limit-

digested genomic DNAs from other species ([163] for example) and is indicative of a tandemly repeated sequence with polymorphisms between the individual repeats which alter the restriction enzyme recognition sequence in some of the repeats. It is interesting to note that the periodicity of pAnml 4 homologous sequences is the same in the two geographically isolated saltwater species. In the species *An. arabiensis* and *An. gambiae* s.s., pAnml 4 homologous sequences are present in much lower copy number. No differences in hybridisation pattern are evident between the different *An. gambiae* s.s. laboratory stocks tested.

Lack of intraspecies variation

For a DNA probe to be of general use for species identification, interspecies differences must be consistently high, whereas intraspecies variation must be absent or negligible. The DNA probes described here do show consistently large differences in copy number between difference species. No differences have been detected between different laboratory stocks of the same species. The eleven different laboratory stocks of *An. gambiae* s.s. used, represent isolates from a wide variety of areas in Africa (see Section 4.5). The absence of intraspecies differences with respect to these DNA probes is therefore likely to be confirmed for most, if not all *An. gambiae* s.s. isolates. The three *An. arabiensis* laboratory stocks and the limited wild-caught material show no intraspecies variation, although further testing on wild-caught specimens is required. The finding that sequences homologous to pAnaFI are conserved in homology and copy number between the different species *An. arabiensis* and *An. merus* and sequences homologous to pAnml 4 are similarly conserved

between *An. melas* and *An. merus* is also evidence that these sequences are unlikely to show significant intraspecies variation. Testing against individual wild-caught specimens from a broad geographical range remains the only certain method of assessing intraspecies variation of a DNA sequence.

Intraspecies variation of sex chromosome heterochromatin has been reported by Gatti *et al* [156], although no qualitative differences have been observed for the *An. arabiensis* 'Y' chromosome. Intraspecies differences are limited to the fluorochrome banding pattern of the 'X' chromosome of *An. arabiensis* and *An. gambiae* s.s. and to the 'Y' chromosome of *An. gambiae* s.s. The stocks used successfully for species identification with the DNA probes described here, include two (IAN and G3) which do show intraspecies polymorphism with respect to 'X' chromosome staining. Differences in sequence arrangement will not affect results obtained using DNA probes against DNA immobilised on dot-blots.

Mosquito specimens may be preserved in isopropanol for long periods at ambient temperatures prior to species identification using DNA probes. This, coupled with the method of squashing specimens directly onto nitrocellulose to liberate the DNA prior to hybridisation using a highly sensitive radiolabelled DNA probe, enables this method of species identification to be used for large numbers of specimens collected in Africa and brought back to a laboratory. Future developments, such as the use of calorimetric DNA probes (see below) may soon make species identification using DNA probes possible in field laboratories. The strategy and techniques employed here may be applied to any species complex, providing the

members of the complex exhibit interspecies polymorphism of repetitive DNA sequences.

CURRENT DEVELOPMENTS

Isopropanol preserved *An. quadriannulatus* individuals have recently been obtained. These were kindly supplied as adults reared from isoenzyme typed females by S.M. Mpofu of the Blair Research Laboratory, Harare, Zimbabwe. Construction of a genomic library and testing of DNA probes against this species is currently underway.

The colorimetric hybridisation systems of Renz and Kurz [164] and Forster *et al* [165] are currently being assessed to ascertain the potential of these techniques for the routine species identification of mosquito specimens squashed onto nitrocellulose.

A method of distinguishing female *An. gambiae* s.s. and *An. arabiensis* using the male specific pAnaF1 DNA probe is currently being assessed. This method involves the removal of the spermatheca from blood-fed female adults. (This is an accessory sex organ used for the storage of sperm introduced by the male). The dissection is very easy to perform (the spermatheca is a distinctive chitinous 'box' cotained within the posterior abdominal segment of *Anopheles* mosquitoes) and represents little extra work if other analyses, such as parous rate and infectivity determinations and blood meal identification are also to be performed on the specimen. The spermatheca is invariably packed with sperm in blood-fed females (M.W. Service, personal communication) and this represents an excellent source of male DNA. This may be liberated onto nitrocellulose by breaking open the spermatheca using a dissecting needle. The very rare occurrence of cross-mating of different *An. gambiae* species in natural populations [122,138,139,141] makes the

identity of this sperm DNA an excellent marker of species identy of the female from which it was obtained.

Future Perspectives

Further work using the strategy and techniques described in this section may be performed in the following areas:

1. Further attempts to isolate a DNA probe to distinguish *An. gambiae* s.s. and *An. arabiensis* females. This should include the construction of larger genomic libraries of randomly sheared (e.g. 2-400 bp) DNA fragments in a bacteriophage vector to facilitate screening larger numbers of recombinants with genomic DNA from females only.

2. Further differential screening combinations (e.g. *An. gambiae* s.s. and *An. melas* or *An. merus*). This may yield new and potentially useful probes.

4. Testing for sex-specificity of DNA probes already isolated in the species *An. melas* and *An. merus*.

5. Further studies on the nature of the sequences which are species specific. These could include restriction mapping of cloned fragments to identify repeat lengths, *in situ* hybridisation to ascertain chromosomal location, and melting curves to ascertain sequence divergence of different copies of the sequences.

6. Evolutionary studies. Sequence comparisons of homologous repeats from the different species may be used to elucidate the route of species evolution in the *An. gambiae* complex. (At present it is

thought that the saltwater species and *An. bwambe* and *An. quadriannulatus* are the oldest species, with *An. gambiae* s.s. and *An. arabiensis* having evolved more recently and displaced the other species in many areas [122,128]).

7. Field tests. Using fly squashes and calorimetric hybridisation probes to compare this method with that of cytogenetics or isoenzyme electrophoresis. Further refinement of the technique (e.g. using M13 bacteriophage to give large amounts of single-stranded probe) and testing against large numbers of individuals from a wide range of localities is required.

8. Testing of pAng 1 - type sequences (which hybridise to all *An. gambiae species) against a wide variety of non - An. gambiae* species to test the specifity of these sequences.

Summary; A DNA Probe Key to the *Anopheles gambiae* Complex

SPECIES	PROBE		
	pAnaFI	pAnml 4	pAng 1
An. gambiae s.s.	-	+	+
An. arabiensis	++	+	+
An. melas	+	+++	+
An. merus	++	+++	+
Non-*An. gambiae* species	-	-	-

KEY:- +++ = very strong hybridisation

++ = strong hybridisation signal

+ = weak hybridisation signal

- = zero hybridisation signal

Notes

1, Sex-specificity not tested for in *An. melas* and *An. merus*.

2, *Wide range of non-An. gambiae* species not tested.

3, pAng 1 should serve as DNA quantity normalisation guide.

4, Positive control required (e.g. minute quantity of clone DNA or small quantity of *An. arabiensis* genomic DNA).

5, Negative control required (e.g. *An. stephensi* genomic DNA).

General Protocol for use of DNA Probes

Sample preparation

1. Take tissue (e.g. fresh, frozen, or rehydrated isopropanol preserved) and divide to give one part for each probe to be used.
2. Place in an ordered array on nitrocellulose filter soaked in 10% S.D.S. (one filter for each probe).
3. Cover with plastic sheet and squash tissue using uniform, downward pressue.
4. Remove plastic. Place filter sample-side up on Whatmans filter paper soaked in denaturing solution (0.5 M NaOH, 1.5 M NaCl) for 3 minutes. Repeat on freshly soaked Whatmans. Remove any large chitinous body parts using forceps.
5. Blot dry. Soak on Whatmans saturated with neutralising solution (0.5 M Tris, 3 M NaCl pH 7.0). Repeat. Blot dry.
6. Bake 2 hrs, 80°C.
7. Wash in 0.1 M NaOH for five minutes. Gently rub surface with gloved hand to remove unbound material.
8. Rinse in water. Rinse in 2 x SSC (0.3 M NaCl, 0.03 M NaCitrate).

Hybridisation and detection

1. Prehybridise (e.g. 1-2 hours at 42°C, in 50% formamide, 3 x SSC, 10 µg/ml denatured herring sperm DNA, 500 µg/ml yeast tRNA, 50 mM HEPES pH 7.0, 5 x Denhardts solution (1 x = 0.1% Ficoll, 0.1% polyvinylpyrrolidine, 0.1% bovine serum albumin).
2. Add labelled probe (e.g. to 50 ng/ml final concentration) and hybridise for 4 hours or longer at 42°C.

3. Rinse filter three times in 0.1 x SSC. Wash in 0.1 x SSC (e.g. 3 times for 20 minutes) at room temperature.
4. Detect probe (e.g. 2-4 hours autoradiography for radiolabelled probe of specific activity 10^8 dpm/μg).

Notes

1. Rehydration of isopropanol preserved material is necessary to facilitate squashing. Small samples (e.g. a mosquito head) may be rehydrated by soaking on 10% SDS for 5-10 minutes. Larger samples require immersion in aqueous solution until rigidity is lost.
2. Prehybridisation, hybridisation and detection steps required may vary according to the sensitivity and specificity of the probe labelling/detection system used.
3. Probe sensitivity may be increased, or hybridisation time decreased, by the addition of a volume exclusion polymer (e.g. polyethylene glycol 8000 to 6% w/v final concentration) to the hybridisation solution. This increases the effective probe concentration the hybridisation solution [164].

MATERIALS AND METHODS

Cont.. PAGE NO.

1. Isolation of Mosquito DNA Ref.[18]

This simple method yields 100-400 µg of pure, high molecular weight genomic DNA from 100-400 fourth instar larvae, pupae or adults. Several modifications of the original method have been made to improve yield and purity of the DNA obtained.

1. Grind 1-2 g of material in liquid nitrogen using a precooled mortar and pestle.
2. Suspend grindate in 3.0 ml of TNESST (10 mM Tris-HCl pH 7.5, 60 mM NaCl, 10 mM EDTA, 0.15 mM spermidine, 0.15 mM spermine, 0.5% Triton X 100). Homogenise in dounce homogeniser on ice, approximately 10 strokes.
3. Collect nuclei by centrifugation (7000 rpm, 7', 0°C).
   Add 2.5 volumes of absolute ethanol to supernatent to precipitate crude total cytoplasmic RNA. Store at -20°C.
4. Resuspend nuclei pellet in TNESST (5.0 ml). Respin 7000 rpm, 7' 0°C, to wash nuclei.
5. Resuspend pellet in TNESST (5.0 ml) and add 10% n-lauryl sarcosine to a final 2% concentration.
   Mix gently to lyse nuclei.
6. Add 12.1 g of Caesium chloride (Analar) and make volume 13.0 ml using TNESST/2% sarcosine. Dissolve by gently inverting tube.
7. Centrifuge 10000 rpm, 10', 4°C. Remove debris pellicle and transfer to ultracentrifuge tube. Check density is 1.68g/ml. Adjust if necessary.
8. Centrifuge 40,000 rpm (Sorvall T865), 40 hours, 18°C.

9. Fractionate through wide-guage needle from base of tube (e.g. 15 x 0.7 ml fractions).
10. Spot 1 µl of each fraction onto 1% Agarose containing ethidium bromide (0.5 µg/ml). Dry, visualise DNA containing fractions over U.V. transluminator.
11. Pool DNA containing fractions. Dialyse against T.E. (e.g. 3 times against 2 litres) to remove Caesium chloride.
12. Extract **gently** with an equal volume of T.E. saturated phenol and once with an equal volume of chloroform : isoamyl alcohol (24:1) to remove any contaminating protein.
13. Add NaCl to 200 mM, then 2.5 volumes of absolute ethanol. Incubate at -20°C for 20 minutes or longer, spin 10,000 rpm, 20',0°C. Wash pellet with 70% ethanol to remove salt. Dry under vacuum and redissolve overnight (4°C) in T.E. to give 2-400 µg/ml concentration.

Notes:

a. Vigorous shaking or pipetting through narrow tips should be avoided as this shears high molecular weight DNA.
b. Large amounts of starting material (more than 400 large larvae, pupae or adults) should be avoided to keep the viscosity of solutions low, prevent an excess of nuceleases relative to inhibitors in the TNESST and to prevent overloading of the $CsCl_2$ gradient.

2. Mosquito DNA Minipreparation

This method, cited in ref. [166], yields high quality DNA from as little as one adult, larva or pupa.

1. Homogenise individual(s) in a 1.5 ml Eppendorf tube containing 100 µl of 10 mM Tris-HCl pH 7.5, 60 mM NaCl, 5% sucrose, 10 mM EDTA using a glass rod.
2. Add 100 µl of 1.25% SDS, 0.3 M Tris-HCl pH 9, 0.1M EDTA, 5% sucrose, 0.8% diethylpyrocarbonate (DEPC, added when required). Incubate 65°C, 30'.
3. Chill on ice, add 30 µl 8M Potassium acetate. Mix. Incubate on ice for 45'.
4. Spin down precipitate (full-speed Eppendorf microfuge). To supernatent add 2 volumes of absolute ethanol. Mix. Incubate at 20°C for 20'.
5. Centrifuge 10'.
6. Redissolve pellet in 100 µl of 0.1 x SSC (1 x SSC = 0.15 M NaCl, 0.015 M Na citrate), 0.2% DEPC. Incubate at 20°C for 30'.
7. Add 3 vols. 100% ethanol. 5', 20°C. Spin 10'.
8. Redissolve pellet in 100 µl of 200 mM NaCl. Add 250 µl of absolute ethanol. Incubate 5', 20°C. Spin 10'. Wash pellet with 70% ethanol.
9. Redissolve dried pellet in 15 µl of T.E. buffer (overnight, 4°C).

Expected yield; 1-2 µg of DNA per mosquito adult, pupae or fourth instar larva. Females are larger and yield 20-50% more DNA than males.

3. Preparation of Mosquito Cell-Line DNA

1. Grow cells to near confluent lysis. (80cm² Falcon flasks).
2. Pour off medium, suspend cells in 5 ml of Hayes saline solution (9g NaCl, 0.26 g $CaCl_2 2H_2O$, 0.2 g KCl, 0.1g $NaHCO_3$ per litre) using a rubber policeman (on ice).
3. Pass suspension through a Pasteur pipette several times to break up clumps of cells (on ice).
4. Spin down cells (1200 rpm, 5') and resuspend in 0.5 ml of Hayes saline.
5. Add 4.5 ml of lysis solution (0.32 M sucrose, 10 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 1% Triton X 100). Homogenise on ice using glass/glass homogeniser.
6. Spin 10,000 rpm, 10', 4°C, to pellet nuclei. Precipitate supernatent to give crude total cytoplasmic RNA (see Materials and Methods 4 for purification of RNA).
7. Resuspend pellet in 2.25 ml of 75mM NaCl, 24mM EDTA. Add 125 µl of 10% SDS, 50 µl of proteinase K (10 µg/ml) and 75 µl of H_2O
8. Incubate at 37°C for 1 hour
9. Extract with phenol three times and chloroform twice.
10. Add NaCl to 200 mM and 2.5 volumes of ethanol. Mix. Spool out DNA precipitate. Rinse in 70% ethanol, dry, and redissolve in T.E. buffer (overnight, 4°C).

Expected yield: 100 µg of DNA per 80 cm² Falcon flask.

Isolation and Purification of Mosquito RNA

A Total cytoplasmic RNA by Phenol/NETS extraction

1. Spin down precipitated crude total cytoplasmic RNA (see Materials and Methods 3, step 3); 10,000 rpm, 10', 4°C.
2. Resuspend in 3 ml of NETS (100 mM NaCl, 1mM EDTA, 10 mM Tris-HCl pH 8.0, 0.2% SDS). Homogenise to dissolve.
3. Extract with an equal volume of NETS saturated phenol. Spin 10,000 rpm, 10', 4°C. Retain aqueous phase.
4. Repeat step '3' until no material is visible at interphase.
5. Extract once with NETS saturated phenol/chloroform/isoamyl alcohol (25:24:1) Spin 10 000 rpm, 10', 4°C
6. Extract with chloroform/isoamyl alcohol (24:1)
7. Add NaCl to 200 mM and then 2.5 volumes of absolute ethanol
8. Spin down precipitate 10,000 rpm, 20', 4°C. Wash with 70% ethanol
9. Dissolve dried pellet in 500 µl of water. Remove aliquot for yield determination or use. Store remainder under ethanol at -20°C.

Expected yield; approximately 4 µg of RNA per fourth instar larva, pupa or adult.

B Guanidine-hydrochloride extraction of total cellular RNA

1. Grind 2-400 larvae, pupae or adults in liquid nitrogen.
2. Suspend in 5 ml of 6M Gu-HCl/2M KAc. pH 5.0 (19 vols./1 vol).
3. Sonicate on ice until viscosity is low.
4. Spin 10,000 rpm, 10', 4°C to clear.
5. Add 0.5 volumes of absolute ethanol. Stand overnight at -20°C.
6. Spin 10,000 rpm, 20, 4°C, to pellet RNA
7. Dissolve pellet in 0.5 ml of 6M Gu-HCl, 2M KAc pH 5.0, 500 mM EDTA (9.5 vols.: 0.5 vols.: 0.5 vols).
8. Add 0.5 volumes of absolute ethanol. Stand for 1-2 hours at -20°C.
9. Spin down precipitate. Repeat steps 7 and 8 (selective precipitation of RNA).
10. Dissolve pellet in 0.5 ml of 20 mM EDTA.
11. Extract with an equal volume of chloroform:butanol (4:1).
12. To separated aqueous phase, add 3 volumes of 4 M NaAc pH 6.0. Stand at -20°C overnight. Spin down pellet. (This removes low molecular weight RNA and contaminating DNA).
13. Dissolve pellet in water, add NaCl to 200 mM and 2.5 volumes of ethanol. Spin down, redissolve in 100 µl of H_2O. Store under ethanol at -20°C.

Expected yield : 1-2 µg of RNA per mosquito.

To prepare nuclear RNA: Prepare nuclei pellet (see mosquito DNA extraction protocols). Wash pellet thoroughly to remove cytoplasmic RNA. Dissolve washed nuclei in 1.0 ml of Gu-HCl and continue as from

step 2 above using reduced volumes. N.B. Nuclei may be resuspended in a small volume (e.g. 50 µl) of TNESST to aid subsequent solvation in Gu-HCl.

Total cytoplasmic RNA may be prepared by ethanol precipitation of the supernatent after pelleting of nuclei and then redissolving the resulting RNA pellet in 500 µl of Gu-HCl and proceeding from step '2' above.

5. Selection of poly A$^+$ RNA by oligo-dT-cellulose column chromatography

1. Equilibrate oligo-dT-cellulose in binding buffer (0.4 M NaCl, 1 mM EDTA, 0.1% SDS, 10 mM Tris-HCl pH 7.5). Pour a 1 ml column.
2. Wash column with 3 volumes of water, 3 volumes of 0.1 M NaOH, and then 3 volsumes of water. Wash column with 5 volumes of binding buffer.
3. Dissolve RNA sample in 250 µl of binding buffer. Heat to 65°C for 2' (to reduce RNA secondary structure), cool to room temperature and apply to column followed by an equal vol. of binding buffer. (Flow rate should be less than 0.5 ml/minute).
4. Reload void material followed by an equal volume of binding buffer. Reload total void material followed by 4 ml of binding buffer.
5. Elute bound fraction with 2 ml of elution buffer (1 mM EDTA, 0.1% SDS, 10 mM Tris-HCl pH 7.5).
6. Make eluate 0.4 M with respect to NaCl. Heat to 65°C for 2'.

Chill, reapply to regenerated column (step 2 above).

7. Wash column with 4 ml of binding buffer, then 4 ml of elution buffer to collect 'bound-bound' fraction.
8. Make eluate 200 mM with respect to NaCl, add carrier (2 µl of 10 mg/ml glycogen) and ethanol precipiate by the addition of (2.5 volumes of absolute ethanol.

Notes:

Up to 10 mg RNA may be loaded per ml of column. Expected yield; 6-30 µg of poly A^+ RNA per mg of total RNA

6. Large-scale Plasmid Preparation

1. Inoculate a 500 ml L-broth culture using the contents of a single colony. Use appropriate antibiotic selection. For a chloramphenicol resistant strain grow overnight at 37°C with agitation. For a chloramphenicol sensitive strain, grow to an $O.D._{600nm}$ of 1.0 and then amplify overnight at 37°C using choramphenicol (170 µg/ml).
2. Spin down cells in 250 ml GSA bottles (7000 rpm, 10', 4°C). Drain pellet.
3. Resuspend pellet in 4 ml of 50 mM glucose, 25 mM Tris-HCl pH 7.5, 10 mM EDTA, 5 mg/ml fresh lysozyme. Incubate at 20°C for 10'.
4. Add 8 ml of 0.2 M NaOH/1% SDS. Mix gently. Incubate 10' on ice.

5. Add 4 ml of 3M K.Acetate pH 4.8. Vortex thoroughly, stand 30',0°C
6. Centrifuge 7000 rpm, 10', 4°C.
7. Carefully decant supernatent. Add a 60% volume of isopropanol. Stand for 5' at 20°C.
8. Centrifuge 10,000 rpm, 10', 20°C.
9. Redissolve pellet in 13 ml of T.E. buffer. Add 13.7 g of $CsCl_2$ (Analar) and 1.3 ml of 10 mg/ml ethidium bromide. Check density is 1.56 g/ml ± 0.02 g/ml.
10. Centrifuge 10,00 rpm, 10', 20°C to clear.
11. Centrifuge 45,000 rpm, 40 hours, 20°C to band DNA (e.g. Sorvall T865 fixed angle rotor)
12. Backlight centrifuge tube with U.V.. Remove the lower of the two bands visible using a syringe. (This should contain supercoiled plasmid whereas the upper band contains open-nicked plasmid and *E.coli* chromosomal DNA.
13. Remove ethidium bromide by extracting three times with isoamyl alcohol.
14. Dialyse 3 times against T.E. buffer to remove $CsCl_2$.

Expected yield; 100-500 µg of plasmid DNA per 500 ml culture.

7. Medium Size, Quick Plasmid Preparation

This method is a convenient way to prepare 5-10 µg of clone DNA suitable for nick-translation or restriction endonuclease digestion.

1. Grow an overnight culture (10 ml), or grow a 20 ml culture to an $O.D._{600nm}$ of 1.0 and amplify using chloramphenicol (170 µg/ml) overnight at 37°C.
2. Harvest cells (3000 rpm, 20', bench centrifuge). Drain pellet.
3. Suspend pellet in 0.35 ml of 50mM glucose, 25mM Tris-HCl pH7.5, 10mM EDTA, 5mg/ml fresh lysozyme. Incubate R.T. 5'.
4. Add 0.7 ml of 0.2M NaOH/1% SDS. Mix gently. Chill 0°C, 5'.
5. Add 0.35 ml of 3M KAcetate pH 4.8. Vortex thoroughly. Incubate 30' on ice.
6. Spin 5' (full speed, Eppendorf microfuge.)
7. Remove supernatent. Add a 60% volume of isopropanol. Stand at 20°C, 5'.Spin 5' as in '6' above.
8. Redissolve pellet in 300 µl of T.E.. Digest on ice with RNase (30', 100 µg/ml RNase).
9. Phenol extract twice, phenol/chloroform extract once, chloroform extract once.
10. Add NaCl to 200 mM. Add 2.5 volumes of absolute ethanol. Snap freeze. Thaw. Spin down precipitate.
11. Wash peellet with 70% ethanol. Dry. Redissolve in 50 µl of T.E. buffer.

N.B. Phenol extraction at stage '6' may be included to reduce nuclease contamination if this is encountered in the final DNA.

Glycerol stocks; Grow a small overnight culture with antibiotic selection. Take 0.5 ml and add to 0.5 ml of sterile glycerol. Mix. Store at -20°C. Renew yearly.

8. Transformation of *E. coli* with Plasmid DNA

This is a very simple protocol of the method originally developed by Mandel and Higa [167]. Control transformations (using supercoiled pAT153 plasmid DNA) typically yield 1 x 10^6 - 2 x 10^6 transformants per microgram of DNA.

1. Grow 2 x 10 ml cultures of the prospective host cells (e.g. *E.coli* MC1060) in L-broth to an $O.D._{600nm}$ of 0.6. The inoculum used should be from fresh, vigorous growth on L-agar.
2. Spin down cells (3000rpm, 5', bench centrifuge).
3. Resuspend cells in 2 x 5 ml of 0.1 M $MgCl_2$ on ice.
4. Pellet as in '2' and resuspend combined pellets in 1 ml of 0.1M $CaCl_2$, on ice, very gently.
5. Leave on ice, 30'.
6. Add 100 µl of competent cells from '5' to DNA in a small volume of T.E. or ligation buffer, with 10 µg of tRNA carrier.
7. Leave on ice 40'. Gently agitate to prevent settling.
8. Heat shock, 42°C, 2'.
9. Leave on ice 20'.

10. Add L-broth to 1 ml. Incubate at 37°C for 50'. (This allows initiation of antibiotic resistance expression).
11. Plate onto L-agar containing appropriate antibiotic to select transformants.

Transformation controls

1. Supercoiled plasmid (e.g. 10 ng, 1 ng and 0.1 ng) to determine tranformation efficiency.
2. No DNA added, to check for contamination of buffers etc. with plasmid DNA and selection of transformants.
3. Digested vector DNA: linearised plasmid should tranform 10-100 fold less efficiently than supercoiled plasmid. This serves to check digestion efficiency.
4. Religated cut vector. This serves to check the ligation reaction efficiency.
5. Vector DNA (e.g. linearised, dephosphorylated plasmid)
6. Re-ligated vector DNA. This should give a low level of transformants (e.g. less than 0.1% of supercoiled efficiency) A large number of tranformants in this control and a low level in '5' above, indicates poor dephosphorylation efficiency.

Controls 1, 4 and 6 are always required. The other controls may be omitted once a new batch of vector has been tested.

Removal of terminal 5' phosphates using calf intestinal alkaline phosphatase (CIAP)

Enzyme used: BCL, 22 units/µl (1 unit of CIAP hydrolyses 1 µmole of p-nitrophenyl phosphate in 1 minute at 37°C under standard assay conditions).

CIAP buffer: 10 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$.

Procedure: typically, 22 units of CIAP were used to dephosphorylate 5-10 µg of linearised plasmid DNA in a 20 µl reaction volume at 37°C for 30 minutes. After incubation, CIAP was inactivated by heat (65°C, 30 minutes) followed by phenolisation.

9. Screening Bacterial Clones

This method, originally developed by Grunstein and Hogness [168] allows 500-1000 colonies to be screened per 9 cm diameter petri-dish.

Making Replicas

1. Suspend diluted cells in 4 ml of L-broth. Apply evenly to sterile, moistened nitrocellulose disc using a Buchner funnel and a vacuum pump. Place filter on L-Agar plate.
2. Grow colonies overnight at 37°C, or for high density screening, for 6 hours at 37°C, to give small colonies.
3. Dry plate briefly. Place fresh nitrocellulose discs on L-Agar plates to moisten. Mark orientation spots on the master plate using a waterfast marker pen.

4. Remove master plate nitrocellulose disc and place colony side up on sterile Whatmans 3 MM filter paper. Place fresh nitrocellulose disc from '4' on top, with Agar-contact surface face up.
5. Place sterile Whatmans filter on top of nitrocellulose discs and glass plate on top of this. Press down. Mark orientation spots on replica filter and peel filters apart. Place each on fresh L-Agar plates, colony side up.
7. Reincubate master 2-4 hours. Incubate replica to give small colonies (e.g. 4-6 hours).

N.B. Two replicas may be made of the master plate for duplicate-screening. Further replicas may also be taken after reincubation of the master plate.

Fixing DNA

1. Place filter colony side up on Whatmans 3MM soaked in 2.0 M NaCl/0.5 M NaOH, 10' to lyse cells and denature DNA.
2. Place filter on Whatmans 3MM soaked in 3M NaCl, 0.5 M Tris-HCl pH 7.0. Leave 5'. Repeat on fresh filter paper.
3. Bake 80°C, 2 hours.
4. Soak in 0.1 M NaOH for 5'. Gently rub surface with gloved hand to remove unbound material.
5. Rinse in distilled water, then 0.5 M Tris-HCl pH 7.0 to neutralise. Blot dry and store or prehybridise at usual.

10. Growth of Bacteriophage λ

Materials

L-broth (1 litre); "Difco Bacto-Tryptone"; 10g, "Difco Bacto Yeast Extract"; 5g, NaCl; 5g, glucose; 1g.

Tryptone Agar (1 litre); "Difco Bacto-Tryptone"; 12g, "Difco Bacto Agar"; 10g, NaCl; 10g.

Tryptone top Agar - as for Tryptone Agar, but add only 6.5g/litre Agar and make 10 mM with respect to $MgSO_4$.

Phage buffer: KH_2PO_4; 3g, Na_2HPO_4; 7g, NaCl; 5g. Dissolve in 979 ml of water, autoclave. Then add sterile 0.1 M $MgCl_2$; 10 ml, sterile 0.01 M $CaCl_2$; 10 ml and sterile 1% gelatin solution; 1 ml.

Plating Cells (e.g. *E. coli* LE392)

Inocculate a 5 ml culture of L-broth-magnesium-maltose (L-broth made with 4g/litre maltose, no glucose and made 10 mM with respect to $MgSO_4$). Incubate overnight at 37°C. Use 400 µl of this to inoculate a 20 ml culture in same medium. Grow to an $O.D._{600nm}$ of 0.3. Harvest cells and resuspend in 2 ml of sterile 10 mM $MgSO_4$. Use 200 µl of plating cells per petri-dish or 2 ml per bioassay tray when plating phage. Store for up to 4 days at 4°C.

Plating Phage

Dilute phage stock in phage buffer (e.g. 100 µl). Add to plating cells. Incubate 37°C, 15' for preadsorption. Add to 3 ml of molten Tryptone top Agar (50°C). Mix gently and pour immediately onto a prewarmed (37°C), thick, dry, Tryptone plate. Allow to

set, dry briefly. Incubate inverted at 37°C. For bioassay tray use 25 ml of Tryptone top Agar. Ensure plates and top Agar are poured on a flat surface.

11. Large-Scale Preparation of Bacteriophage λ DNA

1a. Use titred stock to give multiplicity of infection of 0.1-0.01. Use a 1/1000th culture volume of plating cells. Pre-absorb 15', 37°C. Inoculate L-broth made 10mM with respect to $MgSO_4$.

1b. (Alternative to 1a). Take single fresh plaque (e.g. using sterile Pasteur) and add to 100 µl of phage buffer. Incubate at 37°C for 20' to allow phage to diffuse out of Agar. Use this as inoculum as described in. '1a'.

2. Grow overnight with good aeration. If no lysis (e.g. lumps of debris) is visible, add an equal volume of fresh medium and continue to incubation for 2-6 hours.

3. Add a 1/100th volume of chloroform and shake for a further 10' to lyse infected cells. Stand for 10' to allow chloroform to settle.

4. Decant lysate (avoiding chloroform) into GSA bottles. Centrifuge (8000 rpm, 15', 4°C) to pellet cell debris..

5. To supernatent, add $MgCl_2$ to 10 mM. For 500 ml of lysate; add 17.5 g of NaCl and 62.5 g of polyethylene glycol 6000. Dissolve cold by shaking. Stand at 4°C overnight.

6. Spin down phage precipitate (7500 rpm, 20', 4°C) and drain thoroughly on ice.

7. Resuspend phage pellet in 5 ml of T.M. buffer (10 mM Tris-HCl pH 7.5, 2 mM $MgCl_2$). Mix for 30' to ensure complete dispersion of phage precipitate.
8. Centrifuge 10,000 rpm, 10', 4°C, to pellet debris.
9. Add RNase to 50 µg/ml and DNase to 10 µg/ml. Incubate 20°C, 60'.
10. Prepare stepped $CsCl_2$ gradients. Dissolve 63g of 'Analar' $CsCl_2$ in 50 ml of T.M. buffer to give a density of 1.7g/ml. Take 20 ml of this and add 8 ml of T.M. buffer to give 1.5 g/ml density and 10 ml and add 13 ml of T.M. buffer to give a density of 1.3 g/ml.
    Place 2 ml of 1.3 g/ml $CsCl_2$ in a 13 ml ultracentrifuge tube. Layer 2 ml of 1.5 g/ml $CsCl_2$ underneath and 2 ml of 1.7g/ml $CsCl_2$ underneath this. Layer 2-5 ml of sample on top of stepped gradient and fill tube with T.M. buffer.
11. Centrifuge in a swing-out rotor; 32000 rpm, 18°C, 2 hours, (e.g. Sorvall Tst 41 rotor, 'g' = 1.3×10^5).
12. Collect pale blue phage band, usually visible between 1.5 and 1.7 g/ml steps using syringe.
13. Dialyse twice against 2 l of 25 mM NaCl, 10 mM Tris-HCl pH 7.5, 1 mM $MgCl_2$.
14. Make phage suspension 1% with respect to SDS and 10 mM with respect to EDTA. Heat to 65°C for 15'.
16. Cool to 37°C. Add proteinase K to 50 µg/ml (from 10 mg/ml stock in 10 mM Tris-HCl pH 7.5). Incubate 37°C, 1 hour.
17. Phenol extract twice. Phenol/chloroform extract once. Chloroform extract once.
18. Dialyse three times against 2 l of TE buffer. Store at 4°C.

12. Rapid, Small-Scale Isolation of Bacteriophage DNA

1. Pick a single fresh phage plaque into 100 µl of phage buffer. Incubate at 37°C for 10' to allow phage to diffuse out of Agar.
2. Remove 75 µl from '1' and add to 5 µl of *E. coli* LE392 plating cells. Incubate at 37°C for 10' (preadsorption of phage).
3. Add 2 ml of phage broth, incubate shaking, overnight at 37°C.
4. Add 5 µl of chloroform and shake for a further 10' to complete lysis.
5. Transfer 1.5 ml to Eppendorf tube, spin 5'. Transfer 1.2 ml of supernatent to a fresh Eppendorf.
6. Add 300 µl of polyethylene glycol/NaCl (20%/2.5 M). Incubate at room temperature for 20'. Spin 5'. Remove supernatent. Respin 2' and remove remaining supernatent.
7. Resuspend pellet in TE (150 µl). Phenol extract once. Chloroform extract once. Add NaCl to 200 mM and 2.5 volumes of absolute ethanol. Incubate at -20°C for 20'.
8. Spin 5', wash pellet with 70% ethanol. Redissolve dried pellet in 50 µl of TE/20 µg/ml RNase.

Expected yield; 1-2 µg of DNA suitable for enzymatic reactions.

13. Screening Phage Plaques (Ref.[169])

1. Soak nitrocellulose filters in 3 x SSC, blot briefly to remove excess liquid and sandwich between two sheets of clingfilm. Sterilise over U.V. transluminator.
2. Chill phage plate at 4°C (this prevents top Agar from lifting off subsequently). Dry briefly. Place orientation marks on underside of plate.
3. Carefully lay nitrocellulose on top of phage plate. Avoid trapping air bubbles. Leave 1 minute (1st copy), 2 minutes (2nd copy), 3 minutes (3rd copy) etc. Mark orientation spots on filter using a marker pen.
4. Carefully remove filter and place on Whatmans 3MM filter paper soaked in NaOH/NaCl (0.5 M/1.5 M) phage side up, for 3' to denature the DNA.
5. Neutralise by briefly rinsing filter in 0.5 M Tris-HCl pH 7.4/1.5 M NaCl. Blot dry.
6. Bake; 2 hours, 80°C.

14. *In-Vitro* Packaging of λ DNA

Preparation of Packaging Extracts (Ref. [20])

Strains: Freeze-thaw lysate: BHB2688

Sonicated extract : BHB2690

Checking strains: Boths grow at 32°C, but not 42°C due to cIts 857 mutations and both strains are significantly more susceptable to U.V. damage than are recA$^+$ strains.

Packaging buffers. Buffer 'A'; 20mM Tris-HCl pH 8.0, 3mM $MgCl_2$, 0.05% β-mercaptoethanol, 1 mM EDTA. Buffer 'M1'; 6mM Tris-HCl pH 7.5, 60 mM spermidine, 60 mM putrescine, 18 mM $MgCl_2$, 15mM ATP, 0.1% β-mercaptoethanol. Both buffers should be filter sterilised.

A, Freeze-thaw lysate (FTL)

1. Streak out BHB2688 onto L-Agar and grow overnight at 32°C.
2. Pick a single small colony and use this to inoculate a small overnight L-broth culture (32°C).
3. Use a 1/100th volume of overnight culture to inoculate 3 X 500 ml cultures (L-broth minus glucose, plus 4g/litre maltose, 10 mM $MgSO_4$). Grow to $O.D._{600nm}$ of 0.3 (30°C).
4. Induce lysogen by maintaining cultures at 45°C for 15'.
5. Grow for further 1 hour at 37°C with vigorous shaking.
6. Chill on ice, harvest cells (9000 rpm, 10', 4°C) in sterile GSA bottles.
7. Drain off all supernatent (5' on ice)
8. Resuspend pellet from each 250 ml bottle in 0.5 ml of 10% sucrose, 50 mM Tris-HCl pH 7.5.

9. Pool resuspended pellets. Add 75 µl of fresh lysozyme solution (2 mg/ml in 0.25 M Tris-HCl pH 7.5). Mix gently.
10. Snap freeze in liquid nitrogen. Thaw gradually to 4°C. Add 150 µl of M1 buffer. Mix gently.
11. Centrifuge 17000 rpm, 50', 4°C. Distribute 50-100 µl aliquots of supernatent into pre-chilled, sterile Eppendorf tubes.
12. Snap freeze in liquid nitrogen. Store at -75°C.

B, Sonicated extract (SE)

1. Streak out BHB2690 onto L-Agar. Grow overnight at 32°C. Pick a single small colony and use to inoculate a small overnight L-broth culture.
2. Use a 1/100th volume of overnight culture to inoculate one 500 ml L-broth-maltose-magnesium culture. Grow to an $O.D._{600nm}$ of 0.3 (32°C).
3. Induce lysogen 15', 45°C.
4. Grow for further 1 hour at 37°C with vigorous agitation.
5. Chill on ice, harvest cells (9000 rpm, 10', 4°C,).
6. Drain thoroughly (5' on ice). Resuspend each pellet in 0.5 ml of buffer 'A'.
7. Pool resuspended pellets (1 ml). Add 2.6 ml of buffer 'A'.
8. Sonicate on ice in short bursts until viscosity is low.
9. Pellet debris (10,000 rpm, 10', 4°C).
10. Distribute supernatent into sterile, pre-cooled Eppendorf tubes in 50 µl aliquots.
12. Snap freeze in liquid nitrogen. Store at -75°C.

Packaging Reaction

Thaw FTL and SE to 0°C. Mix in order: 7 µl of buffer 'A', 1-2 µl (100-1000 ng) of DNA to be packaged, 1 µl of M1 buffer, 6 µl of SE and 10 µl of FTL. Incubate at 25°C for 1 hour. Stop reaction by diluting to 500 µl with phage buffer. Add 2 drops of chloroform. Plate phage.

N.B. Ratio of FTL to SE should be optimised for different extract preparations. Buffers should be aliquoted and thawed once only.

15. Agarose gel electrophoresis of nucleic acids

Apparatus

Large gels; BRL series H4. Gel size 20 cm x 25 cm, volume 300 ml. Usually run overnight at 40 V (1.2 V/cm).

Minigels; BRL series H6. Gel size x 7 cm, volume 30 ml. Usually run 1 hour, 70V (4V/cm).

Materials

Agarose; 0.4-2% (Miles Lab. Ltd). **Ethidium bromide;** used at final concentration of 0.5 µg/ml in gel and electrophoresis buffer (Electran). **Tris-acetate electrophoresis buffer;** 4 mM Tris, 2mM EDTA, 2mM NaAcetate, pH adjusted to 7.7 with glacial acetic acid. **Loading buffer;** 0.2% Agarose, 20 mM EDTA, 10 mM Tris-HCl pH 7.5, 20% glycerol, 0.1% bromophenol blue (Melt, allow to set in syringe, pass through 24 guage needle three times).

Photographing gels

Gels were photographed over a short wavelength (302nm) U.V. transluminator (U.V. Products Inc. TM36) using a Polaroid MP4 land camera. The film used was 'Polapan 52' (400 ASA) instant film, or Illford FP4 (125 ASA). Exposures were 1-2S and 4-8S respectively. A 'Wratten' 2A and 15 filter was used for all photographs.

DNA molecular weight markers

A, λ/EcoRI + λ/Hind III (separate digests); 23.7, 21.8, 9.46, 7.55, 6.60, 5.93, 5.54, 4.80, 4.20, 3.38, 2.20, 1.90 and 0.53 Kb.

B. pAT153/HaeIII; 587, 458, 434, 339, 267, 234, 213, 197, 192, 184, 122, 89, 80, 64, 57, 51, 32, 21, 18, 11 and 2 bp.

16, Southern transfer of DNA from Agarose to nitrocellulose or nylon filters

The original method of E.M. Southern [170] was used to transfer DNA from Agarose gels to nitrocellulose or nylon filters. (Schleicher and Schuell or 'Gene-Screen' respectively).

Preparation of gel

1. Mark positions of markers by injecting the U.V. visiualised DNA bands with normal scripting ink.
2. Trim gel to size. Cut through centre of wells laterally.
3. Denature gel in 3 volumes of 1.5M NaCl/0.5M NaOH; 20' for minigels (30 ml volume). 60' for large gels (100-300 ml

volume). Agitate gently.

4. Neutralise in 3 volumes of 3M NaCl/0.5M Tris-HCl pH 7.0 for 30' (minigels) or 60' (large gels). Agitate gently.

Blotting Apparatus

Place glass plate on four plastic bottle tops in the bottom of a tray. Cover in 3 layers of Whatmans 3MM filter paper, overlapping the edges of the plate. Fill tray to just below the level of the glass plate with 20 x SSC.

Blotting

1. Lay gel on platform, smooth away all air bubbles.
2. Place glass plates around gel, leaving a 0.5 cm gap between gel and plates.
3. Soak nitrocelllose filter in 2 x SSC and layer onto gel, overlapping onto glass plates. Smooth away all air bubbles.
4. Soak 3 sheets of Whatmans 3MM in 2 x SSC and layer over nitrocellulose filter. Place a thick wad of dry tissues on top of this and a brick on top of the whole apparatus.
5. Leave overnight or longer. Top up with 20 x SSC if required.
6. Uncover. Place nitrocellulose with gel on top, on a clean glass plate and mark position of wells. Number blot and trim off excess nitrocellulose.
7. Discard gel. Mark positions of markers over U.V..
8. Rinse filter for 5' in 5 x SSC. Blot dry.
9. Bake 2 hours, 80°C. Store or prehybridise.

17. Radiolabelling of DNA and RNA

A. Nick Translation of DNA (Ref. [171])

Solutions

1 x DNase buffer; 10 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 1 mg/ml bovine serum albumin (BSA).

10 x nick-translation buffer; 500 mM Tris-HCl pH 7.5, 500 mM NaCl, 100 mM $MgCl_2$, 70mM β-mercaptoethanol, 500 µg/ml BSA.

dNTP mix; 20 µM each of dGTP, dTTP, dATP

G100 buffer: 20 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA, 0.05% n-lauryl-sarcosine.

$\alpha^{32}P$ dCTP label: 10 µCi/µl, 3000 Ci/mmole. (Amersham Int).

DNase: 1 mg/ml in 10 mM HCl stored in small aliquots at -20°C (Sigma chemicals).

Reaction

1. Activation of DNase; dilute to 200 ng/ml in DNase buffer.
   Leave on ice, 2 hours (activation).
2. Add to tube on ice; 10 x N.T. buffer (1 µl), dNTP mix (1µl), DNA (200 ng), DNase I; 200 ng/ml (1 µl), α ^{32}P dCTP (3 µl), DNA polymerase I; 2 units, H_2O to 10µl.
3. Incubate 2 hours, 15°C.

G100 Desalting

Pour 2.5 ml Sephadex G100 column in a Pasteur pippete. Use nylon filter wool as base. Wash column with 2 column volumes of G100 buffer. Apply reaction mixture to column. Elute column using G100 buffer. Collect 10 x 200 µl fractions. Count. Pool first peak (usually F4 + F5). Denature by boiling 5'.

Specific activity: 1-2 x 10^8 dpm/µg. This may be varied according to DNA/dCTP in reaction.

B, Kinase Labelling of RNA (Ref. [172])

Base cleavage: Take 200 µl of RNA at a concentration of 50 µg/ml or higher. Add 20 µl of 1 M NaOH. Incubate on ice, 30'. (This generates fragments approximately 200 bp in length). Add 20 µl of 1 M Tris-HCl pH 7.5 and 20 µl of 1 M HCl. Ethanol precipitate. Redissolve at a concentration of 50 µg/ml in distilled water. Store at -20°C.

Reaction; To tube on ice, add: 1.25 µg of base cleaved RNA (25 µl), 50 mM $MgCl_2$ (5 µl), 1 M Tris-HCl pH 7.5 (2.5 µl), 200 mM β-mercaptoethanol (2.5 µl), 1 unit of T4 polynucleotide kinase, 50 µCi $\gamma^{32}P$ ATP (5 µl 'Amersham'- 5000 Ci/mmole). Incubate 1 hour, 37°C. Desalt as in 'A' above. Collect front peak. Do not boil before use as hybridisation probe.

Specific activity: 10^7 - 10^8 dpm/µg.

18. Filter hybridisations

Hybridisation solution; 50% deionised formamide, 3 x SSC (0.45 M NaCl, 0.045 M NaCitrate), 50 µg/ml denatured, sheared herring sperm DNA, 10 µg/ml polydeoxyadenine, 50 mM HEPES buffer pH 7.0, 5 x Denhardts solution (1 x Denhardts = 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% BSA). Store -20°C.

Prehybridisations were performed in heat-sealed polythene bags at 42°C overnight using 100 µl of hybridisation solution per cm^2 of filter.

Hybridisations were usually performed using a probe concentration of 50 ng/ml in plastic bags at 42°C. When using heterologous probes, less stringent hybridisation conditions (e.g. 40°C , 5 x SSC) were used. Hybridisations were usually left overnight, although for some probes (e.g. complex DNA or RNA mixtures) hybridisations were left for 2 days. Hybridisation conditions are discussed in Ref. [173].

Washing filters: After hybridisation, filters were rinsed three times in 2 x SSC, washed three times for 20', in 2 x SSC and then once in 0.1 x SSC at 50°C for 20' (homologous probes). When heterologous probes were used, filters were washed astringently (e.g. 2 x SSC at 20°C) and if necessary, re-washed at increased stringency after autoradiography.

Autoradiography was performed at -75°C in 'X-ograph' cassettes using 'Speed-X' intensifying screens and preflashed 'Fuji RX' X-ray film.

Stringency of hybridisation

The stringency of a nucleic acid hybridisation reaction is determined by the proximity of the hybridisation conditions employed to the melting temperature (Tm) of the duplex (DNA/DNA or DNA/RNA) which is formed.

The Tm of a duplex may be calculated using the equation;

$$Tm = 81.5 + 16.6 (\text{Log } M) + 0.41 (\%\ GC) - 0.72 (\%\ \text{formamide})$$

Where Tm = melting temperature in °C

M = molarity of monovalent cation in reaction

%GC = % of guanine and cytosine residues in the duplex and % formamide = % of formamide in the hyrbidisation buffer.

The hybridisation parameters used here (3 x SSC, 42°C, 50% formamide) correspond to a temperature of 20°C below the melting temperature of a perfectly matched DNA/DNA hybrid of 50% GC content under these temperature, salt and formamide conditions. (1% mismatch of hybrid sequences lower Tm by approximately 1°C). These are stringent hybridisation conditions which favour a high rate of reassociation of matched sequences with a low level of reassociation of homologous, but not identical sequences.

For hybridisations using an imperfectly homologous probe (e.g. *Drosophila* rDNA to detect *Aedes*, rDNA) the stringency of hybridisation was lowered to 6 x SSC, 37°C, 50% formamide. These conditions increase the stability of a 50% GC rich, perfectly

matched hybrid by 9°C relative to the stringent hybridisation conditions described above. (i.e. under these conditions, the temperature of hybridistion is 29°C below the Tm of the duplex).

19. Recovery of DNA from low melting agarose (Ref. [174]

The method described is that of Langridge *et al* (Anal. Biochem. 103, 264-271, 1980) with some modifications. The DNA yielded is suitable for nick-translation to very high specific activity, ligation into vector DNA, or digestion with most types of restriction endonuclease. The method relies on the partitioning of DNA into an organic phase, then re-partitioning into an aqueous phase devoid of Agarose.

Materials

Low-melting agarose: 'Seaplaque' (FMC corporation). Hexadecyltrimethylammonium bromide ('CTAB', Sigma chemicals).

Equilibrate butan-1-ol and distilled water: 150 ml/1150ml and separate phases. Dissolve 1g of CTAB in 100 ml of butanol fraction. Equilibrate this with 100 ml of the aqueous fraction, separate phases.

Method

1. Cut band out of LMA gel. Melt at 65°C for 15'
2. Cool to 37°C. Add 10 µg of tRNA carrier per 200 µl Agarose.
3. Add on equal volume of butanol/CTAB and the same volume of aqueous phase/CTAB, both at 37°C. Invert tube for 5'at 37°C. Spin in prewarmed Eppendorf 'microfuge' for 30 secsonds to separate phases.
4. Remove top (butanol) phase. Re-extract aqueous phase twice with an equal volume of butanol/CTAB as in '3'.
5. Pool the 3 butanol extracts and extract with an equal volume of aqueous phase/CTAB. Recover butanol phase.
6. Extract butanol phase with a quarter volume of 200 mM NaCl. Repeat twice, retaining aqueous phases.
7. Extract pooled aqueous phases at 0°C for 30' with 4 volumes of chloroform. (this removes C-TAB).
8. To aqueous phase add 10 µg of glycogen carrier and then 2.5 volumes of absolute ethanol. Stand at -20°C for 20'.
9. Spin 10'. Wash pellet with 70% ethanol. Dry under vacuum. Redissolve in a small volume of TE buffer.

N.B. This method may also be used to recover RNA from normal Agarose-urea gels by melting at 75°C and extracting with butanol/CTAB at room temperature.

20. Protocol for DNA Sequencing using chain terminating inhibitors (Ref. [104])

Reagents and stock solutions (stored at -20°C)

α^{32}P dCTP label: (Amersham International) 3000 Ci/mmole, 10 mCi/ml radioactive concentration, 3.33 nMol. dCTP concentration.

DNA polymerase; (Amersham International) Klenow fragment of *E. coli* DNA polymerase I. 5.4 units per µl (one unit catalyses the incorporation of 10 nmoles of nucleotides into acid insoluble product in 30 minutes at 37°C).

M13 universal primers; (BCL) sequencing primer: d(TCCCAGTCACGACGT) (15-mer), reverse sequencing primer: d(AACAGCTATGACCATG) (16-mer). Dissolved in TE buffer, pH 7.5, at a concentration of 0.5 pmoles/µl.

10 x DNA polymerase buffer: 500 mM NaCl, 70 mM $MgCl_2$, 70mM Tris-HCl pH 7.5.

Deoxynucleotide triphosphates and dideoxynucleotide triphosphates; (BCL) dissolved in double distilled water to give stock solutions of 25 mM and 10 mM respectively.

Dithiothrietol; (Sigma Chemicals) dissolved in double distilled water to give a 100 mM stock.

Working dNTP and ddNTP solutions (from stock solutions)

dCTP/ddCTP :	0.05 mM/0.05 mM	dTTP :	0.5 mM
ddTTP :	0.5 mM	dCTP :	0.5 mM
ddGTP :	0.25 mM	dGTP :	0.5 mM
ddATP :	0.5 mM	dATP :	0.5 mM

Sequencing mixes

'A' mix : 20 µl 0.5 mM dTTP
1 µl 0.5 mM dATP
20 µl 0.5 mM dGTP
20 µl 10 x DNA polymerase buffer

'C' mix : 20 µl 0.5 mM dTTP
20 µl 0.5 mM dATP
20 µl 0.5 mM dGTP
20 µl 10 x DNA polymerase buffer

'G' mix : 20 µl 0.5 mM dTTP
20 µl 0.5 mM dATP
1 µl 0.5 mM dGTP
20 µl 10 x DNA polymerase buffer

'T' mix : 1 µl 0.5 mM dTTP
20 µl 0.5 mM dATP
20 µl 0.5 mM dGTP
20 µl 10 x DNA polymerase buffer

Procedure

Annealing: 200 ng single-stranded template DNA

1 µl (0.5 pmoles) sequencing primer

1 µl 10 x DNA polymerase buffer

H_2O to 10 µl

Mix, boil for 5 minutes, leave at room temperature for 20 minutes. Add: 1 µl of $\alpha^{32}P$ dCTP (3.33×10^{-15} moles), 1 µl of 100 mM dithiothrietol and 1 µl (5.4 units) of *E. coli* DNA polymerase I (Klenow fragment). (Add these to the side of the tube and spin down to mix). Pipette 3 µl of this mixture into each of the following tubes on ice:

'A' tube : 1 µl 'A' mix + 1 µl 0.5 mM ddATP

'C' tube : 1 µl 'C' mix + 1 µl 0.05 mM ddCTP/dCTP

'G' tube : 1 µl 'G' mix + 1 µl 0.25 mM ddGTP

'T' tube : 1 µl 'T' mix + 1 µl 0.5 mM ddTTP

Incubate sequencing reaction 15' at room temperature, then add 1 µl of 0.5 mM dCTP (as chase) to each tube. Incubate for a further 15' at room temperature. Reactions may be stored frozen at -75°C for up to 3 days or size fractionated through 6% polyacrylamide immediately. Prior to electrophoresis, combine reactions with 6 µl of loading buffer (see polyacrylamide gel electrophoresis below), boil 2' and load 3-4 µl of this mixture immediately.

21. Polyacrylamide gel electrphoresis

Stock solutions and buffers

10 x Tris-borate electrophoresis buffer (TBE)

Tris base (Sigma) : 108 g

Borate (AnalaR) : 55 g

EDTA (AnalaR) : 9.3 g

Make volume up to 1 litre. (pH = 8.3)

6% acrylamide stock

Urea (BRL, enzyme grade) : 288 g

Acrylamide (Electran) : 34.2 g

N.N'-methylenebisacrylamide (Electran) : 1.8 g

Dissolve in 500 ml volume at 37°C. Add 60 ml of 10 x TBE. Make volume up to 600 ml. Store at 4°C in dark.

15% acrylamide stock (250 ml); 33.75 g Acrylamide, 3.75 g bis-acrylamide, 144g urea, 25 ml 10 x T.B.E. buffer.

Loading buffer: 80% deionised formamide, 0.01% xylene cyanol (Electran), 0.01% bromophenol blue (Electran), 50 mM Tris-HCl pH 8.3, 1mM EDTA.

Ammonium persulphate (AnalaR)

Tetramethylethylenediamine 'TEMED' (Electran)

γ-(methacryloxy) propyltrimethoxysilane, 'silane' (Electran)

Dimethyldichlorosilane (2% in 1.1.1. trichloroethane) (BDH)

Preparation of glass plates

Long plate (20 cm x 41.5 cm) : clean with detergent, rinse in distilled water. Coat with dimethyldichlorosilane, polish with distilled water.

Short plate (20 cm x 38.5 cm) : clean with detergent, rinse in distilled water. To 10 ml of ethanol add 50 µl of silane. Mix. To this, add 300 µl of fresh 10% acetic acid. Pour over plate, allow to dry. Polish using ethanol.

Spacers: 0.4 mm. Clean with detergent. Rinse in distilled water. Tape plates together using masking tape.

Preparation of gel

1. Prepare 1 ml of fresh 10% ammonium persulphate solution.
2. Degas 60 ml of Acrylamide stock in-vacuo at 20°C for 2 minutes
3. To Acrylamide stock, add 0.5 ml of 10% ammonium persulphate and 28 µl of TEMED. Mix well.
4. Pour gel immediately (e.g. at 45° inclination, to avoid air bubbles). Insert well-former. Clamp using 4 bulldog clips. Leave 1 hour to set.
5. Trim excess polyacrylamide from around well-former.
6. Set up electrophoresis tank. (BRL S1). Remove wellformer and rinse out wells with TBE.

Running gel

Pre-electrophorese (e.g. 30', 1500 volts, Chandos P.S.U.) to obtain 50-55°C gel temperature. Rinse out wells with TBE. Load. Run at 50-55°C (e.g. 1000-1200 V).

Fixing and drying gel

Separate plates. Fix gel in 10% acetic acid for 15 minutes. Rinse in tap water for 30 minutes to wash out urea. Place in 80°C drying oven for 1½-2 hours until completely dry. Autoradiograph overnight at room temperature, without intensifying screen.

22. Growth of M13 Bacteriophage

Plates and media

Enriched broth (EB); 25g Tryptone, 7.5g yeast extract, 6g NaCl, 1g dextrose, 50 ml 1 M Tris-HCl pH 7.5 per litre.

'X-gal' plates; Prepare fresh 'X-gal' (2o mg/ml in diethyl formamide) and IPTG (24 mg/ml in H_2O). Add 25 µl of each to 3 ml of molten top Agar (see Materials and Methods 10) at 50°C. Add plating cells/M13 mixture, mix and pour onto a Tryptone Agar plate prewarmed to 37°C. Recombinant M13 plaques are colourless whereas non-recombinant plaques are blue in this assay.

Host cells; *E. coli* JM103 (containing F' episome required for infection of M13 via the sex pilus, and also required for growth on minimal media).

Plating cells: Take single small colony of JM103 from minimal agar plate and use to inoculate a 20 ml E-broth culture. Grow to $O.D._{600nm}$ of 0.6 (approx. 6×10^7 cells/ml).

Preparation of double stranded replicative form (RF) DNA

1. To 200 ml of plating cells, add M13 phage from a titred stock to give a multiplicity of infection of 50.
2. Incubate 4 hours - overnight at 37°C, shaking.
3. Harvest cells and continue as from stage '3' of a large-scale plasmid preparation (Materials and Methods 6). Retain supernatent for preparation of single-stranded DNA (see below) or to keep as high titre stock (e.g. 10^{11}-5×10^{12} pfu/ml).

Expected yield: 50-200 µg of duoble stranded DNA.

Preparation of single-stranded M13 DNA

High purity single-stranded M13 DNA may be prepared from the filamentous phage liberated from infected cells into the growth medium (e.g. from the supernatent from step '3' above).

1. Precipitate phage by addition of 0.2 volumes of 25% PEG 6000, 2.5 M NaCl, followed by incubation on ice for 1 hour or overnight at 4°C.
2. Spin down precipitate (e.g. 10,000 rpm, 10 minutes, 4°C, GSA rotor).
3. Resuspend pellet in 10 ml of TE buffer. Phenol extract three

times, chloroform extract once.

4. Ethanol precipitate DNA. Redissolve in TE buffer.

Expected yield: approximately 1 mg of DNA per 200 ml culture (3 x 10^{11} phage particles contain 1 µg of single-stranded DNA).

M13 minipreparation

1. Take a single well isolated colourless plaque and add contents (slow-growing, recombinant infected cells) to 2 ml of EB.
2. Incubate overnight at 37°C with shaking
3. Proceed as from step '5' of the 'rapid, small-scale isolation of bacteriophage λ DNA' protocol. This should yield approximately 5 µg of single-stranded DNA suitable for DNA sequencing.

Supernatent from step '5' of this protocol may be used to give a high titre phage stock. The pellet from stage '5' above may be used to prepare double-stranded (RF) DNA using the 'medium size, quick plasmid preparation' protocol, scaled-down appropriately.

PAGE

MISSING

IN

ORIGINAL

REFERENCES

1. Davidson, G. Genetic Control of Insect Pests. (1974) Academic Press Pub.
2. Engels, W.R. (1983) Ann. Rev. Biochem. 17, 315-344. The 'P' family of transposable elements in *Drosophila*.
3. Knight, K.L. and Stone, A. (1977) Suppl. Ent. Soc. Am. Vol VI A catalogue of the mosquitoes of the world.
4. Knight, K.L. (1978) Suppl. Ent. Soc. Am. A supplement to a catalogue of mosquitoes of the world.
5. Mattingley, P.F. (1957) Ann. Trop. Med. Parasit. 51, 392-398 and 52, 5-17. General aspects of the *Aedes aegypti* problem.
6. Christophers, S.R. (1960) *Aedes aegypti* (L.) The Yellow Fever Mosquito. Cambridge University Press.
7. Gutzevich, A.V. (1931) Mag. Parasit. Leningrad. 2, 35-54. The reproduction and development of the yellow fever mosquito under experimental conditions.
8. Craig, C.B. (1967) Science 150, 1499-1501. Mosquitoes: female monogomy is induced by male accessory gland substance.
9. Kershaw, W.E., Chalmers, T.A. and Lavoipierre, M.M.J. (1954). Ann. Trop. Med. Parasit. 48, 442-450. Studies on arthropod survival I. The pattern of mosquito survival in laboratory conditions.
10. Detinova, T.S. (1968) Ann. Rev. Ent. 13, 427-450. Age structure of insect populations of medical importance.
11. McDonald, P.T. (1977) J. Med. Ent. 14, 42-48. Population

characteristics of domestic *Aedes aegypti* (Diptera: Culicidae) in villages on the Kenyan coast I. Adult survivorship and population size.

12. Pandit, C.J. (1971). Ind.J. Med. Res. 59, 1523-1547. India and the yellow fever problem.

13. Tashima, M., Calabretta, B., Torelli, G., Scofield, M., Maizel, A. and Saunders, G.F. (1981) Proc. Natl. Acad. Sci. USA 78, 1508-1512. Presence of a highly repetitive and widely dispersed DNA sequence in the human genome.

14. Crampton, J.M., Humphries, S., Woods, D. and Williamson, R. (1980) Nucl. Acids. Res. 8, 5157-5168. Isolation of cloned cDNA sequences which are differentially expressed in human lymphocytes and fibroblasts.

15. Wichman, H.A., Potter, S.S. and Pine, D.S. (1985) Nature, 317, 77-81. Mys: a family of mammalian transposable elements, isolated by phylogenetic screening.

16. Leder, P., Tiemeier, P.D. and Enquist, L. (1977) Science 196, 175-177. EK2 derivatives of bacteriophage λ: The λgtWES system.

17. Botchan, M., McKenna, G. and Sharp, P.A. (1974) Cold Spring Harbour Symp. Quant. Biol. 38, 383-395. Cleavage of mouse DNA by a restriction enzyme as a clue to the arrangement of genes.

18. Bingham, P.M., Levis, R. and Rubin G.M. (1981) Cell 25, 693-704. Cloning DNA sequences from the white locus of *D. melanogaster* by a novel and general method.

19. Clarke, L. and Carbon, J. (1976) Cell 9, 91-. A colony bank containing synthetic ColE1 hybrid plasmids representing the entire genome.

20. Hohn, B. (1979) In: Methods in Enzymology 68, 299-309. *In*

vitro packaging of λ and cosmid DNA.

21 Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) In: Molecular Cloning; A Laboratory Manual p. 286-288. Cold Spring Harbor Laboratory pubs.

22. Phillips, S.L., Casavant, N.C., Hutchison, C.A. and Edgell, M.H. (1985) An empirical method for the evaluation of the quality of genomic DNA libraries. Nucl. Acids Res. 13, 2699-2708.

23. Jelinek, W.R. and Schmid, C.W. (1982) Ann. Rev. Biochem. 51, 813-844. Repetitive sequences in eukaryotic DNA and their expression.

24. Graham, D.E., Neufeld, B.R., Davidson, E.H. and Britten, R.J. (1974) Cell 1, 127-138. Interspersion of repetitive and non-repetitive DNA sequences in the sea urchin genome.

25. Davidson, E.H., Hough, B.R., Amenson, C.S. and Britten, R.J. (1973) J. Mol. Biol. 77, 1-23. General interspersion of repetitive with non-repetitive sequence elements in the DNA of *Xenopus*.

26. Jelinek, W.R., Toomey, T.P., Leinwand, L., Duncan, C.H., Biro, P.A., Choudary, P.V., Weissman, S.M., Rubin, G.M., Houck, C.M. Deininger, P.L. and Schmid, C.W. (1980) Proc. Natl. Acad. Sci. USA 77, 1398-1402. Ubiquitous, interspersed repeated sequences in mammalian genomes.

27. Crain, W.R., Eden, F.C., Pearson, W.R., Davidson, E.H. and Britten, R.J. (1976) Chromosoma (Berl.) 56, 309-326 Absence of short period interspersion of repetitive and non-repetitive sequences in the DNA of *Drosophila melanogaster*.

28. Crain, W.R., Davidson, E.H. and Britten, R.J. (1976) Chromosoma (Berl.) 59, 1-12. Contrasting patterns of DNA

sequence arrangement in *Apis mellifera* and *Musca domestica*.

29. Wells, R., Roger, H.D. and Hollenberg, C.P. (1976). Mol. Gen. Genet. 147, 45-51. Non-*Xenopus*-like DNA sequence organisation in the *Chironomus tentans* genome.

30. Singer, M.F. (1982) Cell 28, 433-434. SINES and LINES: highly repeated short and long interspersed sequences in mammalian genomes.

31. Young, M.W. (1979) Proc. Natl. Acad. Sci. USA 76, 2274-2278. Middle repetitive DNA: a fluid component of the *Drosophila* genome.

32. Spradling, A.C. and Rubin, G.M. (1981) Ann. Rev. Genet. 15, 219-264. *Drosophila genome*: conserved and dynamic aspects.

33. Jelinek, W.R. and Haynes, S.R. (1982) Cold Spring Harbor Symp. Quant. Biol XLVIII, 1123-1130. The mammalian 'Alu' family of dispersed repeats.

34. Hattori, M., Kuhara, S., Takenaka, O. and Sakaki, Y. (1986) Nature 321, 625-628. L1 family of repetitive DNA sequences may be derived from a sequence encoding a reverse-transcriptase related protein.

35. Finnegan, D.J., Rubin, G.M., Young, M.W. and Hogness, D.S. (1978) Cold Spring Harbour Sym, Quant. Biol. 42, 1053-1063. Repeated gene families in *Drosophila melanogaster*.

36. Robertson, H.D. and Dickson, E. (1984). Mol. Cell. Biol. 4, 310-316. Structure and distribution of 'Alu' family sequences and their analogs within hnRNA of HeLa, KB and L cells.

37. Duncan, C.H., Jagadeeswaran, P., Wang, R.R.C. and Weissman, S.M. (1981) Gene 13, 185-196. Structural analysis of templates and

RNA polymerase III transcripts of 'Alu' family sequences interspersed among the human β-like goblin genes.

38. Walter, P. and Blobel, G. (1982) Nature 299, 691-698. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum.

39. Jelinek, W.R. and Darnell, J.E (1972) Proc. Natl. Acad. Sci. USA 69, 2537-2541. Double stranded regions in heterogeneous nuclear RNA from HeLa cells.

40. Elder, J.T., Pan, J., Duncan, C.H. and Weissman, C.H. (1981) Nucl. Acids Res. 9, 1171-1189. Transcriptional analysis of interspersed repetitive polymerase III transcription units in human DNA.

41. McColl, R.S. and Aronson, A.I. (1974) Biochem. Biophys. Res. Comm. 56, 47-51. Transcription from unique and redundant DNA sequences in sea urchin embryos.

42. Wetmur, J.G. and Davidson, N. (1968) J. Mol. Biol. 31, 349-370. Kinetics of renaturation of DNA.

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

44. Bishop, J.O. (1972) Biochem J. 126, 171-185. Molecular hybridisation of ribonucleic acid with a large excess of deoxyribonucleic acid.

45. Greenberg, J.R. and Perry, R.P. (1971) J. Cell. Biol. 50, 774-786. Hybridisation properties of DNA sequences directing the synthesis of mRNA and hnRNA.

46. Spradling, A., and Penman, S. (1974) Cell 3, 23-30. Repetitous and unique sequences in the heterogenous nuclear and cytoplasmic messenger RNA of mammalian and insect cells.

47. Varma, M.G.R. and Pudney, M. (1969) J. Med. Ent. 6, 432-439. The growth and serial passage of cell lines from *Aedes aegypti* (L.) larvae in different media.

48. Gergen, J.P., Stern, R.H. and Wensink, P.C. (1979) Nucl. Acids. Res. 7, 2121-2136. Filter replicas and permanent collections of recombinant DNA plasmids.

49. Hogan, B. (1975) New Scientist 67 (No. 966), p. 575 Uniqueness and repetition in the organisation of genes.

50. Firtel, R.A. and Kindle, K. (1975) Cell 5, 401-411. Structural organisation of the genome of the cellular slime mold *Dictyostelium discoideum*; interspersion of repetitive and single copy sequences.

51. Efstradiatis, A., Crain, W.R., Britten, R.J., Davidson, E.H. and Kafatos, F.C. (1976) Proc. Natl. Acad. Sci. USA 73, 2289-2293. DNA sequence organisation in the lepidopteran *Antheraea*.

52. Lewin, B. (1985) Genes II 304-311 (J. Wiley and Sons).

53. Tschudi, C. and Pirrotta, V. (1982). The 5S RNA genes of *Drosophila melanogaster*. In: The Cell Nucleus XI, rDNA, Part B, 29-44. Eds. Bush, H. and Rothblum, L. (Academic Press).

54. Brown, D.D. & Dawid, I.B. (1968). Science 160, 272-280. Specific gene amplification in oocytes.

55. Gall, J.G. and Pardue, M.L. (1969) Proc. Natl. Acad. Sci. USA 63, 378-383. Formation and detection of RNA-DNA hybrid molecules in cytological preparations.

56. Renkawitz, R., Gerbi, S.A. and Glatzer, K.H. (1979). Mol. Gen. Genet. 173, 1-13. Ribosomal DNA of the fly *Sciara coprophila* has a very small and homogeneous repeat unit.

57. Degelman, A., Royer, H. and Hollenberg, C. (1979). Chromosoma 71, 263-281. The organisation of the ribosomal RNA genes of *Chironomus tentans* and some closely related species.

58. Maizels, N. (1976) Cell 9, 431-438. Dictyostelium 17S, 25S and 5S rDNAs lie within a 38 Kb repeat unit.

59. Wellauer, P.K. & David, I.B. (1979). J. Mol. Biol. 128, 289-303. Isolation and sequence organisation of human ribosomal DNA.

60. Kominami, R., Urano, Y., Mishima, Y. and Muramatsu, M. (1981) Nucl. Acids. Res. 9, 3219-3233. Organisation of rRNA gene repeats of the mouse.

61. Klootwijk, J., de Jonge, P. and Planta, R.J. (1979). Nucl. Acids Res. 6, 27-39. The primary transcript of the ribosomal repeating unit in yeast.

62. Dawid, I.B., Wellauer, P.K. and Long, E.O. (1978). J. Mol. Biol. 126, 749-769. Ribosomal DNA in *D. melanogaster* I : Isolation and characterisation of cloned fragments.

63. Wellauer, P.K. and Dawid, I.B. (1974) J. Mol. Biol. 89, 379-395. Secondary structure maps of ribosomal RNA and DNA I: Processing of *X. laevis* rDNA and structure of single-stranded rDNA.

64. Wellauer, P.K. and Dawid, I.B. (1973) Proc. Natl. Acad. Sci. USA. 70, 2827-2831. Secondary structure maps of RNA: processing of HeLa ribosomal RNA.

65. Shine, J. and Dalgarno, L. (1973) J. Mol. Biol. 75, 57-72. Occurrence of heat-dissociable ribosomal RNA in insects: The presence of three polynucleotide chains in 26S RNA from cultured *Ae. aegypti* cells.

66. Jordan, B.R., Jourdan, R. and Jacq, B. (1976) J. Mol. Biol. 101, 85-105. Late steps in the maturation of *Drosophila* 26S rRNA: generation of 5.8S and 2S RNAs by cleavage in the cytoplasm.

67. Gourse, R.L. and Gerbi, S.A. (1980) J. Mol. Biol. 140, 321-339. Fine structure of ribosomal RNA, III, Location of evolutionarily conserved regions within rDNA.

68, Brown, D.D. and Blackler, A.W. (1972) J. Mol. Biol. 63, 53-73. A comparison of the ribosomal DNAs of *Xenopus laevis* and *Xenopus mulleri*; the evolution of tandem genes.

69. Moss, T. (1983) Nature 302, 223-228. A transcriptional function for the repetitive ribosomal spacer in *Xenopus laevis*.

70. DeWinter, R.F.J. and Moss, T. (1986). Cell 44, 313-318. Spacer promoters are essential for efficient enhancement of *X. laevis* ribosomal transcription.

71. Schaffer, M. and Kunz, W. (1985) Nucl. Acids. Res. 13, 1251-1266. rDNA in *Locusta migratoria* is very variable; two introns and extensive restriction site polymorphism in the spacer.

72. Wellauer, P.K., Dawid, I.B., Brown, D.D. and Reeder, R.H. (1976) J. Mol. Biol. 105, 461-486. The molecular basis for length heterogeneity in rDNA from *X. laevis*.

73. Glover, D.M. and Hogness, D.S (1977) Cell 10, 167-176. A novel arrangement of the 18S and 28S sequences in a repeating unit of *Drosophila melanogaster* rDNA.

74. Barnett, T. and Rae, P.M.M. (1979) Cell 16, 763-775. A 9.6 Kb intervening sequence in *D. virilis* rDNA and sequence homology in rDNA interruptions of diverse species of *Drosophila* and

other diptera.

75. Beckingham, K. and White, R. (1980). J. Mol. Biol. 137, 349-373. The ribosomal DNA of *Calliphora erythrocephala*; an analysis of hybrid plasmids containing rDNA.

76. French, C.K., Fouts, D.L. and Manning, J.E. (1981) Nucl. Acids. Res. 9, 2563-2576. Sequence arrangement of the rRNA genes of the dipteran *Sacrophaga bullata*.

77. Lecanidou, R., Eickbush, T.H. and Kafatos,F.C. (1984) Nucl. Acids. Res. 12, 4703-4713. Ribosomal DNA genes of *Bombyx mori*; a minor fraction of the repeating units contain insertions.

78. Renkawitz-Pohl, R., Matsumoto, L. and Gerbi, S.A. (1981) Nucl. Acids. Res. 9, 3747-3764. Two distinct intervening sequences in different ribosomal DNA repeat units of *Sciara coprophila* (fungus fly).

79. Roiha, H. and Glover, D.M. (1981) Nucl. Acids. Res. 9, 5521-5532. Duplicated rDNA sequences of variable lengths flanking the short type I insertions in the rDNA of *Drosophila melanogaster*.

80. Kidd, S.J. and Glover, D.M. (1980) Cell 19, 103-119. A DNA segment from *D. melanogaster* which contains five tandemly repeating units homologous to the major rDNA insertion.

81. Browne, M.J., Read, C.A., Roiha, H. and Glover, D.M. (1984) Nucl. Acids. Res. 12, 9111-9122. Site specific insertion of a type I rDNA element into a unique sequence in the *Drosophila melanogaster* genome.

82. Dawid, I.B. and Rebbert, L. (1981) Nucl. Acids. Res. 9, 5011-5020. Nucleotide sequences at the boundaries between gene and insertion regions in the rDNA of *Drosophila melanogaster*.

83. Campbell, G.R., Littau,V.C., Melera, P.W., Allfrey, V.G. and Johnson, E.M. (1979) Nucl. Acids. Res. 6, 1433-1447. Unique sequence arrangement of ribosomal genes in the palindromic rDNA molecule of *Physarum polycephalum*.

84. Wild, M.A. and Gall, J.B. (1979) Cell 16, 565-573. An intervening sequence in the gene coding for 25S ribosomal RNA in *Tetrahymena pigmentosa*.

85. Long, E.O. and Dawid I.B. (1979) Cell 18, 1185-1196. Expression of rDNA insertions in *D. melanogaster*.

86. Roiha, H., Miller, J.R., Woods, L.C. and Glover, D.M. (1981) Nature 290, 749-753. Arrangements and rearrangements of sequences flanking the two types of rDNA insertion in *D. melanogaster*.

87. Renkawitz-Pohl, R., Glatzer, K.H. and Kunz, W. (1981) J. Mol. Biol. 14, 95-101. Ribosomal RNA genes with an intervening sequence are clustered within the 'X' chromosome rDNA of *D. hydei*.

88. Beckingham, K. and Rubacha, A. (1984) Chromosoma (Berl.) 90, 311-316. Different chromatin states of the intron$^-$ and type 1 intron$^+$ rRNA genes of *Calliphora erythrocephala*.

89. Wayne, R.L., Sharp, Z.D. and Procunier, J.D. (1985). Nucl. Acids. Res. 13, 2869-2879. Preferential DNase sensitivity of insert-free ribosomal RNA repeats of *Drosophila melanogaster*.

90. Tartof, K.D. (1979) Cell 17, 607-614. Evolution of transcribed and spacer sequences in the ribosomal genes of *Drosophila*.

91 McDonald, P.T. and Rai, K.S. (1970) Genetics 66, 475-485. Correlation of linkage groups with chromosomes in the mosquito *Aedes aegypti*.

92. Hasan, G., Turner, M.J. and Cordingly, J.S. (1984) Cell 37, 333-341. Complete nucleotide sequence of an unusual mobile element from *T. brucei*.

93. Vincent, A. and Petes, T.D. (1986) Nucl. Acids. Res. 14, 2938-2949. Isolation and characterisation of a Ty element inserted into the rDNA of the yeast *Sachromyces cerevisiae*.

94. Bell, J.R., Bogardus, J.M., Schmidt, T. and Pellegrini, M. (1985) Nucl. Acids. Res. 3861-3871. A new copia-like transposable element found in a *Drosophila* rDNA gene unit.

95. Bosely, P.G., Tuyns, A. and Birnstiel, M.L. (1978) Nucl. Acids Res. 5, 1121-1137. Mapping of the *Xenopus laevis* 5.8S rDNA by restriction and DNA sequencing.

96. Bosely, P.G., Moss, T., Machler, M., Portmann, R. and Birnstiel, M. (1979) Cell 17, 19-31. Sequence organisation of the spacer DNA in a ribosomal gene unit of *Xenopus laevis*.

97. Beckingham, K. (1982) Insect rDNA. In: The Cell Nucleus Vol. 10 p212-213. (Academic Press Inc.) Eds. Busch, H. and Rothblum, L.

98. Manning, R.F., Samols, D.R. and Gage, P.L. (1978) Gene 4, 153-166. The genes for 18S, 5.8S and 28S ribosomal RNA of *Bombyx mori* are organised into tandem repeats of uniform length.

99. Yavachev, L.P., Georgiev, O.I., Bragen, E.A., Avdonina, T.A., Bogomolova, A.C., Zhurkin, V.B., Nosikov, V.V. and Hadjiolov, A.A. (1986) Nucl. Acids. Res. 14, 2799-2810. Nucleotide sequence analysis of the spacer regions flanking therat rRNA transcription unit and identification of repetitive elements.

100. Wu, C., Bingham, P.M., Livak, K.J., Holmgren, R. and Elgin, S.C.R. (1979) Cell 16, 797-806. The chromatin structure of specific genes: 1. Evidence for higher order domains of

defined DNA sequence.

101. Rae, P.M.M. and Steele, R.E. (1979) Nucl. Acids Res. 6, 2987-2995 Lack of methylation of insect DNA. .

102. Adams, R.L.P., McKay, E.L., Craig, L.M. and Burdon, R.H. (1979) Biochem. Biophys. Acta. 563, 72-81. Methylation of mosquito DNA.

103. Viera, J. and Messing, J. (1982) Gene 19, 259-268. The pUC plasmids, an M13 mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers.

104. Sanger, F, Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467. DNA sequencing with chain-terminating inhibitors.

105. Messing, J. (1983) Methods in Enzymology 101, 20-78. New M13 vectors for cloning.

106. Salim, M. and Maden, B.E.H. (1981) Nucleotide sequence of the *Xenopus laevis* 18S ribosomal RNA inferred from gene sequence. Nature 291, 205-208.

107. Ware, V.C., Tague, B.W., Clark, C.G., Gourse, R.L., Brand, R.C. and Gerbi, S.A. (1983) Sequence analysis of 28S ribosomal DNA from the amphibian *Xenopus laevis*. Nucl. Acids Res. 11, 7795-7817.

108. Pavlakis, G.N., Jordan, B.R., Wurst, R.M. and Vournakis, J.N. (1979) Nucl. Acids Res. 7, 2213-2238. Sequence and secondary structure of *Drosophila melanogaster* 5.8S and 2S rRNAs and of the processing site between them.

109. Jordan, B.R., Latil-Damotte, M. and Jourdan, R. (1980) FEBS Letters 117, 227-231. Sequence of the 3' terminal portion of *Drosophila melanogaster* 18S rRNA and of the adjoining spacer.

110. Samols, D.R., Haganbuchle, O. and Gage, L.P. (1979) Nucl. Acids Res. 7, 1109-1119. Homology of the 3' terminal sequences of the 18S rRNA of *Bombyx mori* and the 16S rRNA of *Escherichia coli*.

111. Hasegawa, M., Iida, Y., Taka-aki, Y., Takaiwa, F. and Iwabuchi, M. (1985) J. Mol. Evol. 22, 32-38. Phylogenetic relationships among eukaryotic Kingdoms inferred from ribosomal RNA sequences.

112. Breathnach, R. and Chambon, P. (1981) Ann. Rev. Biochem, 50, 349-383. Organisation and expression of eukaryote split genes encoding proteins.

113. Berk, A.J. and Sharp, P.A. (1977) Cell 12, 721-732. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids.

114. Maden, B.E.H. (1982) In: The Cell Nucleus vol. 10, p342-346. Eds. Busch, H. and Rothblum, L. (Academic Press, New York).

115. Service, M.W. (1982) Bull. Soc. Vector. Ecol. 7, 1-13. Importance of vector ecology in vector disease control in Africa.

116. Davidson, G., Patterson. H.E., Coluzzi, M., Mason, G.F. and Micks, D.W. (1967). The *Anopheles gambiae* complex. In: Genetics of Insect Vectors of Disease (Wright, J.W. and Pal, R. eds.). pp 211-250, Elsevier, Amsterdam.

117. Davidson, G. (1966) *Anopheles gambiae*, a complex of species Bull. Wld. Hlth. Org. 31, 625-634.

118. Davidson, G. and Hunt, R.H. (1973). The crossing and chromosome characteristics of a new, sixth species in the *Anopheles gambiae* complex. Parasitologia, 15, 121-128.

119. Ribbands, C.R. (1944) Differences between *An. melas* (*An. gambiae* var. *melas*) and *An. gambiae*. Annals Trop. Med.

Parasitol. 38, 85-99.

120. Muirhead-Thompson, R. (1948) Studies on *An. gambiae* and *An. melas* in and around Lagos. Bull. Ent. Res. 38, 527-558.

121. Davidson, G. (1956). Insecticide resistance in *An. gambiae* Giles: a case of simple Mendelian inheritance. Nature 178, 705-706.

122. White, G.B. (1974). *Anopheles gambiae* complex and disease transmission in Africa. Trans. Royal Soc. Trop. Med. Hyg. 68, 278-298.

123. Pajot, F. and Bailly-Choumara, H. (1963) Unpublished report WHO/Mal/68.660.

124. Highton, R.B., Bryan, J.H., Boreham, P.F.L. and Chandler, J.A. (1979) Studies on the sibling species *Anopheles gambiae* Giles and *Anopheles arabiensis* Patton (Diptera: Culicidae) in the Kisumu area, Kenya. Bull. Ent. Res. 69, 43-53.

125. Service, M.W. (1970) Ecological notes on species A & B of the *An. gambiae* complex in the Kisumu area of Kenya. Bull. Ent. Res. 60, 105-108.

126. Service, M.W. (1978) A survey of *Anopheles gambiae* (speciesA) and *An arabiensis* (species B) of the *An. gambiae* Giles complex in the Kisumu area of Kenya following insecticidal spraying with OMS-43 (Fenitrothion). Annals. Trop. Med. Parasitol. 72, 377-386.

127. The Garki project. Research on the epidemiology and control of malaria in the Sudan savanna of West Africa. 1980 WHO, Geneva, p311.

128. Coluzzi, M., Sabatini, A., Petrarca, V. and DiDeco, M.A. (1979) Chromosomal differention and adaptation to human environments in

the *Anopheles gambiae* complex. Trans. Royal Soc. Trop. Med. Hyg. 73, 483-497.

129. Davidson, G. (1956) Insecticide resistance in *An. gambiae* Giles. Nature 178, 705-706.

130. Davidson, G. and Hamon, J. (1962) A case of dominant dieldrin resistance in *Anopheles gambiae* Giles. Nature 196, 1012.

131. White, G.B., Tessfaye, F., Boreham, P.F.L. and Lemma, G. (1980) Malaria vector capacity of *An. arabiensis* and *An. quadriannulatus* in Ethiopia. Trans. Royal Soc. Trop. Med. Hyg. 74, 683-684.

132. O. Mastabaum (1957). Past and present position of malaria in Swaziland. J. Trop. Med. Hyg. 60, 190-

133. Brink, C.J.H. (1958) S. Afr. Med. J. 32, 800.

134. White, G.B. (1973). Comparative studies on sibling species of the *Anopheles gambiae* Giles complex (Diptera, Culicidae) III. The distribution, ecology, behaviour and vectorial importance of species D in Bwamba County, Uganda with an analysis of biological, ecological, morphological and cytogenetical relationships of Ugandan species D. Bull. Ent. Res. 63, 65-97.

135. Patterson, H.E., Patterson, J.S. and Van Eden, G.J. (1964) Records of the breeding of saltwater *An. gambiae* at inland localities in South Africa. Nature 201, 524-525.

136. Mosha, F.W. and Petarca (1983) Ecological studies on *Anopheles gambiae* complex sibling species on the Kenyan coast. Trans. Royal. Soc. Trop. Med. Hyg. 77, 344-345.

137. Bushrod, F.M. (1981) The *Anopheles gambiae* Giles complex and Bancroftian filariasis transmission in a Tanzanian coastal village. Annals Trop. Med. Parasitol. 75, 93-100.

138. Bryan, J.H. (1979) Observations on the member species of the *Anopheles gambiae* complex in The Gambia, West Africa. Trans. Royal. Soc. Trop. Med. Hyg. 73, 463-466.

139. Bryan, J.H. (1981) *Anopheles gambiae* and *A. melas* at Brefet, The Gambia, and their role in malaria transmission. Annals Trop. Med. Parasitol. 77, 1-12.

140. Patterson, H.E. (1964) Direct evidence for the specific distinctness of forms A, B & C of the *Anopheles gambiae* complex. Riv. Malar. 44, 191-196.

141. White, G.B. (1972) Comparative studies on sibling species of the *Anopheles gambiae* Giles complex (Diptera, Culicidae): bionomics and vectorial activity of species A and species B at Segera, Tanzania. Bull. Ent. Res. 62, 295-317.

142. Robertson, J.D. (1945) Notes on the gametocyte threshold for infection of *Anopheles gambiae* Giles 1902 and *Anopheles melas* Theobald, 1903 in West Africa. Trans. Royal Soc. Trop. Med. and Hyg. 39, 8-10.

143. Burgess, R.W. (1960) Comparative susceptibility of *Anopheles gambiae* Theobald and *Anopheles melas* Giles to infection by *Plasmodium falciparum* in Liberia, West Africa. Am. J. Trop. Med. and Hyg. 9, 652-655.

144. Gelfand, H.M. (1955) *An. gambiae* Giles and *An. melas* Theobald in a coastal area of Liberia, West Africa. Trans. Royal. Soc. Tropi. Med. and Hyg. 49, 508-527.

145. Coluzzi, M. and Sabatini, A. (1967) Cytotaxonomic observations on species C of the *Anopheles gambiae* complex. Parasitologia, 10, 155-165.

146. Coluzzi, M. and Sabatini, A. (1968) Cytogenetic observations

on species C of the *Anopheles gambiae* complex. Parasitologia 10, 155-165.

147. Coluzzi, M. and Sabatini, A. (1969) Cytogenetic observations on the saltwater species *An. merus* and *An. melas* of the *gambiae* complex. Parasitologia 11, 177-187.

148. Green, C.A. (1972) Cytological maps for the identification of females of the three freshwater species of the *Anopheles gambiae* complex. Annals Trop. Med. Parasitol. 66, 143-147.

149. Davidson, G. and Hunt, R.H. (1973) The crossing and chromosome characteristics of a new, sixth species in the *Anopheles gambiae* complex. Parasitologia 15, 121-128.

150. Mahon, R.J., Green, C.A. and Hunt, R.H. (1976) Diagnostic allozymes for routine identification of adults of the *Anopheles gambiae* complex (Diptera, culicidae). Bull. Ent. Res. 66, 25-31.

151. Miles, S.J. (1978) Enzyme variation in the *Anopheles gambiae* Giles group of species (Diptera: culicidae). Bull. Ent. Res. 68, 85-96.

152. Miles, S.J. (1976) The *Anopheles gambiae* complex: a biochemical key. J. Med. Ent. 15, 297-299.

153. Carlson, D.A. and Service, M.W. (1979) Differentiation between species of the *Anopheles gambiae* complex (Diptera: culicidae) by analsis of the uticular hydrocarbons. Annals Trop. Med. Parasitol. 7, 589-592.

154. Carlson, D.A. and Service, M.W. (1980) Identification of mosquitoes of the *Anopheles gambiae* complex A and B by analysis of the cuticular components. Science 207, 1089-1091.

155. Hamilton, R.J. and Service, M.W. (1983) Value of cuticular and internal hydrocarbons for the identification of larvae of

Anopheles gambiae Giles, *Anopheles arabiensis* Patton and *Anopheles melas* Theobald. Annals. Trop. Med. Parasitol. 77, 203-210.

156. Gatti, M., Santini, G., Pimpinelli, S. and Coluzzi, M. (1977) Fluorescence banding techniques in the identification of sibling species of the *Anopheles gambiae* complex. Heredity 38, 105-108.

157. Bonaccorsi, S., Santini, G., Pimpinelli, S. and Coluzzi, M. (1980) Intraspecific polymorphism of sex chromosome heterochromatin in two species of the *Anopheles gambiae* complex. Chromosoma (Berl.) 76, 57-64.

158. Brutlag, D.L. (1980) Molecular arrangement and evolution of heterochromatic DNA. Ann. Rev. Genet. 14, 121-144.

159. Kennedy, W.P.K. (1984). Novel identification of differences in the kinetoplast DNA of *Leishmania* isolates by recombinant DNA techniques and *in situ* hybridisation. Mol. Biochem. Parasitol. 12, 313-325.

160. Tchen, P., Anxolabhere, D., Nouaud, D. and Periquet, G. (1985) Hybridisation on squashed flies: a method to detect gene sequenes in individual *Drosophila* Analytical Biochem. 150, 415-420.

161. Curtis, C.F. (1982). The mechanism of hybrid male sterility from crosses in the *Anopheles gambiae* and *Glossina morsitans* complexes. In: Recent Development in the Genetics of Insect Disease Vectors (Steiner, W.W.M., Tabachnick, W.J., Rai, K.S and Narang, S. eds). pp 290-312. Stipes, Illinois.

162. Gatti, M. and Pimpinelli, S. (1981). Cytological and genetic analyses of the 'Y' chromosome of *Drosophila melanogaster*. I. Organisation of fertility factors. Chromosoma (Berl.)

163. Pietras, D.F., Bennet, K.L., Siracusa, L.D., Woodworth-Gutai, M.,

Chapman, V,M., Gross, K.W., Kane-Haas, C. and Hastie, N.D. (1983) Construction of a small *Mus. musculus* repetitive DNA library: identification of a new satellite sequence in *Mus. musculus*. Nucl. Acids. Res. 11, 6965-6986.

164. Renz, M. and Kurz, C. (1984) A colorimetric method for DNA hybridisation. Nucl. Acids Res. 12, 3435-3444.

165. Forster, A.C., McInnes, J.L., Skingle, D.C. and Symons, R.H. (1985) Non-radioactive hybridisation probes prepared by the chemical labelling of DNA and RNA with a novel reagent, photobiotin. Nucl. Acids Res. 13, 745-761.

166. Coen, E.S., Thoday, J.M., and Dover, G. (1982) Rate of turnover of variants in the rDNA gene family of *Drosophila melanogaster*. Nature 295, 564-568.

167. Mandel, M. and Higa, A. (1970) Calcium dependent bacteriophage DNA infection. J. of Mol. Biol. 53, 154-.

168. Grunstein, M. and Hogness, D.S. (1975) Colony hybridisation; a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA 72, 3961-3965.

169. Benton, W.D. and Davis, R.W. (1977) Screening λgt recombinant clones by hybridisation to single plaques *in situ*. Science 196, 180-182.

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

171. Rigby, P.W.J., Deickmann, M., Rhodes, C. and Berg, P. Labelling DNA to high specific activity *in vitro* by nick-translation with DNA polymerase I. J. Mol. Biol. 113, 237-251.

172. Van de Sande, J.H., Kleppe, K. and Khorana, H.G. (1973). Kinase

labelling of nucleic acids. Biochemistry 12, 5050-5055.

173. Nucleic acid hybridisation; a practical approach. Eds. Hames, B.D. and Higgins, S.J. (IRL Press).

174. Langridge, J., Langridge, P. and Berquist, P.L. (1980) Extraction of nucleic acids from gels. Anal. Biochem. 103, 264-271.