MOLECULAR ANALYSIS OF THE HUMAN MALARIA PARASITE, *PLASMODIUM FALCIPARUM*

GORDON LANGSLEY

A Thesis presented for the Degree of
Doctor of Philosophy

Department of Molecular Biology
University of Edinburgh

October 1983

CONTENTS

		Page
Chapter 1	Introduction	1
Chapter 2	Materials and Methods	40
Chapter 3	Strain variation characterised by the organisation of a class of repetitive DNA	54
Chapter 4	Isolation and characterisation of rDNA fragments from the genome of *Plasmodium falciparum*	66
Chapter 5	Isolation of putative actin genes from genomic and cDNA libraries of *Plasmodium falciparum*	87
Chapter 6	General Discussion	100
Bibliography	.. ..	105

ABSTRACT

The genomic organisation of the human malaria parasite Plasmodium falciparum has been investigated by the analyses of 3 classes of sequence. The distribution of a repetitive sequence was compared between several clones of a single Thai isolate. The hybridisation patterns displayed to the probed were characteristic of each clone, correlated exactly with all the other parameters of variation, and indicates a high frequency of recombination in P.falciparum. The exception was the invarient pattern presented on EcoR1 digestion, which generated an 11 kb fragment common to all clones. This fragment is suggestive of a transposable element.

A second class of sequence examined was the rRNA genes. These have been demonstrated to be present in approximately 8 copies per haploid genome. This is an unusual low copy number for a eukaryotic organism. The rRNA genes appear to be organised into two classes of unlinked transcription units, and each contains an '18S', '5.8S' and '28S' rRNA gene. The '5.8S' rRNA gene is located between the '18S' and '28S' rRNA genes in a 0.5 kb stretch of A+T rich DNA. The '28S' rRNA gene in at least one of the transcription units is interrupted by approximately 5 kb of non-coding sequence.

Finally, sequences homologous to Dictyostelium discoideum actin DNA have been identified in the genome. The hybridisation patterns to two different D.discoideum actin probes suggests that there may be as many as 5 actin genes, and that 3 of them are clustered on an 11 kb piece of genomic DNA. This DNA has been cloned in phage λ. A 2.8 kb mRNA transcript homologous to

D.discoideum actin has been identified and a cDNA clone isolated. This clone contains a 3.5 kb insert homologous to the *D.discoideum* actin cDNA.

I hereby declare that I alone have composed this Thesis, and that, except where stated, the work presented within is my own.

October 1983

This Thesis is dedicated to my

Mother

(God bless you Ma)

ACKNOWLEDGEMENTS

I would like to thank John Scaife for being much more than an excellent supervisor, rather a good friend and mentor. The whole Scaife group for providing a congenial and convivial working environment. Particularly I would like to thank my friends Mike Goman, Roger Hall and Arve Osland, for being so supportive throughout. John Hyde for answering my many questions whilst patiently putting up with my stream of Malaprops, and Neil Bone for showing me how to grow phage, and for never letting me forget that I am English. I'm indebted to David Walliker and Albert Cornelissen for teaching me parasitology and to Pamela Beattie for the electron micrographs. Thanks also to Kathie for the typing, Alan and Frank for the photographs, and especially Keith Derbyshire for the comradeship during the writing up. Finally, but most importantly, to Sarah, without whom I would not have made it to this point.

ABBREVIATIONS

a.a.	amino acid
A+T	Adenosine plus Thymidine
bp	base pair
kb	kilo base pairs
DNA	deoxyribonucleic acid
RNA	ribonucleic acid
IVS	intervening sequence
ITS	internal transcribed spacer
ETS	external transcribed spacer
NTS	non-transcribed spacer
SSC	standard saline citrate
S	Svedberg, a sedimentation coefficient
s.u.	sub-unit

Some of this work has already been published, or accepted for publication.

Langsley, G. and Goman, M. (1982) 'Initial characterisation of a genomic library of Plasmodium falciparum, Parasitology, 85, XXXVII

Goman, M., Langsley, G., Hyde, J. E., Yankovsky, N. K., Zolg, J. W. and Scaife, J. G. (1982). 'The establishment of genomic DNA libraries for the human malaria parasite P.falciparum and identification of individual clones by hybridisation' Molecular and Biochemical Parasitology, 5, 391-400.

Cornelissen, A. W. C. A., Langsley, G. and Walliker, D. (1983) 'Ribosomal DNA sequences detected in Malaria parasites by cytochemical hybridisation' Biology of the Cell (in press)

Bone, N., Gibson, T., Goman, M., Hyde, J.E., Langsley, G., Scaife, J. G., Walliker, D., Yankovsky, N. K. and Zolg, J. W. (1983) 'Investigation of the DNA of human malaria parasite P.falciparum by in vitro cloning in phage lambda' in 'The Molecular Biology of Parasites' eds. J. Guardida, L. Luzzatto, W. Trager. Raven Press New York.

Hyde, J. E., Goman, M., Hall, R., Osland, A., Hope, I. A., Langsley, G.,Zolg, W. and Scaife, J. G. (1983). 'Messenger RNA from the human malaria parasite Plasmodium falciparum; characterisation, translation studies and construction of a cDNA library'. Molecular and Biochemical Parasitology (in press).

CHAPTER I

INTRODUCTION

1.1 Historical perspective

1.2 Life-cycle of *Plasmodium falciparum*

1.3 Control of malaria

1.3a Anti-vector controls

1.3b Anti-parasite controls

1.4 Development of a vaccine

1.5 Variation in clones of *Plasmodia*

1.6 Ribosomal RNA gene organization

1.7 Actin gene organisation

1.8 Purpose of this thesis

1.1 Historical perspective

Malaria is one of man's oldest diseases, yet today it still presents a threat to the health of some 2,000 million people, and is responsible for the death of about 2 million people annually. The majority of these deaths occur in Africa, where in the tropics malaria is holoendemic and is particularly a threat to the young, killing 1 million children every year. A few years ago the battle against malaria seemed to have been won, but now, despite vigorous anti-malaria campaigns the disease is resurgent and malaria is as much a problem as it ever was (World Health Statistics Annual).

The first noted observations of the relationship between intermittent fevers and marshy areas are attributed to Hipprocates in the 5th century, B.C., but it was not until the 18th century that these fevers took on the Italian name mal'aria; a recognition that they were connected to the foul air of marshes. A treatment for malaria existed long before the cause of the fever was known. At the beginning of the 17th century Jesuit missionaries in Peru noticed that the Indians used a tree bark to treat fevers. This bark, which became known as 'Jesuit's powder', was shown to contain quinine in 1820. However, it was not until 1880 that Laveran, a French Army surgeon in Algeria, observed and described malaria parasites in the red blood cells of man. In 1894 Patrick Manson suggested that Malaria was transmitted by mosquitos, and in 1897 Ronald Ross found a developing parasite in an Anopholean mosquito, linking the mosquito to Malaria (for this Ross in 1902 received the Nobel Prize). Demonstration that the mosquito was the vector for malaria was obtained in 1900 by Manson. He, together with Italian colleagues, transmitted malaria to people in London with

infective mosquitos sent from Italy.

Antimalarial measures in the early part of the 20th century centred on the use of quinine, and anti-mosquito larvae controls. However German losses to malaria in the 1st World War highlighted the need for an alternative to quinine, since they did not have direct access to the drug. The following years in Germany saw the development of parmaquine (1924), mepacrine (1930) and chloroquine (1934). Other anti-malarials were developed in France, Britain and the USA; culminating with the development of pyrimethamine in 1951.

The manufacture in 1942 of dichloro-diphenyl-trichloroethane (DDT), led to what have been described as the 'let us spray years'. The use of insecticides together with anti-malarial prophylactics, gave rise to the hope that malaria could be eradicated, and in 1955 the World Health Organization (WHO) launched a global campaign aimed at the eradication of malaria. This campaign was particularly successful in Europe where malaria has been eliminated in Sicily and Greece, leaving Turkey as the only country where it is actively transmitted today.

In contrast the situation in the rest of the world has worsened rather than improved, and in 1969 the WHO dropped its policy of eradication for one of containment. This failure to control the disease is principally the cumulative result of the development of insecticide resistance in the mosquito and drug resistance in the parasites, compounded by the practical problems involved in operating control procedures over vast areas of a country with very few trained personnel. An inexpensive and long-lasting vaccine is urgently required.

1.2 The life-cycle of *Plasmodium falciparum*

Malaria is caused by a unicellular protozoan parasite, which in the vertebrate host is haploid, and belongs to the genus *Plasmodium*. There are over 100 different species comprising the family *Plasmodiiae*; of which 55 are specific for reptiles, 25 for birds, and 20-25 for a variety of mammals including rodents (Garnham, 1966). The discovery of *P.berghei* in rodents by Vincke in 1948, opened a new way for studies in parasitology, immunology and chemotherapy; and there are now several good rodent laboratory models. There are *ca*. 20 species of *Plasmodia* in primates, and those of the *Aotus* and *Saimiri* monkey again provide good model systems. The four human malarias are listed below and the terms tertian and quartian refer to the periodicity of the fever associated with malaria, and are a reflection of the duration of erythrocytic schizogony (see below).

P.vivax: relapse type, benign, tertian and largely sub-tropical though occurs in S.E. Asia and New Guinea.

P.ovale: relapse, benign and like *P.vivax*, largely confined to Africa.

P.malariae: quartian, benign, cosmopolitan distribution, non-relapse type, and the only one that is also a natural parasite of the monkey.

P.falciparum: malignant, tertian and non-relapse type, largely sub-tropical, and is the great scourge of Africa and S.E. Asia.

P.falciparum if not treated, often causes lethal infections in man, and for this reason is the parasite studied here. Shown below

is the life-cycle of P.falciparum (Fig. 1.1) although it is essentially the same for the other human malarias (for review of the life-cycle of all primate malarias, including P.falciparum see Bray and Garnham, 1982) . There are two factors which I wish to emphasize: (i) the alternation of host-man and mosquito- in the life-cycle; (ii) chronobiology, the time table of events in all stages is highly ordered. Since one of the hosts is the mosquito, a brief summary of its life-cycle is pertinent both here and later when discussing the control of malaria.

The complete life-cycle of the Anopheles mosquito takes 3 weeks to 1 month. Only the female takes a blood feed, after which, she lays her eggs in water, larvae hatch from these eggs and rest on the bottom, coming to the surface to feed. The larvae pupate, and metamorphose into male and female mosquitos. The latter mate once in flight, the female storing the sperm in a spermatheca and taking a blood feed before each egg laying.

The life-cycle of P.falciparum commences when an infected female mosquito takes a blood feed, and along with injection of saliva, injects sporozoites. The sporozoites disappear from the blood stream within 30 minutes, because they are transported to the liver where they enter the hepatocytes (either directly or via Kupffer cells). Once within the hepatocytes, sporozoites round up, and divide mitotically, this process is called ex-erythrocytic schizogony and takes approximately 3 days. The resulting schizonts develop into merozoites which after rupture along with the hepatocytes, are released into the blood stream where they rapidly invade erythrocytes (see Fig. 1). This completes what is known as the 'liver stage' of the life-cycle.

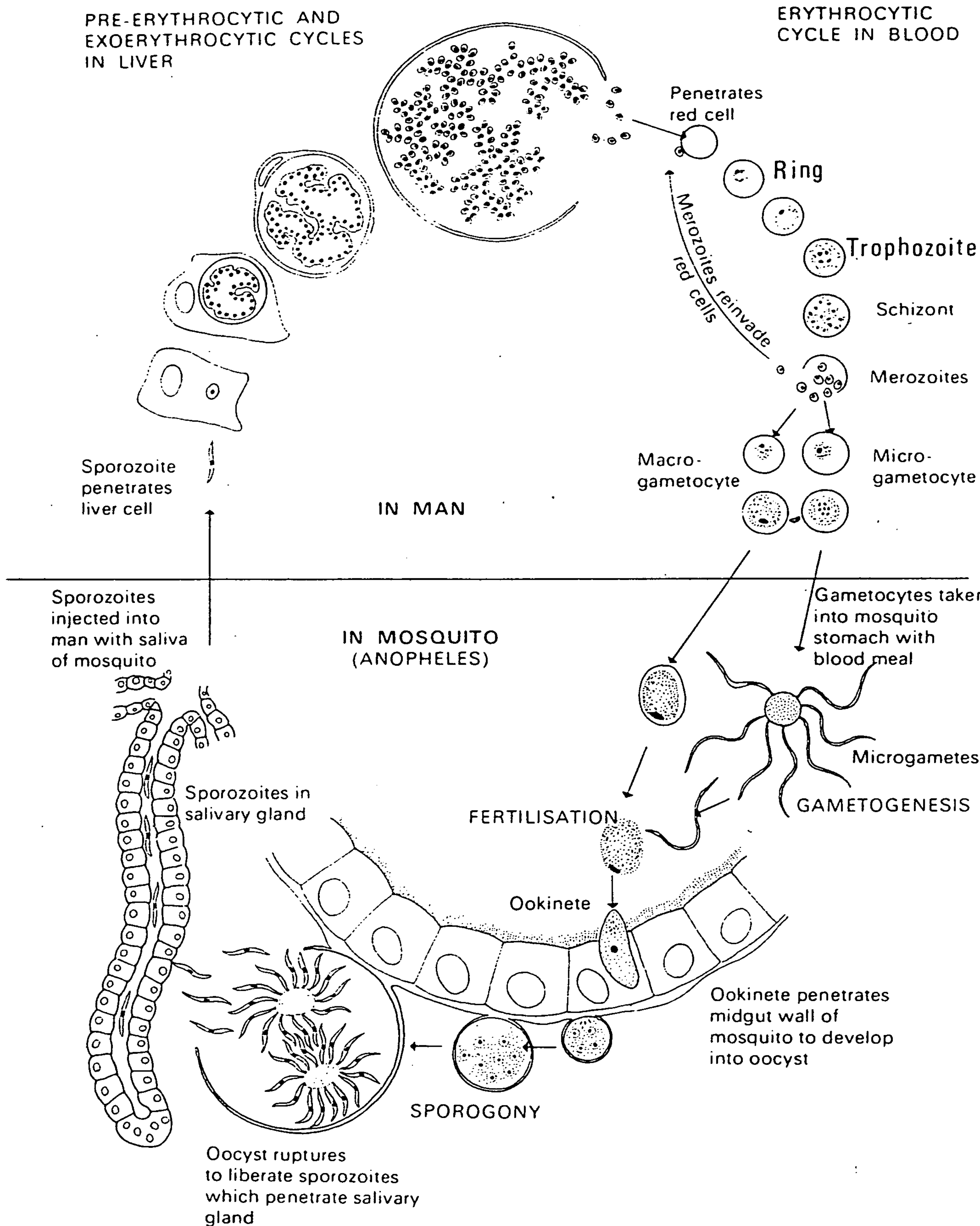


FIG1.1

Life-cycle of Plasmodium falciparum in man and the mosquito. All stages are drawn to approximately the same scale, and the diagram of life-cycle is adapted from Vickerman and Cox (1967)

The merozoites have an apical complex including organelles, which after orientation come into contact with a site on the erythrocyte surface. There are thought to be specific receptors for Plasmodia on the surface of erythrocytes, and for P.falciparum the receptor is believed to be glycophorin A (Pasvol and Wilson, 1982). Merozoites induce invagination, and once inside they are enveloped by what was the erythrocyte plasma membrane, but is now called the parasitophorous vacuole membrane.

The merozoites develop a food vacuole, and although largely parasitic, are capable of some de novo biosynthesis. To compensate for its restricted ability to synthesise amino acids, the parasite has specific proteases which break down host proteins to generate an amino acid pool (Sherman, 1979). Haemoglobin is specifically degraded in this way, and the product heme is harmful to the parasite, and is therefore sequestered (presumably in an innocuous form) in the malaria pigment. The anti-malarial drug chloroquin (see Section 1.3b) is believed to mediate one of its effects by binding to heme, and the heme-chloroquin complex is toxic to the parasite (Orjih, et al., 1981). The parasite can also only synthesise pyrimidines but scavanges for purines. Pyrimidine biosynthesis is the site at which other anti-malarial drugs act, and these are discussed below.

A second round of asexual multiplication occurs within the red blood cell, rings develop into trophozoites which mature into schizonts, followed by segmentation and merozoite release (see Fig. 1.1). This process is called erythrocytic schizogony. The inter-erythrocytic development of the parasite leading to the production of merozoites has a characteristic time period, and for P.falciparum this is 48

hours. After several rounds of asexual multiplication some unknown stimulus triggers the sexual phase. Some of the merozoites develop into gametocytes, either male or female. These male and female gametocytes are infective to mosquitoes. The other merozoites reinvade erythrocytes. The production of gametocytes from merozoites is termed the blood stage of the life-cycle. When a female mosquito feeds on human blood malaria parasites are ingested; following the uptake with the vector blood meal the male produces eight motile micro-gametes and the female a single macrogamete. Fusion of the gametes produces the diploid zygote which rapidly transforms into an ookinete. During the next 24-48 hours the ookinete penetrates the mosquito gut epithehal cells and enters the coelemic cavity, where it encysts forming the oocyst (see Fig. 1). Sporozoites develop within the oocyst (up to 10,000), the oocyst bursts and the sporozoites move to the mosquito's salivary gland. The sporozoites are infective to man, awaiting injection at the next blood feed. Since merozoites are haploid, meiosis must occur at a stage between fertilization and the appearance of merozoites in the erythrocyte, most probably during sporogony. The 'Mosquito stage' and the life cycle are thus completed.

1.3 Control of malaria

The theory of malaria control, as based on mathematical considerations, was first suggested by MacDonald (1957). Control is aimed at achieving

1) reduction in longevity of the mosquito
2) reduction in man-mosquito contact
3) reduction in the longevity of the parasite.

The second can be achieved with repellants and mosquito nets, but are

obviously not widely applicable. The first I shall term anti-vector controls, and the third antiparasite controls; these are discussed below.

1.3(a) Anti-vector controls

The first anti-vector controls to be applied were those directed against the mosquito larvae. These consisted primarily of larvicides such as Paris Green, which coat the surface of water with an oil preventing the larvae from feeding; combined with drainage of the water breeding grounds. However, control by killing the mosquito larvae suffers from two main deficiencies. First, those mosquitoes escaping face no further artificial threat, and some can live long enough to become infective. Secondly, the number of breeding places needing treatment is often astronomically high, especially in rural areas.

The most efficacious approach is to shorten the life of the adult mosquito using insecticide sprays. A daily mortality rate of around 50% is reckoned to be sufficient to break the transmission by most malaria vectors (vector control has been reviewed by Davidson, 1982). This observation is based on the principle of reducing the average life span of the mosquito to below the time taken for the parasite to develop within the mosquito (*ca*. 10 days for *P.falciparum*). From 1945 onwards intensive spraying campaigns eradicated malaria from many temperate areas (the WHO never attempted a large scale spraying campaign in Africa), leaving Turkey as the only European country where malaria is actively transmitted today. Unfortunately insecticide sprays were in such heavy use in agriculture that as early as 1953 some *Anopheles* mosquito vectors were resistant to DDT. Resistance to

insecticides has now spread to over 40 species of Anopheles in 60 countries, and some vector species are resistant to a number of different insecticides.

The resistance of the mosquito vector to most insecticides and the enormous cost of developing environmentally safe new compounds has been the impetus to develop biological controls. The two most favoured candidates are Gambusia affinis, a fish that feeds on mosquito larvae, and Bacillus thuringiensis, a bacterium that kills mosquito larvae. Numerous other predators and pathogens (viruses, protozoa, fungi and nematodes) are also being evaluated, but all of these anti-larvae measures suffer from the same shortcomings as those described earlier, namely difficulty in covering all potential breeding sites.

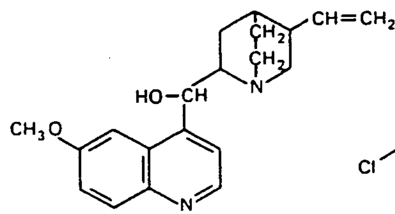
Since the 1950's, laboratory and field research has also focussed on the genetic manipulation of insects. Although the screworm (Cochliomyca hominivoray) was successfully eradicated by releasing sterile males and females, no lasting success has been achieved in the natural environment against Anopheles. The development of Anopheles strains more resistant to Plasmodia is also being actively studied, together with Anopheles strains more susceptible to control; these approaches have been reviewed in detail (Davidson (1982), White (1982)).

1.3(b) Anti-parasite controls

The first anti-malarial drug to be discovered was quinine (see Historical Perspective), and it is still in use today. Many of the more recent anti-malarials have a similar structure to quinine (see Fig. 1.2), and all are based on the phenyl ring. The drugs,

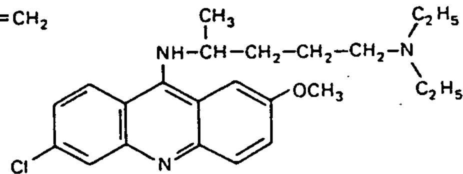
FIG1.2

The Molecular Structure of Some Antimalarial Drugs



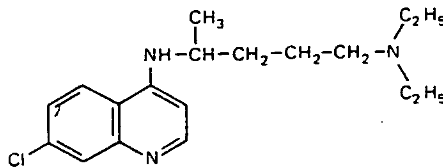
Quinine

6-methoxy-α-(5-vinyl-2-quinuclidinyl)-4-quinolinemethanol



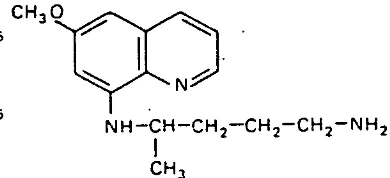
Mepacrine

2-methoxy-6-chloro-9-(4'-diethylamino-1'-methylbutylamino)-acridine



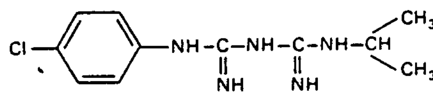
Chloroquine

7-chloro-4-(4'-diethylamino-1'-methylbutylamino)quinoline



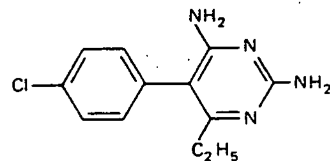
Primaquine

6-methoxy-8-(4'-amino-1'-methylbutylamino)quinoline



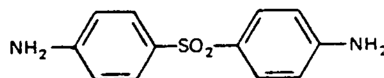
Proguanil

N^1-(*p*-chlorophenyl)-N^5-isopropyldiguanide



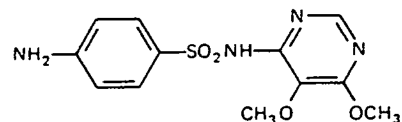
Pyrimethamine

2,4-diamino-5-*p*-chlorophenyl-6-ethylpyrimidine



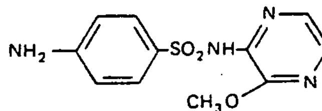
Diaphenylsulfone (Dapsone)

4,4'-diaminodiphenylsulfone



Sulformetoxine (Sulfadoxine)

5,6-dimethoxy-4-sulfanil-amidopyrimidine



Sulfalene

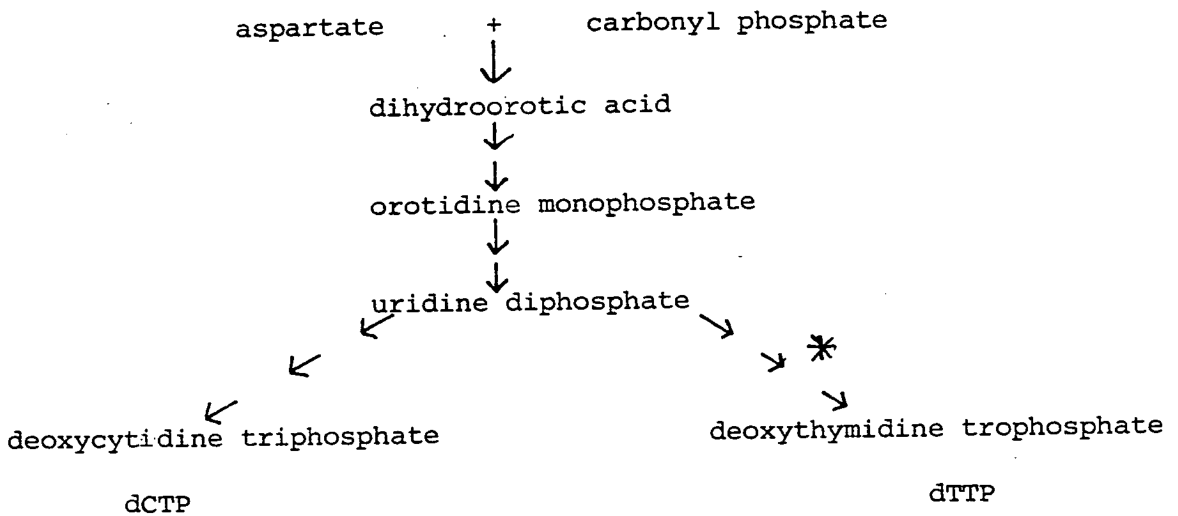
3-methoxy-2-sulfanilamido-pyrazine

Note: taken from Bruce-Chwatt(1980)

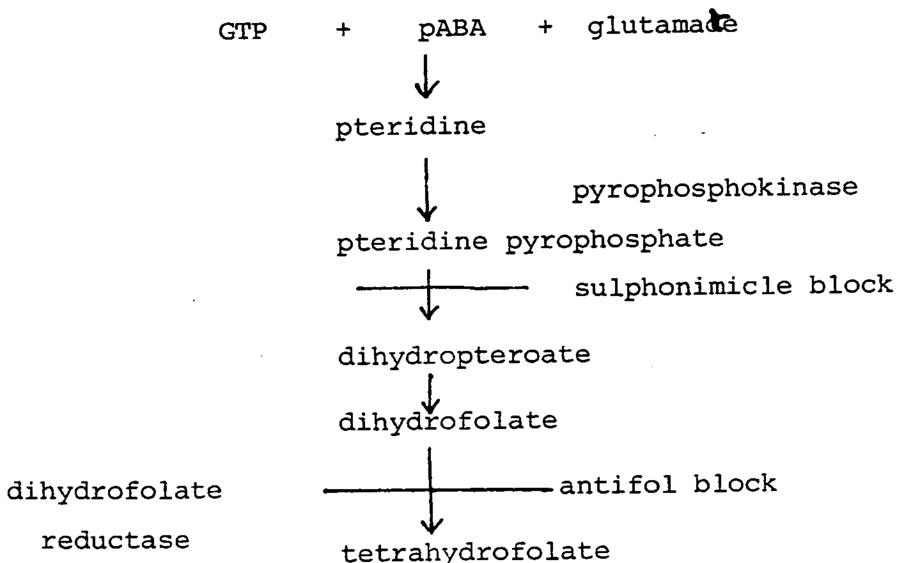
however, are effective against parasites in different stages of development and hence can be grouped by this criterion. Drugs which attack the liver stages of the parasite (see Life-cycle Fig. 1.1) are termed casual prophylactics, and primaquine and proguanil are of this type. Gametocytocidal drugs destroy the sexual forms of the parasite, and so prevent transmission to the mosquito. Primaquine is also effective here, and so is quinine and chloroquine, but the latter two fail to act on the gametocytes of P.falciparum. The majority of the antimalarial drugs destroy the blood forms of the parasite and are termed schizontecides. They can be slow acting, and therefore useful as prophylactics, examples are proguanil, pyrimethamine and the sulphones.

All antimalarial drugs can produce side effects, some only when used in high dose such as pyrimethamine and the sulphones. Mercapcrine produced such toxic side effects that it has been withdrawn. Quinine although highly efficatious, is only used for radical cures, because prolonged use has been associated with blackwater fever, and at lower doses ringing in the ears and restricted vision. Chloroquin is an excellent preventitive and curative drug, with no significant side effects, although it is not gametocytocidal to P.falciparum.

A putative mode of action of chloroquin has already been discussed (see section 1.2). The sulphones plus pyrimethamine and proguanil (antifols), mediate their affects by interferring with folate metabolism. As previously mentioned Plasmodia perform only de novo synthesis of pyrimidines. Pyrimidine biosynthesis is outlined below,



Shown (*) is the point at which the essential cofactor tetrahydrofolate is used. Pyrimethamine and proguanil work by binding more strongly to the parasite dihydrofolate reductase enzyme than to the vertebrate equivalent (see below and General Discussion). The sulphones compete with para-aminobenzoic acid (pABA) for the enzyme which converts it into dihydropteroate via pteridine. Both the folate synthetic pathway, and the pathway for pyrimidine biosynthesis are adapted from Sherman (1979).



The ability to use the antimalarials is being seriously hampered by the emergence of parasite resistance to virtually all the antimalarial drugs. A parasite is defined as resistant when the drug that is required to control the infection reaches or exceeds that which is fully tolerated by the host (for a comprehensive review of drug resistance see Peters, 1980). Some aspects of the genetics of drug susceptibility are discussed in Section 1.5

The emergence between 1961 and 1965 of chloroquin-resistant strains of *P.falciparum* in S.E. Asia, prompted the US military to launch a major programme to develop new antimalarial drugs. Since then screening over 1/4 million compounds, has yielded only 6 with antimalarial activity, and only 1, mefloquine, has reached the clinical trials stages. Resistance to mefloquine (which is structurally related to quinine) has already been demonstrated in the laboratory, so the efficacy of mefloquine may be short lived. In response to the widespread resistance, drug companies have resorted to marketing combinations such as Maloprim, which is pyrimethamine plus dapsone (a sulphone), and Fansidar, which is pyrimethamine plus sulphadoxine. Resistance to both of these emerged quickly, possibly because the antifols and sulphones both inhibit folate metabolism. Chloroquin resistance has taken longer to emerge (probably due to its effect being multifactoral, see Section 1.5), however, isolates can be multiply resistant such as the Thai isolate K1, which is resistant to both chloroquin and pyrimethamine (Thaithong, *et al*. 1981). The need for a new approach is paramount.

1.4 Development of a vaccine

Inhabitants of hyperendemic areas show a consistant pattern of susceptibility and resistance which reflects the slow development of specific immunity. Infants born to immune mothers are relatively resistant during the first three months of life, and thereafter all children suffer severe and recurrent attacks of malaria. These attacks are often fatal (see Section 1). The severity of these attacks diminishes with age, until no clinical symptoms are obvious, although people remain infected. The slow development of naturally acquired immunity in man, and its incomplete nature casts doubts about the effectiveness of any vaccine.

Examination of the life-cycle of *P.falciparum* reveals that at three stages the parasites are 'naked' within the blood stream, and therefore more susceptible to immune attack. These stages are 1) sporozoites 2) merozoites and 3) gametocytes (see Fig.1.1). Vaccination with attenuated forms of each, does result in immunity (for a review see Taylor and Siddiqui, 1982). Obviously before vaccination with attenuated parasites can be considered for wide application, methods will have to be developed for the large scale production of parasites; to date large scale cultivation of malaria parasites is not technically possible.

The development of a continuous *in vitro* cultivation system for the blood forms of *P.falciparum* by Traeger and Jensen (1976), made possible the production of analytical amounts of parasites. Using antibodies raised against *P.falciparum*, studies are now underway aimed at identifying the clinically important antigens (Perrin, *et al.* 1981; Brown, *et al*. 1982; Hall *et al*. 1983).

The recent report that malaria antigens can be expressed in E.coli (Kemp, et al. 1983) gives support to the view that large scale production in bacteria of protective antigens is possible, and that the development of a vaccine by this method is a feasible proposition.

1.5 Variation in clones of Plasmodia

Genetic study of malaria parasites has so far been limited to two of the rodent Plasmodium species, P.yoelii and P.chabaudi, since the entire life-cycle of parasites belonging to these species can be followed in the laboratory. Regarding the human species P.falciparum, while the erythrocytic stages of the parasite can be maintained indefinitely in culture (Trager and Jensen, 1976), it is not ethically permissible to complete the life cycle by deliberate infection of human volunteers. The use of primates such as the Aotus monkey, though possible on a limited scale, is almost impracticable for extensive genetic work which requires large numbers of progeny for analysis.

A comprehensive recent review on the genetic basis of diversity in malaria largely based on studies of the rodent malarias, is given by Walliker (1983), and the salient points are outlined below. Three main characters were used to assay for variation, they are (i) enzymes, (ii) drug susceptibility and (iii) virulence.

Six enzymes have been exploited in this analysis and they are glucose phosphate dihydrogenase (GPI), 6-phosphogluconate dehydrogenase (PGD), lactate dehydrogenase (LDH), glutamate dehydrogenase (GDH), adenosine deaminase (ADA) and peptidase (PEP). These enzymes exist as more than one type of electrophoretic variant,

and the inheritance of a particular variant can be determined by either starch gel or cellulose acetate gel electrophoresis. The electrophoretic variants of each enzyme are denoted by numbers, e.g. GPI-1,-2, etc. according to their position on the gel. When the blood forms were examined only a single enzyme variant was found in each of the cloned progeny, the first indication that the blood forms of the parasite are haploid. Analysis of the progeny of a cross between parental types with different variants revealed that recombination had occurred and the isoenzyme types were inherited in a Mendelian fashion.

Drug-resistant lines of rodent malaria parasites have been obtained either by 'single-step' or 'multi-step' challenges, using predominantly chloroquine and pyrimethamine. Both approaches result in the generation of stable resistance. Crosses between pyrimethamine resistant and sensitive clones showed that the inheritance of resistance was monofactorial and independent of inheritance of the iso-enzymes markers. This contrasts with the inheritance of chloroquine resistance which is multi-factorial. These results showed that the highly resistant clone resulted from a number of successive mutations at specific chromosomal loci (and explains why resistance to chloroquin takes so long to appear in the wild). Progeny from a cross with a sensitive clone displayed varying degrees of resistance. Finally chloroquine and pyrimethamine resistance were shown to be inherited independently.

Variations in virulence have been recorded in many *Plasmodium* species, and the genetics of this character have been investigated in *P.yoelii*. Crosses between virulent and avirulent lines, in

which enzyme and drug-resistant marker genes were present, showed that segregation of virulent and avirulent types occurred, as well as recombination between genes for virulence and markers.

Recombination involving genes at different chromosomal loci has been demonstrated to occur, and is dependent on mosquito transmission (i.e. sexual reproduction is essential). The ease with which different isolates of a species may be crossed in the laboratory, as shown by the frequent occurrence of gene recombination following transmission of genetically diverse parasites through mosquitoes, makes it seem likely that hybridisation and recombination occur abundantly in nature. Evidence to support this view has been obtained with P.chabaudi, in which three forms of the enzyme PGD and four of LDH occur among different wild isolates (Carter and Walliker, 1975). Since no fewer than 9 out of the 12 possible combinations of variants of the 2 enzymes were found, it implies an extensive degree of random mating.

The use of iso-enzyme typing and drug sensitivity has also been applied to study variation between strains of P.falciparum (Rosario, 1981). As mentioned in Section 1.4, monoclonal antibodies raised against different isolates of P.falciparum are now available. These have also been used to assay for variation (McBride, et al. 1982), and it was found that a single isolate could exhibit different types of antigen. Variation has been demonstrated not only between isolates, but also between clones of a single isolate (Rosario, 1981; Thaithong, et al. 1983). The results of characterising isolates of P.falciparum from many parts of the world show that it appears to comprise a single interbreeding population

world-wide, and that considerable diversity exists in enzymes, antigens, and drug sensitivities. In spite of the demonstration of variation, there is as yet no demonstration of linkage between markers, nor any knowledge of the molecular basis for variation; for example is the variation so far characterised reflected in the genomic organisation?

In some instances the ability to distinguish between strains of parasitic protozoa is of extreme clinical importance, and the distribution of repetitive DNA has been used to identify strains of Trypanosomes and Leishmanias (Borst, et al. 1980; Barker and Butcher, 1983). The data presented in Chapter III describes the distribution of a class of repetitive DNA between a number of clones of P.falciparum, and direct comparison is made between the distribution of this repetitive DNA, and the other parameters of variation.

A sequence can be considered as being repetitive if it is present more than once within the genome. In general, repeated sequences are classified into two major categories: satellite DNA, which consists of highly repeated, tandemly linked, simple DNA sequences with a preferential location in the heterochromatin, where it appears to have some structural role (for review see Walker, 1971); and interspersed repetitive DNA.

The relationship between repetitive DNA and single copy DNA is termed the 'interspersion pattern', and it refers to the alternating interspersion of non-repetitive and repetitive DNA. This arrangement was first demonstrated for Xenopus DNA, and hence is often termed the 'Xenopus pattern'. This form of sequence

organisation is characterised by single copy sequence less than 1.5 kb in length. Ajoining this single copy sequence are repetitive sequence elements which themselves have an average length of 300 bp (Davidson, et al. 1973). This short-period interspersion pattern is displayed by most eukaryotes, such as man (Schmid and Deininger, 1975), sea urchin (Graham, et al. 1974), and the housefly Musca domestica (Crain, et al. 1976).

A distinctly different pattern of sequence organisation has been found in Drosophila melanogaster DNA. Here middle repetitive sequences have an average length of 5.6 kb (though about 10% are 500 bp or less), and the single copy sequence elements extend for at least 10 kb on average (Manning et al. 1975). This 'Drosophila pattern' has also been demonstrated for the bee Apis mellifera (Crain, et al. 1976), and for the nematode Caenorhabditis elegans (Schachat, et al. 1978). The long-period interspersion appears to be a characteristic of lower eukaryotes (some lower eukaryotes even lack interspersion), and it seems that the change in genome organisation to a short period pattern may have occurred during the evolution of the Diptera, since both types of interspersion can be found in flies (Drosophila and M.domestica).

There is very little information about the genomic organization of Plasmodia. Only a single study on the renaturation kinetics of the P.falciparum genome has been reported (Hough-Evans and Howard, 1982), and although the authors report that the repeated sequences are on average ca. 400 bp, they merely assume that these are interspersed with single copy sequence. More information is available for P.berghei (Dore et al. 1980 and 1983), and in the latter publication

these authors report that the repetitive sequences (400-1400 bp) are largely interspersed with unique DNA of average length 8 kb (Dore, *et al*. 1983). This implies that the *Plasmodia* genome is organised with a '*Drosophila* pattern' of long period interspersion.

Is there any significance to this pattern of interspersion? The observation that all eukaryotic genomes contain repeated DNA sequences, and that these can account for more than half the nuclear DNA content, suggests that it may have a persistent and basic function. The hypothesis of a role for repetitive sequences in regulation is certainly not new (Britten and Davidson, 1969), and an extensive list of examples where repeated genetic elements have been shown to affect gene expression is given by Shapiro and Cordell (1982). However a difficulty in discussing the role of interspersed repetitive sequences is that the pattern represents an average of many repetitive families. If a repeated sequence does indeed provide a set of related controls, then ultimately the significance of interspersion may lie in the function of the regions adjacent to the different members of individual repetitive families.

An example where a role has been defined for an interspersed repetitive sequence is that described by Zuker and Lodish (1981). They purified a repetitive sequence from the genome of *Dictyostelium discoideum* the clone contained a sequence 300-500 bases long that is also present in polyA$^+$mRNA. This repetitive sequence is absent in message expressed up to 5 hrs of development. The message that contain this element comprises two groups, one expressed initially at 5.5 hrs. and those expressed between 5 and 15 hrs of development. The observation that only message expressed during this developmental

period (5-15 hrs) contains this repetitive sequence argues strongly that it is involved in regulating the expression of these sequences.

1.6 Ribosomal RNA gene organization

The ribosome is the translation machinery of all cells. It is a complex of proteins (ribosomal proteins), and RNA (rRNA). The conservation in function of the ribosome has resulted in an overall conservation of structure, that includes the structure of its constituent parts. There are however significant differences between eukaryotic and prokaryotic ribosomes. Eukaryotic ribosomes are appreciably larger than prokaryotic ones; they contain a greater number of proteins, about 80 rather than 53, and they have extra RNA molecules. In addition, the proteins and nucleic acid are, on average larger. The difference in size is a paradox since both eukaryotic and prokaryotic ribosomes perform the same basic function, namely to catalyse the synthesis of protein.

A few general characteristics of eukaryotic ribosomes follow, and except where stated the references are given in a review by Wool (1979). Ribosomes are composed of two subunits (s.u.) which together contain 4 molecules of RNA and 70-80 proteins. Exceptions are found in *Drosophila*, where there is an extra 2S RNA, and in *Trypanosomes* and *Chlamydomonas* ribosomes which contain several small RNAs (Jordon and Glover, 1977; Cordingley and Turner, 1980; Hernandez *et al*. 1983; Marco and Rochaix, 1980). The sedimentation value (S) varies between species, but generally is close to 80S for the complete ribosome; 60S for the large s.u., and 40S for the small s.u. There is a single 18S RNA molecule and about 30 proteins in the 40S s.u., and one of each of 5S, 5.8S and 28S RNAs plus

45-50 proteins in the 60S s.u.

The rRNA and the proteins interact, and this interaction is poorly defined in eukaryotic ribosomes. Two methods have been used to identify which proteins bind to a particular RNA, affinity chromatography and periodate oxidation. In the former the RNA is immobilised on a sepharose column, and ribosomal proteins passed over it. In this way eukaryotic 5S RNA has been shown to bind to eukaryotic ribosomal proteins L6 and L19, and 5.8S rRNA to eukaryotic ribosomal proteins L19, L8, L6, S13 and S9. Since 5.8S rRNA binds to E.coli ribosomal proteins L18 and L25, it implies that 5.8S rRNA is the eukaryotic analogue of the prokaryotic 5S rRNA which also binds these proteins. Such binding studies have to be interpreted with caution, since on extraction of the 5S RNA-protein complex from the 40S s.u., the RNA proves to be associated with eukaryotic L5 and not L6 or L19. Moreover, periodate oxidation cross links the 3' terminus of 5.8S rRNA to eukaryotic L7 and L23, 18S to 53a and 28S to L3. A possible explanation for the discrepancy is that the 3' binding site in the RNA is destroyed when it is immobilised on the sepharose column. What is clear is that in vivo the association between the rRNA and the ribosomal protein is very rapid, the ribosome is assembled in the nucleus on the large precursor rRNA, even as it is being transcribed from the DNA; for a recent review of ribosome structure see Lake (1983).

The rRNAs not only interact with proteins, they also interact with each other. Probably the best postulated interaction is that between the 3' end of the 16S rRNA, and the 5' end of mRNA in prokaryotes. A recognition that there is a canonical sequence,

which is the mRNA binding site in the 16S rRNA was proposed by Shine and Dalgarno (1974). In eukaryotes only a few 18S rRNA sequences have been elucidated, such as Xenopus (Salim and Maden, 1981) and yeast (Rubstov, et al. 1980), however a large number of eukaryotic mRNAs have now been sequenced (Breathnach and Chambon, 1981), and there is no recognisable eukaryotic equivalent to the Shine-Dalgarno sequence. The 5.8S rRNA interacts with both the 5' and the 3' ends of the 28S rRNA, and one of its roles may be to confer a particular conformation on the 28S rRNA, The interactions between the 5.8S rRNA and the 28S rRNA, and discussions of the possible function of this interaction have recently been reviewed (Walker and Pace, 1983).

I will use the term 'rRNA genes' to designate the sequences encoding the 18S, 28S, and 5.8S rRNAs, and 5S RNA genes to designate the sequences encoding the 5S rRNA. In most eukaryotes, the 5S RNA genes are separate in the genome from the rRNA genes (Long and Dawid, 1980). The exceptions are yeast and the slime mould Dictyostelium where the 5S gene is 3' to the 25S and 26S rRNA genes respectively (Nath and Bollon, 1977; Cockburn, et al. 1976; Maizels, 1976). The 5S RNA and rRNA genes are present in multicopy in almost all eukaryotes studied so far, but their location varies (Long and Dawid, 1980). In the lower eukaryotes Physarum, Dictostelium, and Tetrahymena, the multiple rRNA genes are predominantly or entirely extrachromosomal (Campbell, et al. 1979; Cockburn, et al. 1978; Gall, 1974). When chromosomal they can be located on autosomes, such as in humans where the rRNA genes are on the telomeres of chromosomes 13, 14, 15 20 and 21 (Henderson and Atwood, 1976); or they can be on the sex chromosomes such as in Drosophila

(Wellauer and Dawid, 1978). In both humans and *Drosophila* as well as *Xenopus laevis*, the rRNA génes are clustered at the nucleous organiser, NO (*loc. cit.* and Wallace and Birnsteil, 1966).

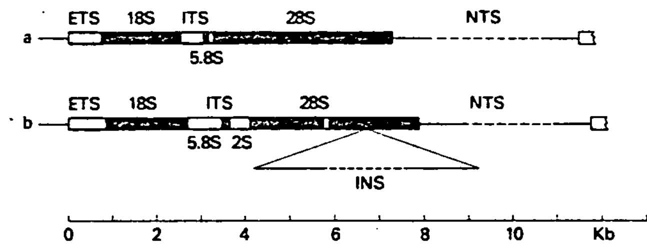
In eukaryotes there are three RNA polymerases which were originally distinguished according to the inhibitory effect of α-amanitin. Although slightly different in size the three polymerase each have a molecular mass approximately 500,000 (eukaryotic RNA polymerase reviewed by Chambon, 1975). RNA polymerase A (1) transcribes the rRNA genes, RNA polymerase C (III) transcribes the 5S RNA and tRNA genes, and RNA polymerase B (II) transcribes most mRNAs (Chambon, 1975). In all of the arrays of rRNA and 5S RNA genes analysed so far, the length of the DNA repeated exceeds the length of the transcription unit or units (Arnheim and Southern, 1977; Wellauer, *et al.* 1976; Long and Dawid, 1980). The length of the extra sequence repeated along with the transcription unit varies from a few hundred to several thousand nucleotides for the 5S RNA genes, and from less than 5 kb to more than 58 kb for the rRNA genes (for a review of spacers see Fedoroff, 1979). A repeating unit generally contains a single transcription unit, although the oocyte 5S RNA gene family in *X.borealis*, which can have four or more genes per repeating unit, is an exception (Korn and Brown, 1978). In yeast the 5S RNA transcription unit is separated from the rRNA transcription unit by a spacer, and is transcribed in the opposite direction (Valenzuela, *et al.* 1977). Henceforth I will refer to the repeating units containing the rRNA genes as rDNA, and the 5S RNA genes as 5S DNA.

Two well-studied rDNA repeats are those of *Xenopus* and

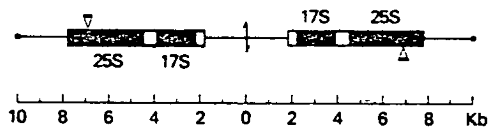
Drosophila, maps of these reports are presented in Fig. 1.3, and are taken from Long and Dawid (1980). As noted previously transcription units alternate with spacer regions. The latter are referred to as either transcribed spacers, which are sub-divided into external (ETS) and internal (ITS) regions, or non-transcribed spacers (NTS), depending on whether the sequence appears in the initial rRNA transcript. In X.leavis there is a single autosomal NO, which comprises 400-500 tandem copies of rDNA (Wallace and Birnsteil, 1966). The transcription unit is flanked by short non-repetitive spacer sequences, the remainder of the non-transcribed spacer sequence varies in length from about 2.7 kb to 9 kb in different repeating units. Despite this difference in length, the spacers differ little in sequence because they are internally repetitive (Wellauer, et al. 1976; Botchan, et al. 1977).

In Drosophila melanogaster there are two NO, one on the X and one on the Y chromosome; each containing 150-200 rDNA genes, although the number varies somewhat between strains (Ritossa and Spiegelman, 1965). The rDNA of D.melanogaster can be of two classes, one resembles that of X.leavis (Fig. 1.3), but the other differs in that the 28S rRNA gene contains an intervening sequence (IVS). The IVS can vary in size from 0.5 to 6 kb, and the sequences are of two groups: Type 1 insertions included the most abundant 5 kb insertion, plus homologous small 1 kb and 0.5 kb insertions, and occur exclusive in the X chromosome NO; Type II insertions have no homology with Type I, occur in several sizes between 1.5 and 4 kb and are located on both the X and the Y chromosomes (Glover and Hogness, 1977; Dawid, et al. 1978; Wellauer and Dawid, 1978). Sequence

FIG 1.3



Maps of repeating units of ribosomal genes of *Xenopus laevis* (*a*) and *Drosophila melanogaster* (*b*). Coding regions are shown as solid bars; transcribed spacers, i.e. regions transcribed into nonconserved segments or pre-rRNA, are shown as open bars; and nontranscribed regions are shown as lines. The transcribed spacers are distinguished as external (*ETS*) and internal (*ITS*). Conventionally, the *ITS* is thought of as including the coding regions for 5.8S and 2S RNAs, the latter only present in *D. melanogaster* but not in *X. laevis*. In *b* the point of interruption of some of the 28S rRNA regions is indicated. NTS and insertions (*INS*) occur in variable length as indicated by a broken line.



Map of the extrachromosomal rDNA in *Tetrahymena*. The molecule is a palindrome, i.e. it has a center of symmetry as indicated on the figure. Two transcription units are directed outward, with the two pre-rRNAs copied from opposite strands. Transcribed spacer regions are indicated

Note, maps taken from Long & Dawid (1980)

elements homologous to the Type 1 insertions also occur in multiple copies outside the NO (Dawid, *et al*. 1981). In *D.melanogaster* rDNA repeats interrupted by Type 1 insertions are interspersed on the X chromosome with uninterrupted rDNA, this is in contrast with the situation in *D.hydei* where 90% of the Type I genes are clustered (Hawley and Tartof, 1983). The repeats containing the Type II insertions are significantly clustered on the Y chromosome, and possibly also on the X chromosome in *D.melaongaster* (Wellauer and Dawid, 1978).

Although insertions sequences constitute the major source of length heterogeneity in the rDNA of *D.melaongaster*, there is also length heterogeneity in the non-transcribed spacer sequence of about 3.5 kb. As in *X.laevis*, the non-transcribed spacers are internally redundant, and are very similar in the X and Y NO (Wellauer and Dawid, 1978).

In addition to the 2S rRNA, *D.melanogaster* is also unusual in that the 28S rRNA gene contains an IVS, and the 28S rRNA has a break, lacking 120 nucleotides present in the rDNA (Wellauer, and Dawid, 1977). Although unusual these features are not uncommon. I have previously mentioned the presence of extra small RNAs in *Trypanosomes* and Algae; the presence of a break in the large rRNA has also been noted in *Leishmama donovani* (Leon, *et al*. 1978), *Trypanosomes* (Cordingley and Turner, 1980; Castro, *et al*. 1981), *Plasmodium berghei* and *P.chabaudi* (Miller and Ilan, 1978; DaSilveira, *et al*. 1983). Maturation of the 26S rRNA of *Tetrahymena pyriformis* also results in a break to give two fragments (Eckert, *et al*. 1978). As well as IVS in *Drosophila*, interrupted large rRNA genes are observed in some strains of *Tetrahymena* (Wild and Gall, 1979), and *Physarum*, which is unusual in having two

IVS (Campbell *et al.* (1979)). Surprisingly the mitochondrial rRNA gene of *Neurospora crassa*, and the chloroplast gene of *Chlamydomonas reinhardii* also contain an IVS (Hahn, *et al.* 1979; Rochaix and Malnoe, 1978).

The possession of a Type I interruption in the 28S rRNA gene in *D.melanogaster* renders that gene transcriptionally inactive (Long *et al.* 1981), whereas Type II IVS genes are variently transcribed, though at a low level (Kidd and GLover, 1981). Since all the rRNA genes of *P.polycephalum* contain IVS, they are obviously expressed. In *Tetrahymena* the rDNA of some species such as *T.thermophila* contain an IVS which is transcribed and then spliced out, while the rDNA of other species such as *T.pyriformis* contain no IVS. The *T.pigmentosa* species is heterogeneous, only some strains having an IVS in the rDNA (Wild and Gall, 1979; Din *et al.* 1979). The removal of an IVS has been best-studied in *Tetrahymena*, and will be discussed later together with transcription and processing of rRNA genes.

Extrachromosomal rDNA occurs quite frequently in unicellular eukaryotes. In *T.pyriformis* the micro nucleus contains only a single integrated copy of the ribosomal genes (Yao and Gall, 1977). In the macro nucleus this single gene is excised and amplified on a linear extrachromosomal palindromic DNA molecule that contains two copies of the rRNA genes (Fig. 1.3 and Pan and Blackburn, 1981). A repeated hexanucleotide C_4A_2 is present at or near the free ends of these palindromes, it is also found in the micro nucleus, and is believed to be involved with the chromosome breakage that generates the excised copy (Yao and Yao, 1981; Blackburn and Gall, 1978).

The excised copy is first replicated as a single 11 kb molecule, which is subsequently replicated to form palindromic rDNA molecules. This replication results in the loss of 0.3 kb of DNA from the end that becomes the centre of the palindrome, and the resultant molecule has the hexanucleotide C_4A_2 repeated at its termini (Pan and Blackburn, 1981).

Palindromic rDNA occurs in several unrelated organisms such as P.polycephalum (Campbell, et al. 1979), and D.discoideum (Cockburn, et al. 1978). A different structure occurs in Paramecium tetraurelia, here the rDNA is in both linear and circular extrachromosomal molecules (Findlay and Gall, 1978). In another ciliated protoxoan Oxytrica fallax, the rDNA in the macronucleus is extrachromosomal, but O.fallax is a special case, as all the genes in the macronucleus appear to be amplified as single gene-pieces of DNA (Prescott and Murti, 1973). Amplified rDNA in oocytes is always extra-chromosomal, and its properties are discussed below.

There is a single cluster of 5S RNA genes in D.melaongaster, which contains approximately 160 copies of the gene, and is located at band 56F on the right arm of chromosome 2 (Wimber and Steffenson, 1970). In X.laevis the arrangement of 5S RNA genes is quite complex. There are two distinct 5S RNA gene families, comprising separate tandem arrays; these families have been termed 'oocyte' and 'somatic' based on the observation that the former is only expressed during oogenesis (Ford and Southern, 1973). The somatic 5S RNA genes are much less abundant than the oocyte 5S RNA genes, which form a large family. In X.laevis these genes occur in clusters at the telomeric regions of most if not all chrososomes

(Pardue, et al. 1973).

Amplification of rDNA is a common feature of oogenesis, where it is thought to provide sufficient DNA templates for the large accumulation of ribosomes in oocytes. Amplified DNA in X.laevis occurs in large circles containing many hundreds of repeating units, which are homogeneous in length. As chromosomal rDNA is heterogeneous in length, this implies that a single repeat is amplified (Wellauer, et al. 1976). I have already discussed a particular case of rDNA amplification which takes place in Tetrahymena, where the rDNA genes in the macronucleus are amplified on extrachromosomal palindromes (Yao and Gall, 1977).

Dosage regulation of rDNA genes is exhibited by both Xenopus and Drosophila somatic cells. In X.laevis heterozygous anucleolate females lacking a NO accumulate the same amount of amplified rDNA in their oocytes, as do wild-type frogs, even though their somatic cells only have half the number of genes (Perkowska, et al. 1968). In the polytene chromosomes of salivary glands of D.melaongaster the hetero chromatin containing rDNA is under-replicated compared to the euchromatin. This results from only the rDNA of the Y chromosome NO being replicated (Endow and Glover, 1979). Since the Y NO can be replicated independent of the X NO, this may explain the 'compensation' described by Tartof (1971). He observed that when flies lacked a copy of the X NO, they compensated for this loss of rDNA by increasing the remaining rRNA genes.

Drosophila also exhibit another mechanism of varying the rDNA content. If the genome contains less than half the normal number of rDNA units, rRNA synthesis is inadequate and a mutant, bobbed (bb)

fly develops, which phenotypically have shorter bristles. This loss of rDNA can be made up, by a process which Ritossa (1968) termed 'magnification'. The ability to be able to magnify the rDNA is inheritable, but the underlying mechanism of magnification is not known.

I have previously cited that the rRNA genes are transcribed by RNA polymerase I(A) (Chambon, 1975). Sequences that promote RNA pol I transcription have yet to be identified, but deletion, in vitro transcription, and sequencing studies of mouse (Miller and Sollner-Webb, 1981), Drosophila (Kohorn and Rae, 1982), Xenopus (Bakken, et al. 1982) and yeast (Verbeet, et al. 1983), reveal that information essential for the initiation of rRNA genes is present within the region ca. -200 to +100, where transcription begins at +1. This implies that the promoter for RNA pol I lies in the non-transcribed spacer region, and this region has been shown to promote in vitro transcription of both Drosophila and Xenopus rDNA (Kohorn and Rae, 1982; Miller and Sollner-Webb, 1981).

The expression of rRNA genes in eukaryotes is a multistep process. It begins with the transcription of a large RNA molecule which serves as the precursor of the three (or more) RNA species found in the ribosome. The rRNA precursor is assembled into a ribonucleoprotein (RNP) particle during transcription. This primary transcript undergoes a series of processing steps, including modification of nucleosides, and removal of spacer sequences by nucleotide cleavage. In cases where the genes possess an IVS, this is also removed by a cleavage-ligation step, termed splicing. Finally the fully processed rRNP is exported from the nucleus to the cytoplasm.

The above mentioned processes have been most studied in Tetrahymena, and I will summarise here the transcription and splicing of the Tetrahymena rRNA precursor. This process has been comprehensively reviewed, and the salient points are outlined below (Cech, et al. 1982). Electron microscopy combined with hybridisation and R-loop mapping of the rDNA showed that transcription proceeds bidirectionally from the centre of the palindrome, and that the 17S rRNA is transcribed first (Engberg et al. 1976). Pulse-chase labelling experiments revealed that the primary transcript is a 35S RNA molecule which is rapidly methylated (Eckert, et al. 1978). Methylation occurs principally in regions that are conserved throughout processing, and the primary transcript encodes the 17S, 5.8S and 26S rRNAs. The 5' end of the 35S RNA has been mapped (by S1 nuclease digestion) to a site approximately 110 bp from the axis of symmetry of the rDNA, and using the same technique the terminator region has been located to approximately 13-15 nucleotides downstream from the 3' terminus of the 26S rRNA (Niles, et al. 1981).

After its synthesis, the pre-rRNA (35S) undergoes a series of processing events, the first of which appears to be the splicing out of the IVS (when present). In T.thermophila the IVS is excised as a discrete linear RNA molecule, which is subsequently circularised (Grabowski, et al. 1981). Since there is no evident delay in the processing kinetics of 35S rRNA with and without an IVS, it implies that the splicing event is very rapid (Cech and Rio, 1979). The next processing step generates the 17S rRNA, which is rapidly exported to the cytoplasm, and then the pre-26S rRNA is cleaved to give the 5.8S and 26S rRNAs.

The rate of these processing steps appears to be dependent on the growth rate of the cells, it decreases 36 fold when the cells are shifted to starvation medium (Cech et al. 1982).

In a remarkable series of experiments Cech and his co-workers demonstrated that the splicing of the IVS in T.thermophila is an autocatalytic event (Kruger, et al. 1982). A 1.6 kb segment of T.thermophila rDNA that included the 413 bp IVS was cloned into a plasmid which contained the lac UV5 promoter-operator region. Purified recombinant plasmid DNA was transcribed in vitro with E.coli. RNA polymerase, and the resultant RNA fractionated by denaturing polyacrylamide gel electrophoresis. The linear and circular IVS were detected in this way, moreover using deproteinized RNA they showed that excision was not dependent on transcription. Incubating the linear IVS RNA at 39°C in a buffer containing Mg^{++}, they further demonstrated that circularisation was independent of any enzyme.

Since the IVS has no open reading frames, its secondary structure is thought to catalyse the autoexcision event. It is noteworthy that Tetrahymena IVS shows structural homology with fungal mitochondrial IVS, which themselves appear to be self-spliced (Cech, et al. 1983). This relationship suggests a common origin for some nuclear and mitochondrial introns, and possibly similar mechanisms for their splicing.

There have already been several references throughout this Introduction to Protozoan rDNA organization. Since protozoan rDNA organisation is most germane to the work presented in this

Thesis, I will reiterate some of the points whilst considering the available data on protozoan rDNA. I shall confine myself to the uninucleate protozoa, as those that possess a macronucleus appear to have the rDNA almost exclusively on extrachromosomal elements, and these have been discussed in some depth.

The first detailed report on the arrangement of rDNA in uninucleate protozoa was that given by Leon, et al. (1978). They described the organisation of rDNA sequences in Leishmania donovani, where the rRNA genes are present as 166 integrated chromosomal copies. The genes were found to be ordered 16S-5.8S-26Sa-26Sb. Note that the 26S rRNA has a gap. The authors presumed that the order was 5'-16S-26S-3', by comparison with Xenopus. A single repeat was shown to be composed of (a) a 16S gene of length 2.12 kb, (b) an internal spacer of 1.23 kb, (c) a 26S gene of 4.3 kb with a 0.587 kb gap, and (d) an external spacer of average length 5.85 kb. A number of 160 copies for the rRNA genes has also been determined for L.braziliensis (Villalba and Ramirez, 1982).

In the Trypanosomatids the rRNA genes are again integrated in the chromosome, and present at approximately 114 copies (Castro, et al. 1981). The large rRNA is gapped and a number of small rRNAs are also present (Cordingley and Turner, 1980; Hernández, et al. 1983). The rRNA gene organization has been studied in T.brucei (Hasan, et al. 1982) and T.cruzi (Hernandez and Castaneda, 1983). In T.brucei the length of the rDNA repeat is at least 21 kb, and the genes are ordered in the repeat in the same general pattern as L.donovani, and again the large rDNA has a gap. In T.cruzi 3 of the 5 small rRNAs map within the rRNA cistron, the

others S4 and S5, map elsewhere independently of one another. Hernandez and Castaneda (1983) also report very large spacer DNA ($>$ 20 kb), this prevented them from discerning whether the rDNA repeats were in tandem. There is also no information as to whether the rDNA repeats in *T.brucei* are in tandem, however Hasan *et al*. (1982) report the interruption of the large rRNA gene by a mobile element (RIME) that is present at approximately 20 copies in the genome. This insertion is believed to inactivate the gene.

In *Plasmodia*, apart from the work presented in this thesis, there are just two reports from the same authors on the rDNA of the mouse malaria *P.berghei* (Dame and McCutchan, 1983a and 1983b). In the 1983b paper they report the cloning of a rDNA fragment containing part of both the 18S and 28S rRNA genes, and that the 18S is 5' to the 28S gene. No homology to *Xenopus* 5S DNA was detected, indicating that the 5S RNA gene may be absent in *Plasmodia*, since 5S RNA sequences are highly conserved (Long and Dawid, 1980). Further interpretation is unfortunately complicated as the rDNA cloned had rearranged (the 3' half of the 28S gene had been replaced by some other DNA). The genomic hybridisations were also complicated by cross-homology to contaminating mouse DNA. Interestingly rearrangement of rDNA clones has also been reported for mouse, due to repeat sequences in the non-transcribed spacer (Arnheim and Kuehn, 1979). In their other paper, Dame and McCutchan report the number of the rDNA transcription units as four. This is a remarkable number, considering that in all other eukaryotes the rRNA genes are present in multi-copy (Long and Dawid, 1980). Further more each of the four transcription units appear to be different,

and this conclusion was drawn from genomic hybridisation patterns displayed to total RNA, and the previously cloned rDNA. They were also unable to detect any link between the transcription units.

These studies suggest that uninucleate parasitic protozoa all have their rRNA genes ingegrated in the chromosome, but that the precise arrangement of the rDNA sequences differ, such as the number of small rRNA genes, and surprisingly the number of rDNA copies per genome. Analysis of the rDNA sequences of the human malaria Plasmodium falciparum constitutes a major part of this thesis work. They represent the first detailed analysis of P.falciparum genes of known function, and the low copy number reported for P.berghei rRNA genes is consistent with the number reported here for P.falciparum, and it appears that Plasmodia may be different from other parasitic protozoa in regard to the number of rRNA genes per haploid genome.

1.7 Actin gene organization

Movement is a fundamental property of living cells. It is expressed in such diverse activities as cytoplasmic streaming, phagocytosis, amoeboid motion, cytokinesis and mitosis. Electron microscopic and immunological evidence has implicated actin in all of these processes and myosin in many of them (for a review see Clarke and Spudich, 1977). Three types of filamentous structures have been identified in eukaryotic cells, microtubules, intermediate filaments and microfilaments.

The principal constituents of all microtubules to date, are equal amounts of α- and β-tubulin which are assembled as heterodimers, and are the products of distinct genes (Luduena and Woodward, 1973;

Cleveland, et al. 1978). Microtubules are integral parts of diverse eukaryotic subcellular structures, including cytoskeletal elements, cilia and flagella, and mitotic and meiotic spindles. Many cells also contain structures termed intermediate filaments. This is a lose term describing structures which are neither microtubules nor microfilaments. Prominent examples of this class of filament are neurofilaments of neurones, and glial filaments of glial cells. The intermediate filaments are often associated with desmosomes (which are involved in attachment between adjacent cells), and as a consequence the sub-unit protein of the filaments is sometimes termed desmin. Chick smooth muscle desmin has a molecular weight of 50,000 and is resolved into two closely related spots by two dimensional electrophoresis (Lazarides and Balzer, 1978).

Cells have a sub-membraneous network of fibres which are identical to microfilaments, and the presence of actin within these microfilament fibres was revealed by immunofluorescent anti-actin antibodies (Lazarides and Lindberg, 1974). This cytoskeletal network is also known as stress fibres, and the involvement of stress fibres in the maintenance of cell morphology is graphically illustrated in fibroflast cell lines. Fibroblasts are motile in culture, and move by 'spreading' over the substratum. The organization of stressfibres is related to the ability of a fibroblast to spread out ,and this relationship is reflected in a striking manner by the changes that occur when anchorage dependent cells are transformed by virus. Using affinity purified anti-chicken actin antibody Boscheck et al. (1981) clearly showed that the arrangement of stress fibres in chicken embryo cells radically alters during

transformation with Rous sarcoma virus. The correlation between the arrangement of the cytoskeletal network and the cells morphogy suggests that actin is also involved in the maintenance of form.

Sequencing studies of a number of actins from a range of organisms, demonstrated that a number of different actins can exist in a single species (iso forms), and that actin is a highly conserved protein (Vandekerckhove, et al. 1978 and 1980). The identification of different forms of actin in a single species suggested that actin is encoded by a multigene family, and the sizes of the actin gene families in different organisms are given in Table 1. With the exception of yeast ,which has only a single actin gene (Ng and Abelson, 1980; Gallwitz and Seidel, 1980), other organisms have actin multigene families of varying sizes. There is a general increase in actin gene number with genome size and evolutionary complexity, although there are exceptions such as D.discoideum which has more actin genes than chicken (see Table 1).

D.discoideum also provides an example in which multiple actin genes appear to code for the same actin isoform (McKeown and Firtel, 1981) ,though none of these genes has been completely sequenced. This may not simply be a reflection of a requirement for a large number of genes to produce sufficient actin, since these genes are differentially regulated during the D.discoideum life-cycle (McKeown and Firtel, 1981). In contrast S.cerevisiae functions with only one actin gene, whereas in D.melanogaster, at least 5 of the 6 genes encode different actin polypeptides (Ng and Abelson, 1980; Gallwitz and Seidel, 1980; Fyrberg, et al. 1981). The use of probes from the 3' non-coding region of chicken β- and γ-actin mRNAs

Table 1. Actin gene number and organisation

Organism	No. of genes	Organisation	References
Saccharomyces cerevisiae (yeast)	1	-	Ng and Abelson (1980) Gallwitz and Seidel (1980)
Oxytricha fallax (ciliated protozoan)	2-3	-	Kaine and Spear (1980)
Dictystelium discoideum (slime mould)	17	2 linked, others possibly	Kindel and Firtel (1978) McKeown and Firtel (1981)
Caenorhabditis elegans (Nematode)	4	3 genes linked	Files *et al*. (1983)
Drosophila melanogaster (fruit fly)	6	all dispersed	Tobin, *et al*. (1980) Fyberg, *et al*. (1980)
Strongylocentrotus purpuratus (sea urchin)	11	some linked	Durica, *et al*. (1980) Schuler and Keller (1980)
Chicken	>7	>3 dispersed	Cleveland, *et al*. (1980)
Man	>20	partially dispersed	Humphries, *et al*. (1981)

has shown that there are 1β and 3 γ-like actin genes (Cleveland, et al. 1980). However, the γ-like genes have not been definitely shown to code for γ-actin.

In most species it is not clear what proportion of the detected actin genes is actually expressed. In the cases where an expressed gene has been identified, it has been done by using 5'- and 3'- non-coding regions of the mRNA (present in the actin cDNA), which is highly specific for an individual gene. Such experiments show that all the D.melanogaster actin genes (Zulaf, et al. 1981), and at least eight of the D.discoideum actin genes (McKeown and Firtel, 1981) are expressed. A negative result is not conclusive, since the gene could be expressed in another cell type, or transcribed at a level too low to be detected.

There are examples where the gene is known not to be expressed, because the gene lacks the regulatory signals (or has internal stop signals). These are sometimes known as 'pseudogenes' and D.discoideum contains at least one pseudogene (pDd actin 2-sub 2) which lacks a TATA box (a TATA or Goldberg-Hogness box is a sequence believed to be required for efficient initiation of transcription, and is located at approximately -30 where transcription starts at +1), shows multiple amino acid substitutions and from which expression cannot be detected (McKeown and Firtel, 1981). When pseudo genes also lack an IVS and regulatory signals they are further defined as 'Processed-pseudo genes'. The pseudo gene pDd actin 2-sub 2 is linked to a functional actin gene, pDd actin 2, from which it presumably arose by duplication (McKeown and Firtel, 1981a).

Two quite distinct patterns of multigene family organisation have been established, with genes being either clustered or dispersed in different chromosomal regions (see Table 1). The clustered pattern has been demonstrated for *D.discoideum* (McKeown and Firtel, 1981), *C.elegans* (Files *et al*. 1983), and *S.purpuratus* (Schuler and Keller, 1981), by the isolation of recombinant phage linking the genes together. In contrast, *in situ* hybridisation experiments have shown that the six *D.melanogaster* actin genes are at separate chromosomal locations (Fyberg, *et al*. 1980; Tobin, *et al*. 1980). The actin genes are also dispersed in chicken (Cleveland, *et al*. 1980).

A number of actin genes from a range of different organisms have now been sequenced, and this permits comparison for evolutionary conserved traits; namely the location of regulatory signals and the position of any introns. A good example is *D.melanogaster*, where S1 mapping has indicated that there may be multiple initiation sites for transcription (McKeown and Firtel, 1981). Most eukaryotic mRNAs have a sequence at the 3' terminus which is thought to be the site for polyadenylation. A polyadenylation signal sequence has also been found at the 3' end of the actin genes (McKeown and Firtel, 1981a).

With respect to the location of introns, actin genes in most organisms are exceptional in that the intron position is highly variable both within, and between species (Durica, *et al*. 1980; Fyrberg, *et al*. 1980 and 1981; Ng and Abelson, 1980; Schuler, *et al*. 1981). In *C.elegans* although the intron position is conserved the intron sequence is not. The intron in gene II is completely

different to that of gene I (Files, *et al*. 1983). The actin genes I and II of *S.purpuratus* each have 2 introns, although the position of one is conserved, the other is not (Schuler and Keller, 1981). Finally, in *D.melaongaster* both the size and the position of the introns vary between the actin genes (Fyberg, *et al*. 1981). In contrast the actin genes of *D.discoideum* are unusual in that they do not possess introns (Firtel, 1981).

Earlier in this discussion of actin genes I referred to the observation that *D.discoideum* actin genes are differentially expressed (McKeown and Firtel, 1981). Differential expression of actin genes has also been demonstrated in sea urchins, where actin synthesis increases with early sea urchin development (Merlino, *et al*. 1980).

Developmental regulation of actin gene expression is also a feature of *D.melaongaster* (Zulaf., *et al*. 1981). Here actin gene III appears to be specific for the thorax, whereas actin gene I is only expressed in larvae. Furthermore Sanchez, *et al*. (1983) report that 2 actin genes, those located at positions 79B (on the 3 L chromosome) and 88F (on the 3 R chromosome) show similar transcriptional activity until the pupal stage, thereafter the 88F genes produce more transcripts. Since *D.discoideum* and *S.purpuratus* differ from *D.melanogaster* in that their actin genes are clustered, clustering does not appear to be a requirement for developmental regulation. Finally Couderc, *et al*. (1983) have recently reported that the expression of the 5C actin gene (on the X chromosome) is regulated by ecdysterone. It remains to be shown whether the increase in actin gene expression is specific, or reflects on overall stimulation of transcription by ecdysterone.

The actin genes therefore have been shown to be highly conserved, and are usually present as a multigene family. Members of this family can be either clustered or dispersed around the genome. Not all genes have been shown to be expressed but of those that are, some appear to be developmentally regulated. Some actin genes (the majority) possess introns, and the number, type and position of these is highly variable. The conserved nature of the actin genes, and their ubiquity plus the ability to be able to use them as a model system for developmental regulation of gene expression, have encouraged me to seek actin genes in the genome of P.falciparum.

1.8 Purpose of this Thesis

In the course of this Introduction I have described some of the characteristics of 3 classes of sequence. One illustrates a highly repetitious class, the second (rRNA genes) are moderately repetitive, transcribed but not translated; the third (actin genes) are developmentally regulated and code for a protein which is thought to have a crucial role in the life-cycle. A need for a better understanding of the biology of the human malaria parasite has also been clearly demonstrated, and by the study of the above mentioned P.falciparum sequences, a more complete understanding of the genomic organisation of the parasite should result.

CHAPTER 2

MATERIALS AND METHODS

2.1 Bacterial strains and phages

2.2 Continuous culture of *P.falciparum*

2.3 Extraction of parasite DNA and RNA

2.4 Cloning vectors

2.5 Media and solutions

2.6 Plating cells

2.7 Phage lambda titration

2.8 Preparation of lambda DNA

2.9 Preparation of plasmid DNA

2.10 Small scale preparation of plasmid DNA

2.11 Ethanol precipitation

2.12 Restriction enzyme digest

2.13 Ligation of DNA

2.14 Cloning

2.15 Transformation

2.16 *In vitro* packaging of lambda DNA

2.17 DNA gels

2.18 RNA gels

2.19 Protein gels

2.20 DNA transfer to nitrocellulose

2.21 RNA transfer to nitrocellulose

2.22 Protein transfer to nitrocellulose

2.23 Plaque screening by hybridisation with DNA or RNA probes

2.24 Nick translation of DNA

2.25 Hybridisation

2.26 End labelling of RNA by kinase

2.27 Immuno detection of actin on Western blots

2.28 Extraction of DNA/RNA from agarose gels by electroelution

2.29 Mapping of multiple sites for a single enzyme

2.30 Gel photography

2.31 Autoradiography

2.32 Electron microscopy

2.1 Bacterial Strains and Phages

All bacterial strains and phages used in this study are listed in Table 2.1. The yeast strain used S41 has the genotype mata/MATα, HO/HO arg 4.1/arg 4.1, cyh 1/cyh 1, and was a kind gift of Dr. Ian Dawes.

2.2 Continuous Culture of P.falciparum

Growth Medium (Traeger and Jenson 1976) RPM1 Medium 1640 (GIBCO Laboratories), 10.4 gms in 900 mls of double distilled water; HEPES (Sigma), 6 gms; gentamycin (Sigma) 100 mgs. The solution is diluted to 960 ml and sterilised by filtration through a Millipore filter 0.22 μm porosity. This solution may be stored for up to 1 month at 4°C. The medium is completed by the addition of filter sterilised 5% sodium bicarbonate to 0.2% final concentration. This medium without serum (RP) can be stored up to a week in the refrigerator. Complete medium (C.M.) is prepared from it by the addition of pooled serum (Blood Transfusion Service, BTS) to a final concentration of 10%.

Culture System - Preparation and Maintenance

Parasites are cultured in group Orh^+ blood (BTS) which has been washed three times with RP to remove the buffy coat layer. Each 75 cc tissue culture flash (Falcon) contains 10 mls of blood and 50 mls of C.M. Parasites are fed by removing spent medium, and adding fresh C.M. every 24 hr. The parasites are diluted every 48 to 72 hr. in fresh blood plus C.M. (haematocrit of 40%) to give a parasitemia of about 1%. Parasites are harvested when the desired quantity of red blood cells has a parasitemia of 10%. All cultures are grown under an atmosphere of 3% CO_2, 3% O_2, 94% N_2 (BOC).

2.3 Extraction of parasite DNA

Parasitized blood was harvested by centrifugation, washed, and resuspended in an equal volume of R.P. To this saponin is added

Table 2.1 Strains and Vectors

Strain/Vector	Genotype	Reference
ED8654	*supE*, *supF*, *hsdR*$^-$, *M*$^+$, *S*$^+$, *met*$^-$, *trp*R	Murray *et al* (1977)
HB101	*ara*, *arg*, *gal*, *hsdS*, *pro*, *recA*, *str*R	Boyer *et al* (1969)
NM514	*lyc7*, *hsdR*$^-$	Murray, N. M. (1983)
BHB 2688	N205 *recA* (λ*imm*^{434}, *c*I$_{ts}$, *b2*, *red3*, *Eam*$_4$, *Sam*$_7$)λ	Hohn (1979)
BHB2690	N205 *recA* (λ*imm*^{434}, *c*I$_{ts}$, *red3*, *Dam*$_{15}$, *Sam*$_7$/λ	Hohn (1979)
NMλ788*	*trpE*, (*att-red*), *c*I, *nin5*, *Wam*$_{403}$, *Eam*$_{1100}$, *Sam*$_7$	Murray *et al* (1977)
NM 1149*	λ*imm*^{434}, (*b538*)	Murray N.M. (1983)
pBR322		Bolivar *et al* (1977)

* Maps of vectors shown in Fig. 2.1

to give a final concentration of 0.1%, and the mixture left on ice for five minutes. Parasites released from lysed r.b.c. were pelleted at 10,000 g for 5 mins. washed 3 times in RP, and finally resuspended in 3.2 mls of RP. To this 0.8 mls of 0.5M EDTA was added giving a final concentration of EDTA of 100 mM.

Parasites were lysed by sodium lauroyl sarcosine (Sigma) at a final concentration of 4%, in the presence of Proteinase K (1 mg/ml) (Boehringer). The solution made up to 10 mls. 1 ml of ethidum bromide (5 mgs/ml) added, plus caesium chloride to give a density of ρ = 1.55. The DNA was centrifuged to equilibrium at 38K rpm for 48 hrs.

2.3(b) Extraction of Parasite RNA

Parasite RNA was extracted by J. Hyde as described by Hyde, *et al*. (1981 The rRNA species were purified by gel electrophoresis, see RNA gels.

2.4 Cloning vectors

The lambda cloning vectors used were either NM788 (Murray *et al*. 1977) or NM1149 (Murray 1983) (see Fig.2.1) The plasmid vector was pBR322 (Bolivar *et al*. 1977).

2.5 Media and Solutions

L Broth - Difco Bacto Tryptone, 10 gms, Difco Bacto yeast extract, 5 gms; NaCl 5 gms l^{-1}, pH 7.2.

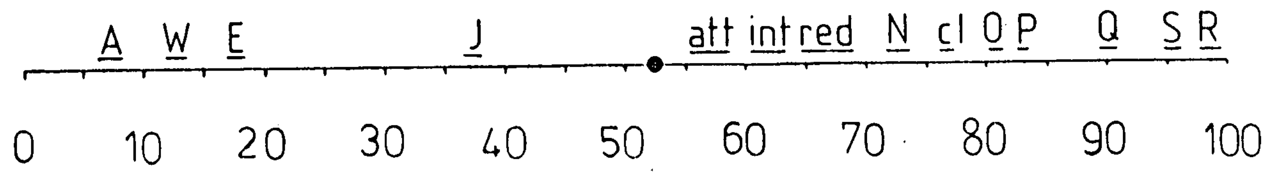
L-agar - Difco Bacto Tryptone, 10 gms; Difco Bacto yeast extract, 5 g; NaCl 10 gms; Difco agar 15 gms. per litre, pH 7.2

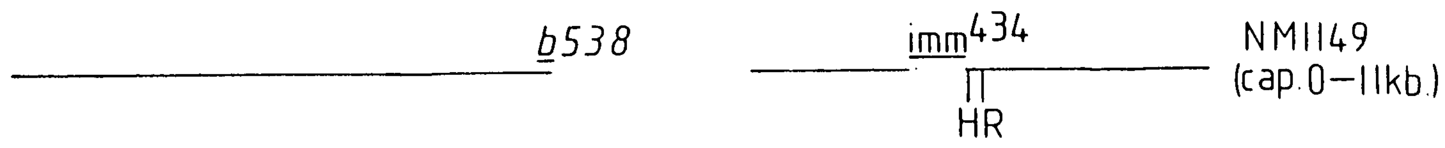
BBL agar -Baltimore Biological Laboratories trypticase 10 g; NaCl 5 g; Difco agar, 10 gms. per litre.

BBL top agar - As for BBL agar but only 6.5 gms Difco agar per litre.

Fig. 2.1

Lambda vectors. The map at the top of Figure shows some of the important genetic markers of phage λ, the scale 0-100 represents the length of λ^+DNA; ● is the phage attachment site. The HindIII (H) and EcoRl (E) cloning sites are shown. Given in parenthesis is the capacity of each vector.





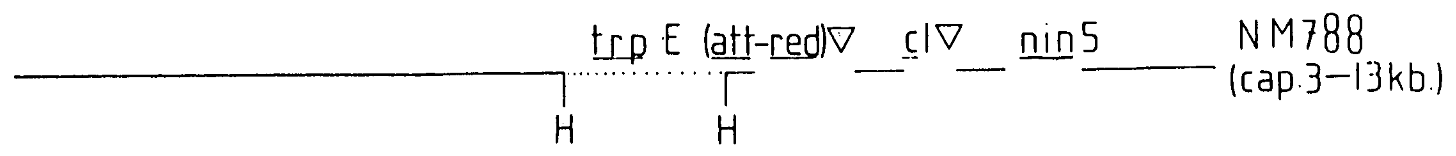


FIG 2.1

Phage buffer - KH_2PO_4, 3 g; Na_2HPO_4 (anhydrous), 7 gms; NaCl, 5 g; 10 ml 0.1M $MgSO_4$; 10 ml 0.1M $CaCl_2$; 1m 1% gelatin solution per litre.

Bacterial buffer - KH_2PO_4, 3 g; Na_2HPO_4 (anhydrous), 7 g; NaCl, 4 g; $MgSO_4$ $7H_2O$, 0.2 g; per litre.

YEPD - Difco Bacto yeast extract, 10 g; Difco Bacto peptose, 20 g; glucose, 2%; per litre.

TE buffer - 10 mM Tris-HCl pH 8.0, 1mMEDTA.

1 x SSC 0.015M tri-sodium citrate, 0.15M NaCl. Prepared as 20 x SSC and diluted as appropriate.

Phage Dialysis buffer - 10 mM NaCl, 50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$.

2.6 Plating cells

A fresh overnight culture of TGL70 or NM514 was diluted 20-fold in 100 ml L broth and incubated with aeration at 37°C until an A_{650} of 0.5-1.0 was reached. The cells were then centrifuged and resuspended in an equal volume of 10 mM $MgCl_2$. Bacteria prepared in this manner were stored at 4°C and could be used for up to two weeks.

2.7 Phage Lambda titration

The phage stock to be titred was diluted in phage buffer; and 0.1 ml of the appropriate dilution mixed with 0.1 ml of plating cells. 2.5 ml of molten BBL top agar was added and the mixture poured on to a BBL agar plate.

2.8 Preparation of lambda DNA

A single plaque of the relevant phage was picked in to phage buffer, and then plated for confluent lysis on 10 BBL agar plates. After overnight incubation the phage in the top agar were harvested,

and the phage eluted from the top agar into phage buffer, the top agar being removed by centrifugation at 5000 rpm for 15 minutes. This phage stock solution was then titred. The phage were added to a 250 ml culture of the plating cells when the A_{650} was 0.4. The multiplicity of infection (moi) was 3. The growth was followed until the culture lysed. The bacterial debris was removed by centrifugation at 5000 rpm for 15 minutes. The phage were pelleted by centrifugation at 18,000 rpm for 3 hours. After removal of the supernatant, the phage were resuspended in 10 mls of phage buffer, by gentle shaking overnight at 4°C. DNAse I (50 μg/ml) and RNAse (100 μg/ml) were added at this stage. CsCl was added to the resuspended phage to give a final density ρ = 1.45. The gradient was then centrifuged to equilibrium at 38K rpm for 48 hrs. The phage were removed from the gradient by withdrawal through the side of the centrifuge tube using a needle and syringe. The CsCl was removed by dialysis against phage dialysis buffer at room temperature for 2 hrs., two changes of 1l of buffer being sufficient. The phage DNA was extracted using an equal volume of redistilled phenol, which had been equilibrated against TE buffer. Three phenol extractions were followed by two ether extractions (to remove phenol), and the DNA ethanol precipitated. The DNA was resuspended in TE buffer, and the concentration estimated by absorbance at OD_{260}.

2.9 Preparation of plasmid DNA (Humphreys *et al*., 1975)

An overnight culture of the plasmid carrying strain was diluted in 500 ml of L broth (containing ampicillin at 50 mg per litre) and shaken at 37°C overnight. Chloroform (5 ml) was added and the culture shaken for a further 15 mins. The cells were centrifuged and resuspended in 6 ml 25% sucrose, 50 mM Tris-HCl, pH 8.1, 40 mM EDTA. Lysozyme

(1 ml of 10 mg/ml solution in sucrose mix) and 1 ml 0.5M EDTA pH 8.0 were then added. The mixture was kept on ice for 5 mins, swirling occasionally. Thirteen ml of Triton mix (2 ml 10% Triton X-100, 25 ml 0.5m EDTA pH 8.1, 10 ml 1MTris-HCl pH 8.1, H_2O to 200 ml) were added, and the solution kept on ice for 10 mins. swirling occasionally. The lysate was centrifuged for 30 mins at 15k rpm in a 8 x 50 rotor in a Sorvall RCSB centrifuge. The supernatant was decanted and, after addition of 0.95 g CaCl and 0.1 ml 5 mg/ml ethidum bromide per ml of supernatant, was centrifuged to equilibrium at 38k rpm for 48-72 hours at 18°C. The plasmid DNA was visualised by long wave u.v., and removed by side puncture. Ethidum bromide was removed by extraction with an equal volume of butan-1-ol, and the CsCl by dialysis against TE buffer. Three changes over three hours at 4°C, was sufficient to remove CsCl. Concentration of DNA was estimated by absorbance at OD_{260}.

2.10 Small scale preparation of plasmid DNA

The method used was essentially that of Birnboim and Doly (1979). 1 ml of an overnight culture was centrifuged in a microfuge tube for 1 min. The cell pellet was resuspended in 100 µl 2 mg/ml lysozyme, 25 mM Tris-HCl pH 8.0, 10 mM EDTA, 50 mM Glucose and left on ice for 30 mins. AFter addition of 200 µl of alkaline SDS solution (0.2 M NaOH, 1% SDS), the solution was left on ice for a further 5 minutes. This was followed by the addition of 150 µl of high salt solution (3M sodium acetate pH 5.0), and another hour on ice. The white precipitate that formed was removed by 5 mins centrifugation in an Eppendorf microfuge. To the supernatant 1 ml of cold ethanol was added, and the solution placed at -20°C for 30 mins. The DNA was pelleted by 2 mins. centrifugation, and then resuspended in 100 µl of 0.1M sodium acetate pH 6.0. To this

200 µl of cold ethanol was added and the solution placed at -20°C for 10 mins. The DNA was again pelleted by 2 mins. centrifugation, dried, amd resuspended in 50 µl of TE buffer. Half of this material was adequate for restriction, and could be visualised on agarose gels.

2.11 Ethanol precipitation

DNA or RNA was precipitated by either ethanol or iso-propanol. For the former a 1/10 volume of sterile 3M sodium acetate was added, followed by at least 2 volumes of cold 100% ethanol. The mixture was placed at -70°C for 15 mins. The DNA was pelleted by centrifugation in a microfuge for 5 mins. After removing the supernatant the pellet was washed in cold 70% ethanol, and then re-centrifuged. For iso-propanol precipitation a ½ vol. of sterile .6M ammonium acetate was added, followed by 2 volumes of isopropanol. The solution was left at room temperature for 10 mins., then centrifuged for 10 mins in a microfuge. The pelleted DNA washed with 70% isopropanol and recentrifuged.

2.12 Restriction enzyme digest

All restriction enzymes were purchased from Bethesda Research Laboratories with the exception of HindIII and Pst I, which were made in this Department. See Table 2.2 for the buffer conditions used. All digestions were at 37°C with incubation periods of 1-5 hrs. Reactions were terminated by heating at 65°C for 10 mins. Reaction volumes were from 10 µl-100 µl, and multiple digestions were performed simultaneously when possible. Otherwise those enzymes demanding lower salt concentrations were added first, followed later by addition of the appropriate amount of salt plus final enzyme.

2.13 Ligation of DNA

Only sticky ended ligations were performed and these were in 66 mM

Table 2.2 Buffer conditions for restriction enzymes

Enzyme	Recognition Sequence	Tris-HCl mM	pH	NaCl mM	KCl mM	$MgCl_2$ mM	DTT* 5mM	BSA* 100 µg/ml
AvaI	C↓PyCGPuG	6	7.4	60	-	10	+	+
BamHl	G↓GATCC	6	7.9	150	-	6	-	+
BglII	A↓GATCT	10	7.4	-	66	10	+	+
DraI	TTT↓AAA	6	7.6	100	-	6	+	+
EcoRl	G↓AATTC	100	7.5	50	-	5	-	+
HpaI	GTT↓AAC	10	7.5	-	20	10	+	+
HindIII	A↓AGCTT	10	7.4	60	-	7	-	+
Pst 1	CTGCA↓G	6	7.4	50	-	6	+	+
SalI	G↓TCGAC	6	7.9	150	-	6	+	+
HincII	GTPy↓PuAc	10	7.9	60	-	6	+	+

* DTT - dithiothreitol

BSA - bovine serum albumin

+ Py - Pyrimidine

Pu - Purine

Tris-HC pH 7.2, 1 mM EDTA, 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.1 mM ATP. Reactions were incubated overnight at 10°C with 10 units of T4 DNA ligase.

2.14 Cloning

(i) Cloning into phage lambda.

The vector and donor DNAs were digested with the appropriate restriction enzymes, and mixed to optimise the ratio of arms to potential inserts and the concentration of the DNA species (Maniatis *et a* (1982) Molecular cloning)). For NM1149 1 µg of insert was mixed with 2 µgs of arms and ligated in 10 µl. For NM 788 the arms were purified over a sucrose gradient, and the stuffer fragment removed. The cloning of a *Hind*III library into NM788 has previously been described (Goman *et al*. 1982).

(ii) Subcloning into pBR322

Plasmid and donor DNAs were digested with the appropriate restriction enzymes and then mixed in equimolar amounts to a total concentration of 40 µg/ml, and ligated. The ligated DNA was then transformed into HB101 and transformants selected on L agar plates containing ampicillin (100 µg/ml). Transformants were screened for the desired recombinant by small scale plasmid preparations (Birnboim and Doly, 1979).

2.15 Transformation

A 500 ml culture of HB101 was grown at 37°C with aeration until an A_{650} of 0.4 was reached. The cells were chilled on ice for 10 minutes, pelleted, resuspended in 200 ml of cold 0.1M $CaCl_2$ and left on ice for 25 mins. The cells were again pelleted and resuspended in 5 ml of $CaCl_2$-15% glycerol. These cells could be used immediately or stored at -70°C. They remain competent for up to two months. The

transforming DNA was mixed with 50 µl of competent cells, mixed, and kept on ice for 30 mins. A 2 min. heat shock of 42°C was followed by a further 30 mins. on ice. One ml of L broth was added, and the cells grown at 37°C for 1-3 hrs, before plating on selective plates.

2.16 In vitro packaging of lambda DNA

Freeze Thaw Lysate (FTL)

Three 500 ml cultures of BHB2688 were grown in L broth at 30°C. At A_{650} 0.3 the phage were induced by a 15 min. heat shock at 45°C. The induced bacteria were grown for 1 hr at 37°C, before harvesting by centrifugation. They were resuspended in 6 ml of cold 10% sucrose, 50 mM Tris-HCl pH 7.5, and dispensed into 2 10 ml ultracentrifuge tubes. To each tube 75 µl of fresh lysozyme solution (2 mg/ml in 0.25M Tris-HCl, pH 7.5), was added and the solution mixed well. A quick freeze in liquid N_2 was followed by gentle thawing. To each tube 75µl of buffer M1 (110µl H_2O, 6 µl 0.5 M Tris-HCl pH 7.5, 300 µl 50 mM spermidine, 100 mM putrescine neutralised with Tris base, 9 µl 1M $MgCl_2$, 75 µl 0.1M ATP neutralised with NH_4OH, 1 µl β-mercaptoethanol) was added and mixed gently. The samples were then centrifuged at 35000 rpm for 35 mins. The supernatant removed and dispensed in 50-100 µl aliquots and stored at -70°C.

Sonicate extract (SE)

A 500 ml culture of BHB2690 was grown at 30°C to an A_{650} of 0.3. The phage induced and harvested as before. The pellet was resuspended in 4.6 ml of buffer A (20 mM Tris-HCl pH 8.0, 3 mM $MgCl_2$, 0.5% (v/v) β-mercaptaethanol, 1 mM EDTA, pH 7.0). The resuspended cells were sonicated without foaming until the suspension was no longer viscous. The cellular debris was removed by centrifugation at 6000 rpm for 6

mins., and the supernatant aliquoted and stored at -70°C.

2.17 DNA gels

Large gels - DNA samples were prepared for electrophoresis by adding 10 µl of loading buffer (10% ficoll, 0.025% bromophenol blue). Electrophoresis was through horizontal gel slabs (28 x 14 x 0.5 cm) connected by wicks at each end to tanks containing 40 mM Tris, 20 mM tri-sodium acetate, 1 mM EDTA adjusted to pH 8.2 with glacial acetic acid, and ethidium bromide (500 µg per litre). The agarose, which had been melted in the same buffer was 0.5, 0.7, 1.0, 1.5, 2.0%. The samples were run into the gel at 10 volts/cm and then run overnight at 4 volts/cm. Phage λcI_{857} DNA restricted with HindIII and/or EcoRI provided molecular weight markers.

Mini gels - Small amounts of DNA (0.05 µg- 0.5 µg) were separated using a gel (10 x 5 x 0.15 cm) made up in Tris-borate buffer (89 mM Tris, 89 mM boric acid, 25 mM EDTA pH 8.2), containing ethidum bromide (500 µg per litre). The electrophoresis was performed under buffer at 40 volts/cm, usually for 30 mins. In this way aliquots of DNA could be checked for restriction, ligation or size.

2.18 RNA gels

The RNA was fractionated under denaturing conditions using an agarose-formaldehyde gel. The gel buffer (GB) used was 20 mM MOPS, 5 mM sodium acetate, 1 mM EDTA pH 7.0. Agarose of the required concentration (0.8%, 1%, 1.5% or 2%) was melted in 1 x GB and then cooled to 50°C. Formaldehyde was added to give a final concentration of 2.2M. The RNA samples were taken up in sample buffer (50% formamide 2.2 M formaldehyde, 1 x GB) and heated at 60°C for 5 mins. Loading buffer (16% ficoll, 0.025% bromophenol blue) was then added and the

samples run in at 5 volts/cm. The gel (28 x 14 x 0.5 cm) was run overnight at 2 volts/cm.

2.19 Protein gels

Proteins were fractionated according to their size using the sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) system first described by Laemmli (1970). However, in this case 3.75% stacking and 10% slab running gels were used. A mixture of high and low molecular weight size markers (Pharmacia) gave a range in molecular weight from 14,400 to 330,000 daltons. The proteins were visualised by silver staining, using the revised protocol of Morrisey *et al*. (1981).

2.20 DNA transfer to nitrocellulose

(i) Southern blots - The method used was essentially that described by Southern (1975). The agarose gel was prepared for transfer by denaturing in 0.5M NaOH, 1.5M NaCl, and then neutralising in 1M Tris-HCl, 3M NaCl pH 5.5. After soaking for 10 mins. in 2 x SSC the gel was blotted overnight in 20 x SSC. After transfer the nitrocellulose was rinsed in 2 x SSC before baking for 2 hr. at 80°C under vacuum.

(ii) An alternative method was also used (Smith and Summers, 1980). In this case the gel was also denatured in 0.5M NaOH/ 1.5M NaCl, but neutralised in 1M ammonium acetate/0.02M NaOH. The transfer requires no wicks, and can be bidirectional if required. DNA transfer takes 1 hour at room temperature. The filter was rinsed in 2 x SSC and baked at 80°C under vacuum.

2.21 RNA transfer to nitro cellulose

Northern blots - The agarose/formaldehyde gel was soaked in

20 x SSC for 30 mins. The gel was blotted in 10 x SSC as described above. After transfer the nitrocellulose was rinsed in 3 x SSC, air dried, and baked for 2 hrs. at 80°C under vacuum.

2.22 Protein transfer to nitro cellulose

Western blots - The method used was the electrophoretic transfer of proteins from PAGE gels described by Towbin et al. (1979). The gel was soaked for 10 mins. in the transfer buffer (192 mM glycine, 25 mM Tris), to reduce the amount of SDS, before electro-blotting. The transfer was carried out at 60 volts (0.3 amp) for 6 hrs.

2.23 Plaque screening by hybridisation with DNA or RNA probes

The method used was that described by Benton and Davis (1977). The nitro cellulose filter was placed against the BBL plates without prior treatment. After 10 mins. the filter was carefully removed, and placed on blotting paper soaked in 0.5 NaOH, the surface that had been in contact with the plaques being uppermost. After a further 10 mins. the filters were placed sequentially in 0.1M NaOH/1.5M NaCl, 0.5M Tris/1.5M NaCl pH 7.5, 0.2M Tris/2 x SSC pH 7.5. The filters were then air-dried before baking at 80°C under vacuum for 2 hrs.

2.24 Nick translation of DNA

DNA used for probing nitrocellulose filters was labelled with α-^{32}P-dCTP (Amersham), of specific activity 410 Ci/mM (0.8 mC/Ml). The method employed was essentially that of Rigby et al. (1977). Usually 0.5 µg of DNA were labelled and the activity incorporated was 1-10 x 10^6 cpm/µg. The incorporated counts were separated from the unincorporated by passing the reaction mixture over a Sephadex G-50 (fine) column.

2.25 Hybridisations

Whether DNA or RNA was used as a probe against southern or northern blots, the hybridisation was the same. It was 4 x SSC, 50% formamide, 0.1% SDS, 1 x Denhardt solution (0.02% BSA 0.02% polyvinyl pyrolidone 0.02% ficoll). When probing for sequences thought to be present in low copy number Dextran sulphate (Sigma) was added to the hybridisation solution at 5% concentration. The blots were hybridised overnight at 37°C on a rotary shaker. After hybridisation the filters were washed at 37°C with shaking in 0.1% SDS, SSC (2x, 1x, 0.5x), 4 x ½ hr. washes. This was followed by 4 x ½ hrs. washes in SSC (2x, 1x, 0.5x), before air drying and exposing. The stringency was varied depending on whether a homologous or heterologous probe was used. Blots could be dehybridised by boiling in 0.1 x SSC for 30 mins. and reprobed if required.

2.26 End labelling of RNA by kinase

RNA to be end-labelled was first hydrolysed to generate sufficient ends. Heating for 10 mins. at 90°C in hydrolysis buffer (5 mM glycine) 0.01 mM EDTA pH 9.5) yielded RNA fragments of approximately 100 bp. After hydrolysis the RNA was ethanol precipitated and resuspended in buffer (70 mM Tris-HCl pH 8.1, 10 mM $MgCl_2$, 50 mM KCl, 5 mM DTT, 1 mM spermidine). Addition of ^{32}P-ATP (0.01 mM) and 5 units of polynucleotide kinase (P.L. Biochemicals) was followed by incubation at 37°C for 1-3 hrs. The incorporated counts were separated from the unincorporated by passage over a Sephadex G-25 column. The column having first been treated with the RNAse inhibitor diethyl pyrocarbonate (Sigma) at 0.2% concentration.

EX BIBL. UNIV. EDINBURGEN.

2.27 Immuno detection of actin on Western blots

The technique used was enzyme linked immunosorbant assay (ELISA) described by Towbin *et al.* (1979). The western blot was first incubated with 10% egg albumin, (Crude powder grade II Sigma) 0.15M NaCl 0.01M Tris pH 7.4 for 1-3 hrs. Then it was washed for ½ hr in 0.15M NaCl, 0.01M Tris pH 7.4, with multiple changes, before incubating with the rat α yeast actin (a kind gift of John Kilmartin). The antibody (Ab) incubation buffer was 0.15M NaCl, 0.01M Tris pH 7.4, 5% egg albumin. After 1 hr. incubation, the filter was again washed and the rabbit α rat (Miles-Yeda) added at a 1:500 dilution. A further 1 hr incubation was followed by another ½ hr wash before the last Ab was added. This was goat α rabbit IgG conjugated with horse radish peroxidase (Miles-Yeda). The final incubation was for 1 hr, after washing the western blot was stained with 500 μl of imidazole buffer pH 7.4, 250 μl o-dianisidine dihydrochloride, 50 μl hydrogen peroxide (Sigma) 4.2 mls of H_2O. A brown colouration develops within 10 mins where the original Ag (actin) was located.

2.28 Extraction of DNA/RNA from agarose gels by electroelution

The fragment to be electroeluted was visualised under uv, and excised from gel. The piece of agarose containing the fragment was placed inside dialysis tubing and electrophoresis buffer added sufficient to cover agarose. The clipped dialysis sack was placed under electrophoresis buffer and electroeluted at 300 volts for 2 hrs. The agarose was removed, the tubing resealed and electrophoresis continued for a further 10 mins. in the opposite polarity. The buffer was removed from tubing, given a phenol extraction to remove traces of agarose, and then ethanol precipitated.

2.29 Mapping of multiple sites for a single enzyme

The method used was based on that described by Smith and Birnsteil (1976). The pBR322 recombinant was restricted with BamH1 and the end 'filled in' using the Klenow fragment of DNA polymerase (Boehringer). The 3' end labelled fragments were then partially digested with DraI (NBL enzymes), and fractionated over 0.7% agarose gel. Since the DraI sites in pBR322 were known, the partials due to the plasmid were easily distinguished from those derived from the insert. Only those partials with a common ^{32}P-labelled 3' terminus were seen, when the dried down gel was autoradiographed.

2.30 Gel photography

Gels were photographed under short wave uv light using Ilford HP5 professional 5" x 4" sheet film, and a red filter. Exposure was for 30 secs. followed by 5 mins. in Ilford Microphen developer. The relative mobilities of the DNA fragments were measured directly from the negative.

2.31 Autoradiography

Autoradiography was performed using CRONEX 4- X-ray film, cassette and lightning plus intensifying screens. The cassettes were stored at -70°C, for the appropriate time.

2.32 Electron Microscopy (Highton, et al. 1975)

The phage λ clones were partially denatured in a sodium hydroxide/formaldehyde solution for 90 secs. at 27°C. The DNA was spread on a hyperphase of 50% formamide. The single stranded size marker was M13 (Denhardt et al. 1978) and the double stranded size marker was pAT153 (Twigg et al. 1980).

CHAPTER 3

STRAIN VARIATION CHARACTERISED BY THE ORGANIZATION OF A CLASS OF REPETITIVE DNA

3.1 Introduction

3.2 Results and discussion

3.3 Conclusions

3.1 Introduction

It has been known for many years that striking changes occur in the physical properties of DNA when a solution of it is heated. This denaturation or melting of DNA represents the disruption of the double helix into its complementary single strands (Marmur et al. 1963). The process is characterised by the temperature of the midpoint of transition, Tm, which is also described as the melting temperature. When denatured DNA is cooled slowly, a substantial proportion of the original duplex reforms, this process is described as renaturation. The reassociation of complementary single strands can be assayed either by the amount of double stranded DNA retained on a hydroxyapatite (HAP) column, or by absorbtion of ultraviolet light. The rate of this reassociation follows second-order kinetics, and is given a rate constant k (Wetmur and Davidson, 1968), such that

$$Cot_{\frac{1}{2}} = {}^{1}/k$$

where Co is the initial DNA concentration and $t_{\frac{1}{2}}$ is the time required for half the DNA to be reassociated. Reassociation of DNA is usually followed in the form of a Cot curve, where the fraction of that which remains single stranded ($^{C}/Co$) or the fraction that has renatured ($1-{}^{C}/Co$) is plotted against the log of the Cot (Britten and Kohne, 1968).

Cot analysis of eukaryotic genomes has revealed them to be composed of 3 reassociating components (for review see Davidson et al. 1975). The slow component that renatures at high Cot values, usually represents approximately 60% of the genome, and comprises non-repetitive or unique sequences. An intermediate component renatures over a range of Cot values, corresponds to approximately 20% of the genome, and is

made up of middle repetitive sequences. Finally there is a fast renaturing component composed of highly repeated sequences that has the lowest Cot values. There is obviously some degree of overlap between the moderate and highly repetitive components of the genome.

In Plasmodia to date Cot analysis has only been performed on P.berghei and P.falciparum (Chance et al., 1978; Dore et al., 1980; Dore et al., 1983 and Hough-Evans et al. 1982). These studies reveal that both P.berghei and P.falciparum possess DNA of 3 reassociating components, i.e. the genome of Plasmodia comprises unique, moderately repetitive, and repetitive sequences. In P.berghei the repetitive proportion was estimated at 18% (Dore et al., 1980), whereas in P.falciparum it was estimated at 10% (Hough-Evans et al. 1982). This discrepancy is probably due to the fact that Hough-Evans et al. observed an initial reassociation of 9%, which they ignored since they also observed an initial reassociation of 5% (as defined by binding to the HAP column) in the E.coli standard. The observation of approximately 20% repetitive DNA has also been reported for P.falciparum (Goman et al. 1982). This rough estimate was derived from the proportion of a genomic library of P.falciparum cloned into phage λ (λHPF8), which rapidly hybridised to radio-labelled total parasite DNA. One of the clones identified in this way, λPFH8rep20, was used here to examine strain variation at the genomic DNA level.

Variation between and within strains of P.falciparum, has been characterised previously using biochemical and immunological markers (Wilson, R. J. M. 1980; Tait, A. 1981; McBride, et al. 1982), and genetic variation in malaria parasites has been recently reviewed (Walliker, D., 1982). The salient details are outlined below.

The two principle requirements for a conventional type of genetic analysis are controlled matings between strains of the organism concerned; and the availability of stable genetic markers differentiating the strains. Note that since a mosquito is allowed to feed on gametocytes from the two different strains, as it is technically difficult to use just male and female gametocytes from each strain, a certain degree of self as well as cross fertilization occurs. A certain number of recombinants will therefore be of parental type. For genetic markers, extensive use has been made of different forms of enzymes as revealed by their mobility when subjected to electrophoresis. The use of these and other markers demonstrated that *Plasmodia* undergo a eukaryotic type of life cycle, the blood forms being haploid. The enzyme studies showed that a considerable degree of random mating occurs in populations of a single species, allowing new types of variant forms to arise readily.

An interesting new observation is that a single strain of *P.falciparum* is capable of displaying different forms of antigenic determinants (Hommel *et al.* 1983). This type of antigenic variation has been known for some time in *P.knowlesi* (Brown *et al.*, 1965). Both cases concern antigenic determinants on the surface of infected erythrocytes. In *P.falciparum* it is not clear whether the antigenic determinant is parasite encoded, or results from parasite modification of an erythrocyte antigen, since its presence was detected by fluorescent antibody only. In *P.knowlesi* the antigenic determinants are parasite encoded, and the major variant antigens (210,000 and 190,000), have recently been identified (Howard *et al.*, 1983). It has yet to be determined if the alteration in size of the variant

antigen is a consequence of expression of antigens of different size, or a difference in modification of a single antigen. The work reported by Hommel, et al.(1983) is the first indication that P.falciparum may be capable of varying the expression of particular antigenic determinants. However, the full implications of this observation can only be assessed after the antigens involved have been identified and characterised.

An understanding of strain variation and expression of parasite antigens, is important in research aimed at recognition of clinically important antigens as a prelude to the development of a vaccine against malaria. The aim of the work presented here was to establish if it is possible to correlate strain variation with an alteration in the genomic organization of P.falciparum.

3.2 Results and Discussion

The construction of a genomic library from the Gambian strain of P.falciparum HG13 (Butcher 1981), in phage lambda has previously been reported (Goman et al. 1982). This library was constructed from DNA extracted from an asynchronous in vitro culture of asexual inter erythrocytic blood stages of P.falciparum. Three clones containing repetitive DNA were identified and isolated from the library by their strong homology to total parasite DNA labelled with ^{32}P-containing dCTP by nick-translation (see Methods 2.24). Since a clone containing repetitive DNA, unlike those with unique sequences, should readily find homologous, labelled molecules in the probe and become heavily labelled itself (Goman et al. 1982). Three clones were selected for further study, and were found to have inserts of 7.0 kb, 5.0 kb and 7.1 kb. All three clones cross hybridise to each other, and λPFH8rep13 (5.0 kb) was shown to be more than 70% A+T (Bone et al., 1983). One of these clones

λPFH8rep20 (7.1 kb) was used in the study presented here. Its parasite DNA insert has recognition sites for BglII and HpaI, but no sites for BamHl, SalI, PstI, and EcoRl (M. Goman, pers. comm.).

When genomic DNA extracted from P.falciparum, was digested with HindIII, and fractionated on a 0.7% agarose gel, a large number of discrete bands are seen, indicating the presence in P.falciparum of repetitive DNA (Fig. 3.1). Although the pattern is very complex differences can be observed between the two isolates (parasite lines derived from separate patients) compared. Both isolates are from Thailand but one, Kl (Thaithong et al. 1981), is resistant to the anti-malarial drug pyrimethamine, and is shown in track 3. The other, Tak 9 (Rosario, 1981) is sensitive to pyrimethamine, and is shown in track 4.

When the DNA is transferred to nitrocellulose and hybridised with λPFH8 rep 20, a finite number of bands are observed (tracks 5 and 6), moreover dramatic differences can now be seen between the two isolates (compare tracks 5 and 6). These differences must be due to sequences homologous to λPFH8rep20 having different genomic environments in the two isolates, i.e. the sequences, and/or their locations are not conserved. When the same DNA was hybridised to pPFribl (see Chapter 4) an identical pattern of hybridisation is observed (tracks 7 and 8). This pattern of hybridisation is always observed when pPFribl is hybridised to HindIII digested genomic DNA (see Chapter 4), and allows one to conclude that the difference in the hybridisation patterns for Tak9 and Kl reflect differences in the isolates, and are not a result of partial digestion by HindIII.

The data presented so far demonstrate that λPFH8rep20 can recognise variation between two isolates of *P.falciparum*. However isolates have previously been shown to contain several different genetic types of parasite. Rosario (1981) prepared clones from the isolate Tak9 (T9), and showed that while the original uncloned material exhibited a mixture of several enzyme variants, single clones showed only one form of each enzyme. These and other clones of T9 have been examined for variation in other characters 1) proteins, 2) drug susceptibility, 3) antigens, 4) knobs on infected erythrocytes 5) electrophoretic variants of iso-enzymes (Thaithong *et al*. 1983). This offered a unique opportunity to compare the hybridisation pattern of λPFH8rep20 to their genomic DNA, and to see if this could be correlated with the other parameters of variation.

Genomic DNA was extracted from 6 clones of T9, digested with a set of restriction endonucleases and fractionated on agarose gels. After transfer to nitrocellulose the DNA was hybridised with labelled λPFH8rep20 and pPF*ribl*. Hybridisation with pPF*ribl* was included as an assay for limit digestion, as its pattern of hybridisation to genomic DNAs restricted with these enzymes is known (see Chapter 4). Genomic DNA from the 6 clones of Tak9 was first restricted with *Hind*III and fractionated on a 0.7% agarose gel (Fig. 3.2). Note that some of the genomic DNAs were poorly digested particularly clone T9.96 (tracks 2 and 9). This was confirmed by hybridisation with pPF*ribl* (Fig. 3.3b), compare the pattern with that in Fig. 3.1 tracks 7 and 8. Inspite of the poor restriction T9.94 (Fig. 3.3a track 1) is clearly different from T9.106 (track 6), and both differ from T9.97 (track 3) and T9.102 (track 5).

Fig. 3.1

Gel and hybridisation analysis of total *P.falciparum* DNA. Compared are DNA extracted from two different *P.falciparum* isolates, K1 (tracks 1, 3, 5 and 7) and Tak9 (tracks 2, 4, 6 and 8). Tracks 1 and 2 show unrestricted total DNA. Tracks 3 and 4 show *Hind*III digests. The fractionated DNA was transferred to nitrocellulose and hybridised to either the repetitive clone λPFH8rep20 (tracks 5 and 6) or the ribosomal RNA clone pPF*rib1* (tracks 7 and 8). The size markers are *Hind*III fragments of λDNA.

Fig 3.1

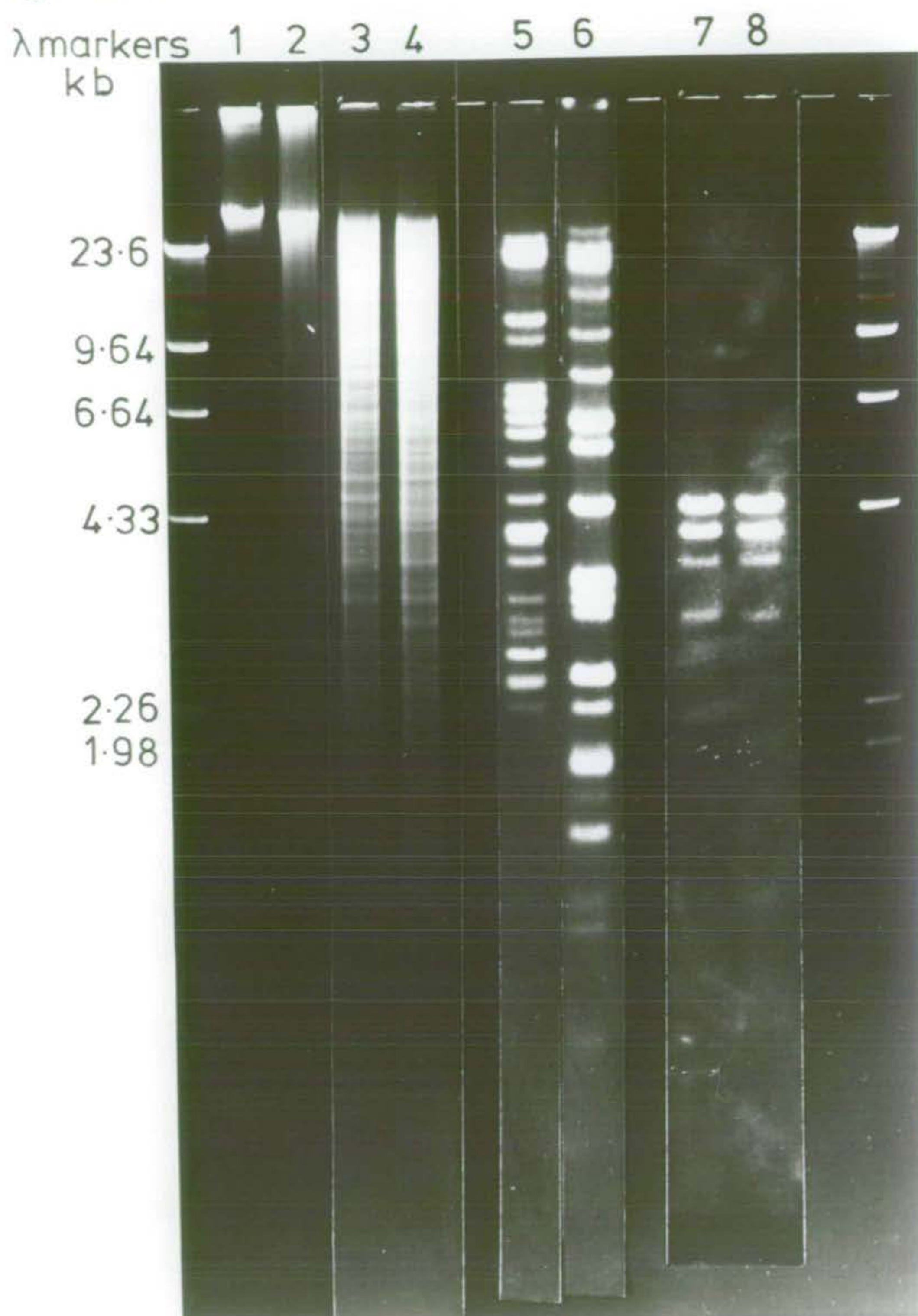
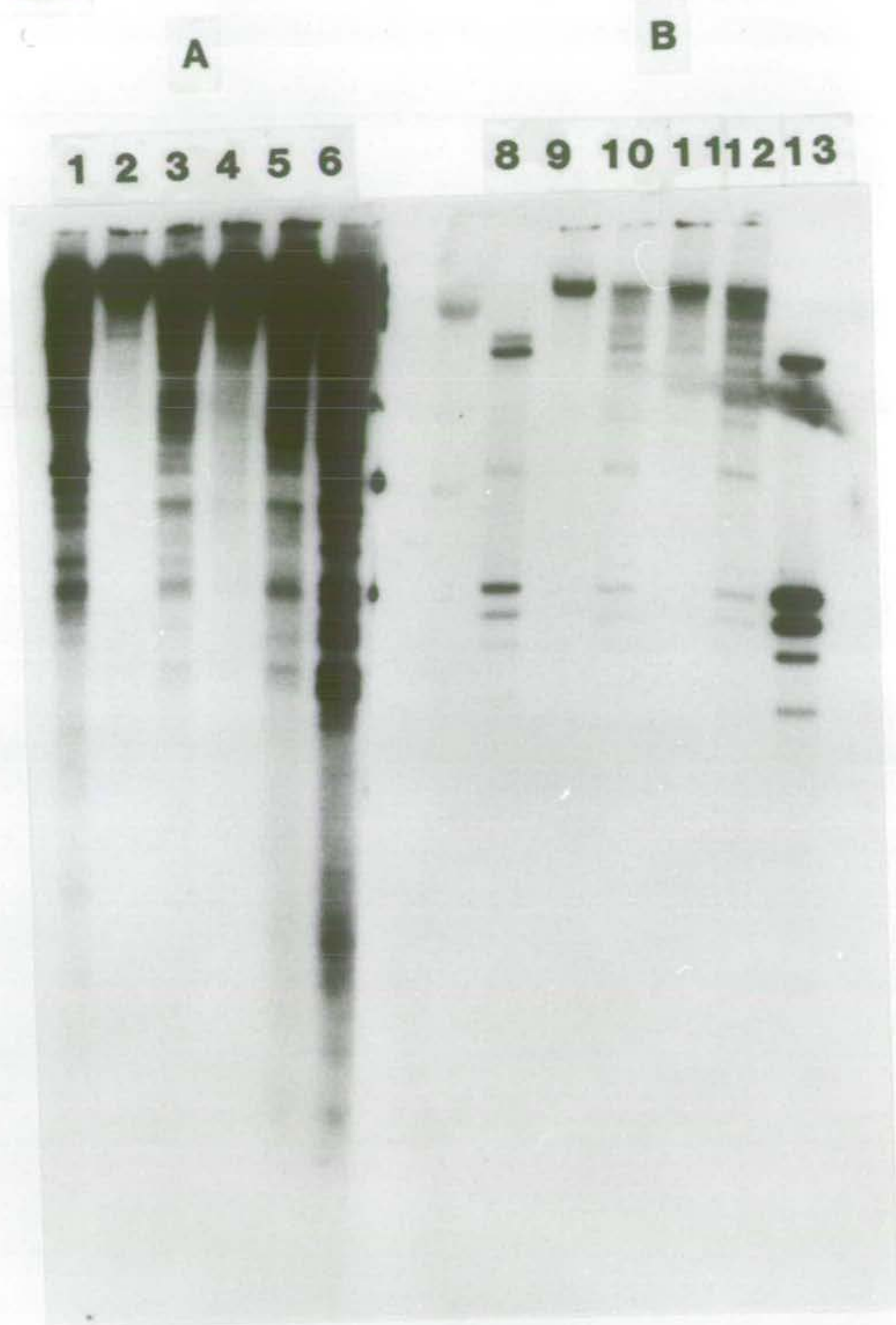


Fig 3.3



The same clones of T9 were then restricted with EcoRl and fractionated on a 0.7% agarose gel (Fig. 3.4). Again the DNA from T.9.96 was poorly digested (tracks 2 and 9) as well as T9.94 (tracks 1 and 8). The other appeared to have been digested to completion, and this was confirmed by the hybridisation pattern shown to pPFribl (Fig. 3.5b). The pattern of hybridisation shown to λPFH8rep20 (Fig. 3.5a) is quite significant. All clones displayed an identical pattern with the exception of track 2 where some EcoRl partials are seen. The majority of hybridisation was to high molecular weight fragments poorly resolved by the gel system used. λPFH8rep20 also showed homology to a fragment of approximately 11 kb whose size was invariant in the 6 clones examined. The presence of this fragment is dependent upon restriction as no homology is seen in that region when unrestricted DNA is probed with λPFH8rep20 (Fig. 3.7a, track 6). The low degree of hybridisation to the unrestricted DNA is probably due to poor transfer to nitrocellulose of this very high molecular weight material. The significance and possible interpretations of this observation are discussed in Section 3.3.

Finally genomic DNA from 5 of the clones was restricted with the enzyme DraI. This enzyme was chosen as its recognition sequence is TTT▼AAA, and considering that the genome of *P.falciparum* has an A+T content of 81% (Goman *et al*. 1982; Pollock *et al*. 1982), was expected to cut at high frequency. As predicted the bulk of the genomic DNA was reduced to fragments of small size (Fig. 3.6), surprisingly a number of discrete high molecular weight fragments were generated. Transfer of this DNA to nitrocellulose and

Fig. 3.4

Ethidium bromide stained gel of EcoRl digestions of clones of T9.

Track		
Track 1	T9.	94
2	T9.	96
3	T9.	97
4	T9.	99
5	T9.	102
6	T9.	106
7	HindIII λ DNA size marker fragments.	

Tracks 8-13 are a direct repeat of tracks 1-6.

Note DNA in tracks 1, 2, 8 and 9 poorly digested.

Fig 3.4

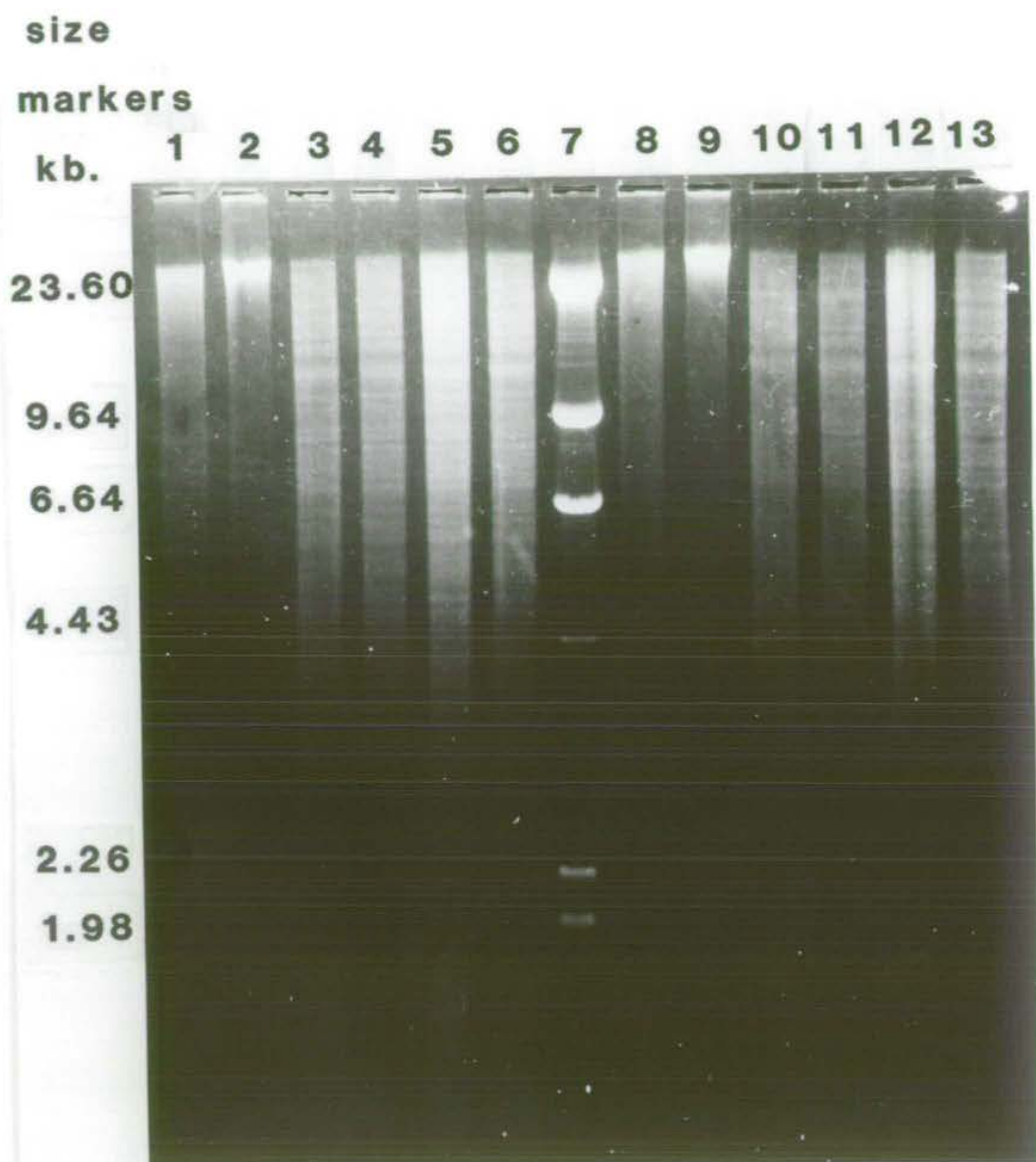


Fig. 3.5

Autoradiograph of Southern transfer of gel in Fig. 3.4.

Panel A: hybridised to λHPF8rep20. Note extra bands observed in track 2, are due to partial digestion. Note also common band on invariant size.

Panel B: hybridised to pPF*ribl*. Hybridisation pattern observed for tracks 10-13 indicate complete digestion of DNA (see text).

Fig3.5

Fig. 3.6

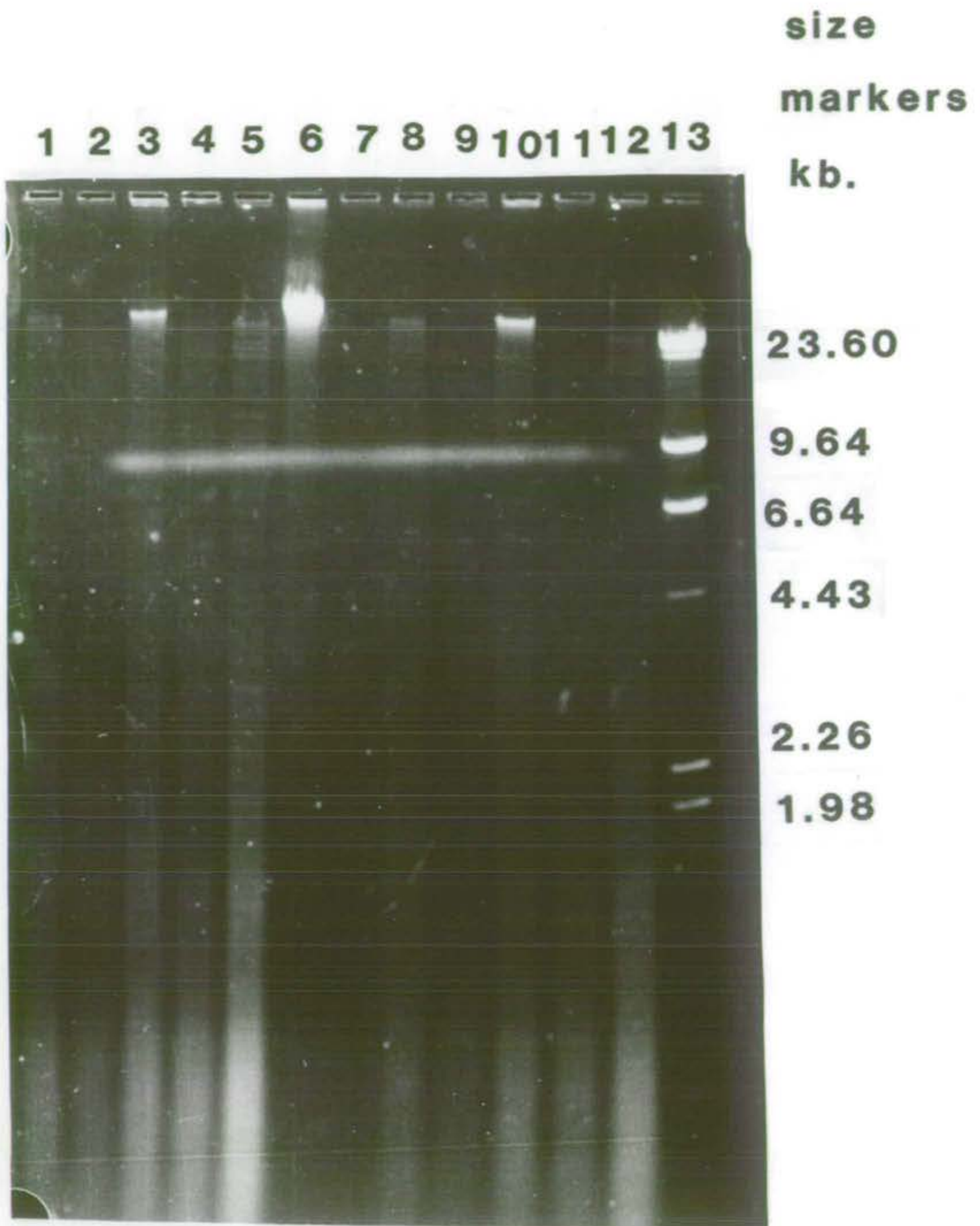
Ethidium bromide stained gel of DraI digestions of clones of T9.

Track	
1	T9. 94
2	T9. 96
3	T9. 97
4	T9. 99
5	T9. 106
6	undigested T9

Tracks 8-12 direct repeat of tracks 1-5.

Track 13 HindIII λ DNA size marker fragments. Note line across gel is an artifact introduced during photography.

Fig 3.6



hybridisation with λPFH8rep20 revealed that it is to these high molecular weight fragments that homology is observed (Fig. 3.7). The pattern generated by the restriction enzyme DraI most clearly differentiates between the clones of T9. The patterns reveal that T9.94, T9.96 and T9.106 are clearly different (tracks 1, 2 and 5), whereas T9.97 and T9.99 appear to be related (tracks 3 and 4). A longer exposure of the same blot identifies more differences between lower molecular weight fragments, and the pattern observed when pPFribl is used as a probe indicates that the digest was complete (Fig. 3.8). Although the differences between the clones of T9 may be more clearly revealed by the use of some other restriction endonuclease, those demonstrated by DraI are adequate to differentiate between clones of a single isolate.

3.3 Conclusions

As noted previously these clones were also being characterised by Thaithong and her co-workers using a different set of parameters. With kind permission a summary of their results is presented in Table 3.1, note that clone T9.99 is not included in the table. Depending on the response of each clone to the 5 different parameters used to assay for variation, each was assigned to a clonal type. The clones used in this study are circled.

The important observation from the hybridisations with λPFH8rep20 is that there is perfect correlation between the genomic hybridisation analysis and the other assays of variation. This is illustrated by T9.99, which was identified by the DraI pattern to λPFH8rep20 as being the same as T9.97. Subsequent work by Walliker *et al.* (pers. comm.) has confirmed that T9.99 and T9.97 show the same variation in enzymes,

Fig. 3.7

Autoradiograph of Southern transfer of gel in Fig. 3.6.

Panel A: hybridised to λPFH8rep20. Note clear differences in hybridisation patterns, with the exception of tracks 3 and 4.

Panel B: hybridised to pPF*ribl*. Note at exposure presented very little hybridisation observed. For longer exposure see Fig. 3.8.

Fig 3.7

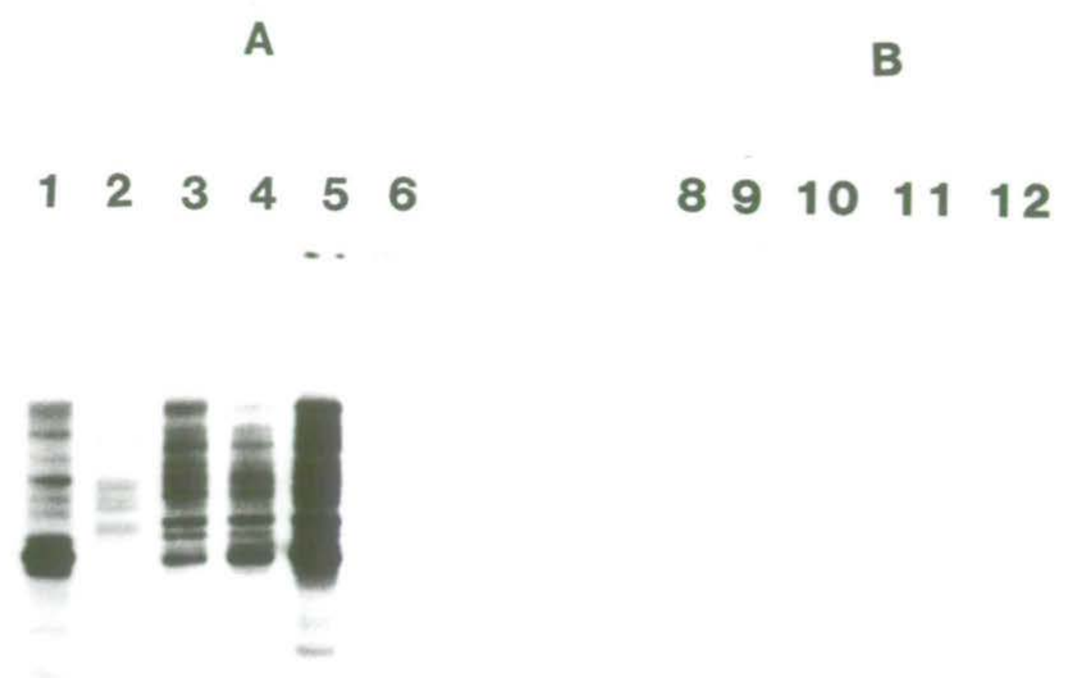


Fig. 3.8

A longer exposure of the autoradiograph presented in Fig. 3.7.

Panel A: this exposure reveals further differences between low molecular weight fragments, compare tracks 1 and 5. Note also that homology is only detected to unrestricted T9 DNA track 6.

Panel B: hybridisation detected to pPFribl is now observed. Pattern indicative of complete digestion (see text).

Fig 3.8

Table 3.1

The reference numbers of the cultures are given as they appeared in the dilution cloning experiments (Thaithong *et al*. 1983). The enzymes shown are glucose phosphate isomerase (GPI))EC5.3.1.9); adenosine deaminase (ADA) (Ec3.5.4.4); peptidase (PEP) (EC3.4.11 or 13). The numbers refer to variants of each enzyme. Drug susceptibility values refer to minimum inhibitory concentration (MIC) in molar (M) units. Antigen diversity is shown by indirect immunofluorescence test with ten monoclonal antibodies. Presence or absence of reaction is indicated by + or - respectively. -/+ indicates mixtures containing a small (*ca*.) 2% proportion of + parasites. Knobs: presence or absence is indicated by + or - respectively. Proteins by 2D gel electrophoresis; A-K are six selected protein spots on the 2D gels, the numbers refer to variants (see Thaithong, *et al*., 1983). The inferred 'clone-types' are indicated by a roman numeral, and indicates clones which are alike in all characters observed.
NT = not tested.

Table 3.1

CHARACTERIZATION OF CLONES PRODUCED BY DILUTION OF ISOLATE T9 OF P.FALCIPARUM

		Original uncloned isolate (T9)	Reference nos. of cultures produced by dilution of original isolate									
			(102)	(106)	100	101	98	(97)*	(96)	107	(94)	105*
Enzyme types	GPI	1/2	1	1	1	1	2	1	2	2	1	1
	ADA	1/2	1	1	1	2	1	2	1	1	1	1
	PEP	1/2	1	1	1	1	2	1	2	2	1	1
Drug susceptibility (MIC)	Chloroquine	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-7}	10^{-6}	10^{-7}	10^{-8}	$5x10^{-7}$	10^{-6}
	Pyrimethamine	NT	$5x10^{-5}$	$5x10^{-5}$	$5x10^{-5}$	$5x10^{-5}$	$5x10^{-5}$	$5x10^{-5}$	10^{-9}	10^{-9}	10^{-9}	10^{-5}
Antigens	Monoclonal antibodies 5.1	+	-	-	-	-	+	-	+	+	+	+
	7.3	+	-	-	-	-	-	-	-	-	+	+
	9.2	+	+	+	+	+	+	+	+	+	-	-/+
	9.5	+	-	-	-	+	-	+	-	-	-	-
	12.2	+	-	-	-	-	+	-	+	+	-	-
	12.1	+	+	+	+	-	+	-/+	+	+	-	-/+
	10.4	+	+	+	NT	+	NT	+	+	+	-	-/+
	9.4	+	+	+	NT	+	-	+	-	-	-	-/+
	12.3	+	+	+	+	+	+	+	+	+	+	-/+
	9.21	+	+	+	+	+	+	+	+	+	-	+
Knobs		+/-	+	+	+	+	+	+	+	+	-	+
Proteins by 2-D electrophoresis	A	2/4	1	1	1	1	1	1	2	2	4	2/4
	C	1/2/3	2	2	NT	2	2	2	1	1	2	2
	D	2/4/5	5	5	5	4	5	4	4	4	5	4
	E	2/3/4	4	4	4	4	1	4	3	4	2	2
	F	1/2/3	2	2	2	2	2	2	2	2	3	1
	G	1/2	2	2	2	2	2	2	2	2	2	1
	K	3/5/6	6	6	6	6	3	6	3	3	6	5
Inferred clone type			I	I	I	II	III	II(+III?)	IV	V	VI	VII(+?)

drug susceptibility, antigens, possessions of 'knobs', and enzyme variants, and have been assigned the same clonal type. Note that T9.99 and T9.97 are not pure clones and are mixtures of clonal types II and III.

The evidence presented here, although preliminary, raises interesting questions about the genetic basis of the observed variation between strains of P.falciparum. The process underlying the variation in hybridisation patterns of λPFH8rep20, has at least 3 possible explanations. The first is that λPFH8rep20 is a transposable element which transposes at high frequency. Transposition elements are known to exist in eukaryotes, in yeast there are about 35 copies of the transposable element Ty1 per genome, and their chromosomal distribution differs between strains (Cameron, et al. 1979). According to this view the insertion of λPFH8rep20 at different points within the genome of P.falciparum, results in differing patterns of hybridisation between the strains compared.

In one case λPFH8rep20 is inserted into a piece of repetitive DNA defined by EcoR1 sites, which are conserved and 11 kb apart. This is the invariant 11 kb fragment identified by λPFH8rep20 in all clones. In the majority of cases λPFH8rep20 is inserted in EcoR1 fragments of high molecular weight. Note that λPFH8rep20 itself does not possess EcoR1 sites.

If the frequency of transposition were high enough it should be possible to detect it in the following way. A pure clone of P.falciparum could be maintained by in vitro culture of the asexual blood stages. At various points over a time course, a proportion of the parasites could be harvested, their DNA extracted, and the

hybridisation pattern of λPFH8rep20 determined. Variation in the pattern would indicate that transposition had occurred. It is evident that only a high rate of transposition in the blood stages could be detected in this way; however, using essentially this protocol, transposition of Ty1 was detected in yeast (Cameron *et al*., 1979). The observation that T9.97 and T9.99 give identical hybridisation patterns, even after continuous *in vitro* culture of the pure clones, argues against this hypothesis.

A second possibility is suggested by the copia family of elements in *Drosophia* (for review see Rubin, G. M., 1983). Here λPFH8rep20 is dispersed around the genome by a transposition event that occurs at low frequency. The strain variation shown by the hybridisation pattern of λPFH8rep20 is stable, and due to the location of λPFH8rep20 at different sites within the genomes of the strains compared. The observation that strain differences are observed with both *Hind*III and *Dra*I, make it difficult to accept restriction site polymorphism as an explanation for the variation in hybridisation pattern. In depth analysis of the invariant 11 kb *Eco*R1 fragment may reveal whether λPFH8rep20 is part of a transposable element (e.g. by the possession of direct repeats at its termini), this 11 kb fragment should be readily identified in an *Eco*R1 genomic library constructed from T9 DNA.

The third possibility does not have to involve active transposition. The heterogeneity between the strains can be explained providing sequences homologous to λPFH8rep20 are repeated and located at different loci in the genomes of the strains compared. Recombination between the strains would then generate the diverse

patterns observed. The invariant 11 kb fragment results from the occurrence of λPFH8rep20 in a repeated stretch of genomic DNA with 2 conserved EcoRl sites. The majority of homology to high molecular weight DNA is due to a lack of EcoRl sites in these regions. Why sequences homologous to λPFH8rep20 occur only in stretches of DNA bereft of EcoRl sites is unknown at this time.

The involvement of recombination in strain variation can be tested directly. This is most easily done using the rodent malaria P.chabaudi, as it is technically difficult to maintain the complete life cycle of P.falciparum in the laboratory. Mice could be infected with pure haploid clones of two different strains of P.chabaudi, and after the establishment of a reasonable parasitaemia, uninfected mosquitos allowed to feed on the gametocytes in the blood. The parasites go through sexual reproduction in the mosquito, before it is allowed to feed on uninfected mice. The parasite DNA from the two groups of mice could be extracted and the hybridisation patterns between the pre-meiotic and post-meiotic parasites compared. An alteration in the hybridisation patterns would indicate that recombination had occurred.

The evidence presented here suggests that strain variation in P.falciparum may be due to transposition of a particular DNA sequence, and/or a high frequency of recombination between inter-breeding strains. A high frequency of recombination can also be inferred from the failure to detect any linkage between the 10 genetic markers so far examined (D. Walliker, pers. comm.). The recent demonstration of the variation in an antigenic determinant in P.falciparum may be significant in the light of current efforts to

produce a protective vaccine to malaria. However, the molecular basis of this antigenic variation in malaria has yet to be determined. Antigenic variation in *Trypanosomes* is known to be due in some cases to transposition to an 'expression site', followed by recombination (for review see Borst and Cross, 1982). It is not unreasonable therefore to suppose that transposition combined with recombination may have a role in antigenic variation in *Plasmodia*.

CHAPTER 4

ISOLATION AND CHARACTERISATION OF rDNA FRAGMENTS FROM THE GENOME OF *PLASMODIUM FALCIPARUM*

4.1 Introduction

4.2 Identification of genomic fragments homologous to rRNA probes

4.3 Isolation of rDNA clones from λPFH8 library

4.4 Mapping of the A+T rich regions by partial denaturation

4.5 Physical characterisation of the 4.1 kb and 4.2 kb rDNA fragments

4.6 Mapping of the rRNA coding sequences in pPF*rib1* and pPF*rib2*

4.7 Isolation and characterisation of a rDNA fragment homologous to 28S rRNA from the λRIKI library.

4.8 Estimation of the number of copies of the 4.4 kb and 4.2 kb fragments in the genome of *P. falciparum*.

4.9 Conclusions

4.1 Introduction

The sequences of eukaryotic genomes can be grouped into 3 classes (1) highly repetitive, (2) middle repetitive, (3) unique or single copy, as defined by renaturation kinetics. The proportion of each varies from organism to organism, the single copy class ranging from 50-70%, and the highly repetitive class comprising 10% of the genome (for review see Davidson *et al*. 1975). The remainder of the genome is comprised of middle repetitive sequences, a significant proportion of which is rDNA. For example the rDNA sequences in *Drosophila melanogaster* (*Dm*) comprise 2% of the total genome, and 20% of the middle repetitive DNA (Spradling *et al*. 1981).

Studies on the rRNA genes from a wide range of organisms reveal them to be repeated and conserved in their coding sequence. The abundance and conservation of rDNA sequences, has made them readily accessible for analysis (Long and Dawid 1980). In all species studied so far, the major rRNA species of eukaryotic organisms are called 28S, 18S, 5.8S and 5S according to their approximate sedimentation coefficients. The 28S, 5.8S and 5S are found in the large (60S) subunit and 18S in the small (40S) subunit of the mature ribosome. The three larger rRNA species (28S, 18S and 5.8S) are generated from a single precursor rRNA molecule (Perry, R. P. 1976). DNA complementary to the precursor is subdivided into 5' external transcribed spacer, an 18S rRNA gene, an internal transcribed spacer, and a 28S rRNA gene, in that order. The 5.8S rRNA gene is located in the internal transcribed spacer region. In many organisms rRNA genes are located in tandem, separated by a sequence called the non-transcribed spacer. In some cases the repeating unit also codes for 5S rRNA, but in

general pre rRNA and 5S rRNA genes in eukaryotes are not linked. For a more comprehensive review of rRNA see Chapter 1.

The rRNA genes in P.falciparum were chosen for study since their conserved sequence should permit their indentification by homology to cloned rRNA genes from another organism. The availability of their gene product (rRNA) enables characterisation of their coding regions, and the highly conserved nature of these regions allows comparative study between P.falciparum and other eukaryotes. In particular it will be of interest to determine whether the rRNA genes are organised in the characteristic eukaryotic sequence of 5'-18S-5.85S-28S-3'. Is this sequence dispersed (like E.coli), or tandemly repeated? In addition are the rRNA genes present in multi copy and what is their location? Does the copy number vary with development? Furthermore, is the 28S gene interrupted, and what kind of rRNA processing takes place? The work presented in this Chapter addresses itself to these questions.

4.2 Identification of genomic fragments homologous to rRNA probes

Previous work in this laboratory (Hyde et al. 1981) has shown that the base composition of P.falciparum (P.f.) rRNA is more similar to Drosophila melanogaster (Dm) rRNA, than human, Xenopus or E.coli, when the G+C content is measured. For this reason the rRNA genes of Dm cloned into pBR322 were chosen as a potential probe for the rRNA genes of P.f. This recombinant plasmid pDm 238/29 was a kind gift of Dr. D. Glover. For convenience a map of pDm 238/29 is presented in Fig. 4.1. Note that it contains an EcoRl genomic fragment (Dm 103) encoding the 18S, 5.8S and 28S rRNA genes (Glover et al. 1977). To demonstrate that pDm 238/29 showed homology to P.f. rRNA, 0.5 μg of

Fig. 4.1

Map of Dm103 showing the arrangement of the 18S and 28S rDNAs and indicating its origin from the chromosomal rDNA. Dm103 is shown here as a complete repeating unit excised from the chromosomal rDNA by cleavage with EcoR1. These sites correspond to the two EcoR1 sites in pDm238/29. Note that the 28S rRNA gene contains an intervening sequence (ivs) and that the non-transcribed spacer is shown by thicker line. Coding regions are shown hatched.

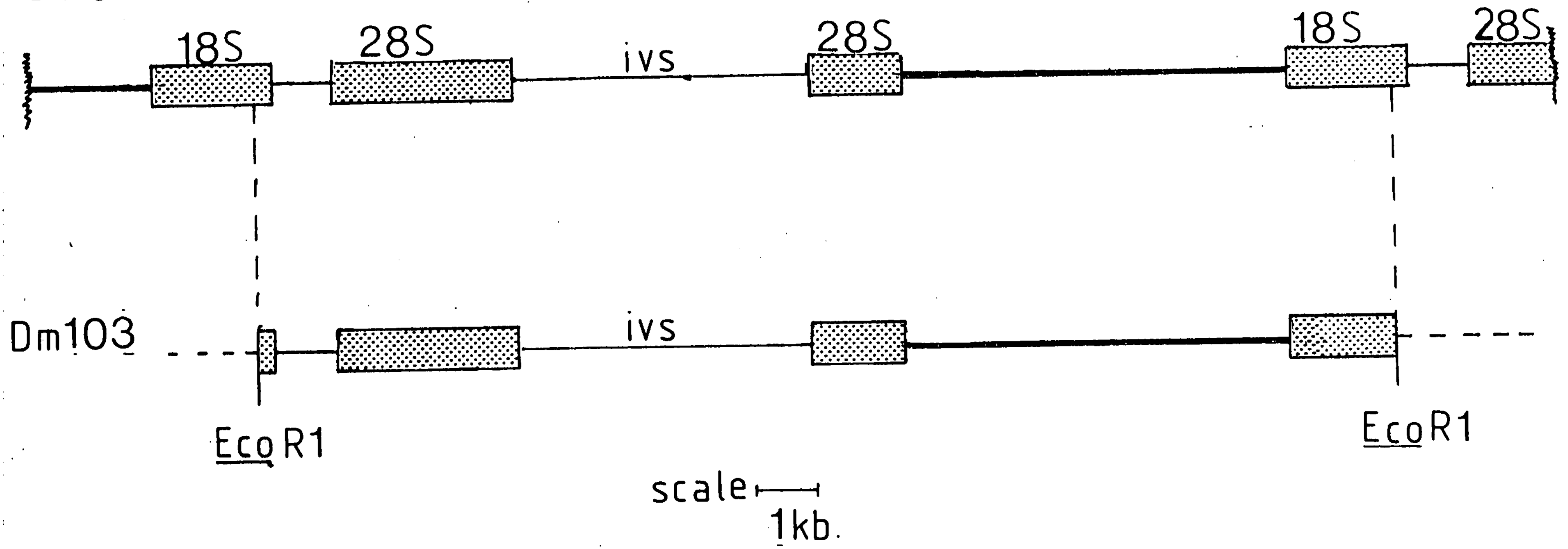


FIG. 4.1

rRNA was fractionated on an agarose/formaldehyde gel, transferred to nitrocellulose and probed with ^{32}P-labelled pDm 238/29 (Methods). An autoradiograph of the transfer is presented in Fig. 4.2, track 1, where it can be seen that there is good homology to both the large rRNA subunit (s.u.) and the small rRNA s.u. These s.u's. were operationally defined as 28S and 18S due to their co-migration on gels with known rRNA standards (Goman et al. 1982).

It is possible to show which fragments of P.f. genomic DNA are homologous to pDm 238/29 and to P.f. rRNA; genomic DNA from strain K1 (Thaithong et al. 1981) was digested with a set of restriction endonucleases, fractionated on a 0.7% agarose gel (Fig. 4.3a) and transferred to nitrocellulose. The filter was first probed with ^{32}P-labelled pDm 238/29 under low stringency (2 x SSC) conditions. Certain genomic fragments were identified and the autoradiograph is shown Fig. 4.3b. The high background is probably due to the low stringency of washing.

The filter was then dehybridised and probed with parasite rRNA end-labelled with polynucleotide kinase (Methods). Characterisation of this rRNA appears in Section 4.6. The stringency of the washing regime was increased (0.5 x SSC), and the same pattern of hybridisation observed (Fig. 4.4a). To see which if any of the genomic fragments visualised were 28S or 18S rRNA specific, the blot was again dehybridised and probed with end-labelled 28S rRNA (Fig. 4.4b).

Comparison of these three autoradiographs reveals firstly that pDm 238/29 and rRNA give rise to the same hybridisation pattern confirming that pDm 238/29 is a good probe for the rRNA genes of P.f.

Fig. 4.2

Autoradiographs of *P.falciparum* polyA$^-$ RNA and genomic DNA hybridised with pDm238/29. pDm238/29 shows homology to both the 28S and 18S rRNAs. Note that the 5.8S rRNA is not shown; for more detailed analysis of polyA$^-$ RNA see Fig. 4.16. When *HindIII* digested genomic DNA is hybridised to pDm238/29 6 fragments are identified (numbered 1 to 6). The strongest homology is to fragments 1 (4.4 kb) and 2 (4.2 kb).

FIG 4.2

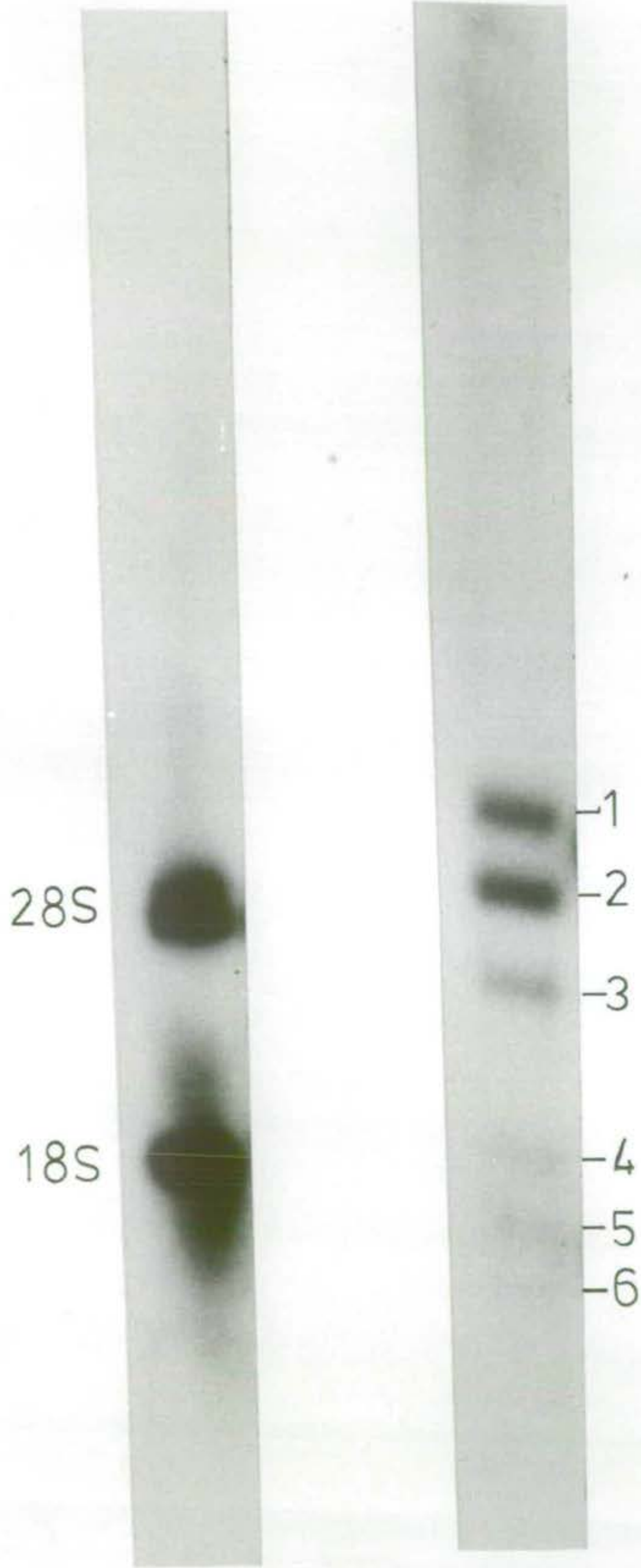


Fig. 4.3

Ethidium bromide stained gel and autoradiograph of K1 DNA. Panel A shows K1 DNA digested with HindIII (track 2), EcoR1 (track 3), SalI (track 4) and HpaI (track 5). Size marker fragments (tracks 1 and 6) are HindIII/EcoR1 fragments of λ^{+} DNA. Panel B is an autoradiograph of above DNA hybridised with pDm238/29. Note highest ★ molecular weight fragment in track 2 results from partial digestion, as it is reproducibly absent in other HindIII digests of K1 (see Fig. 4.2 track 2).

FIG 4.3

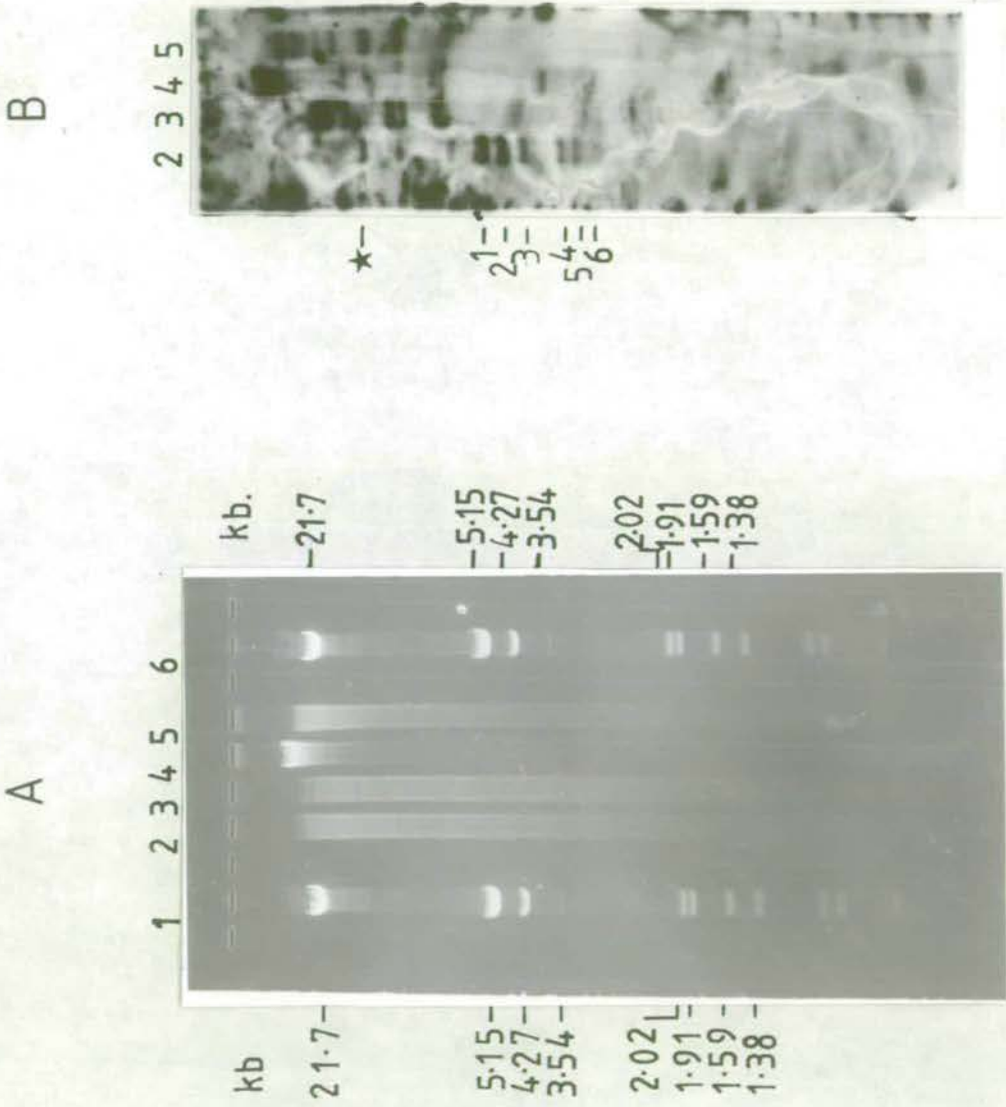
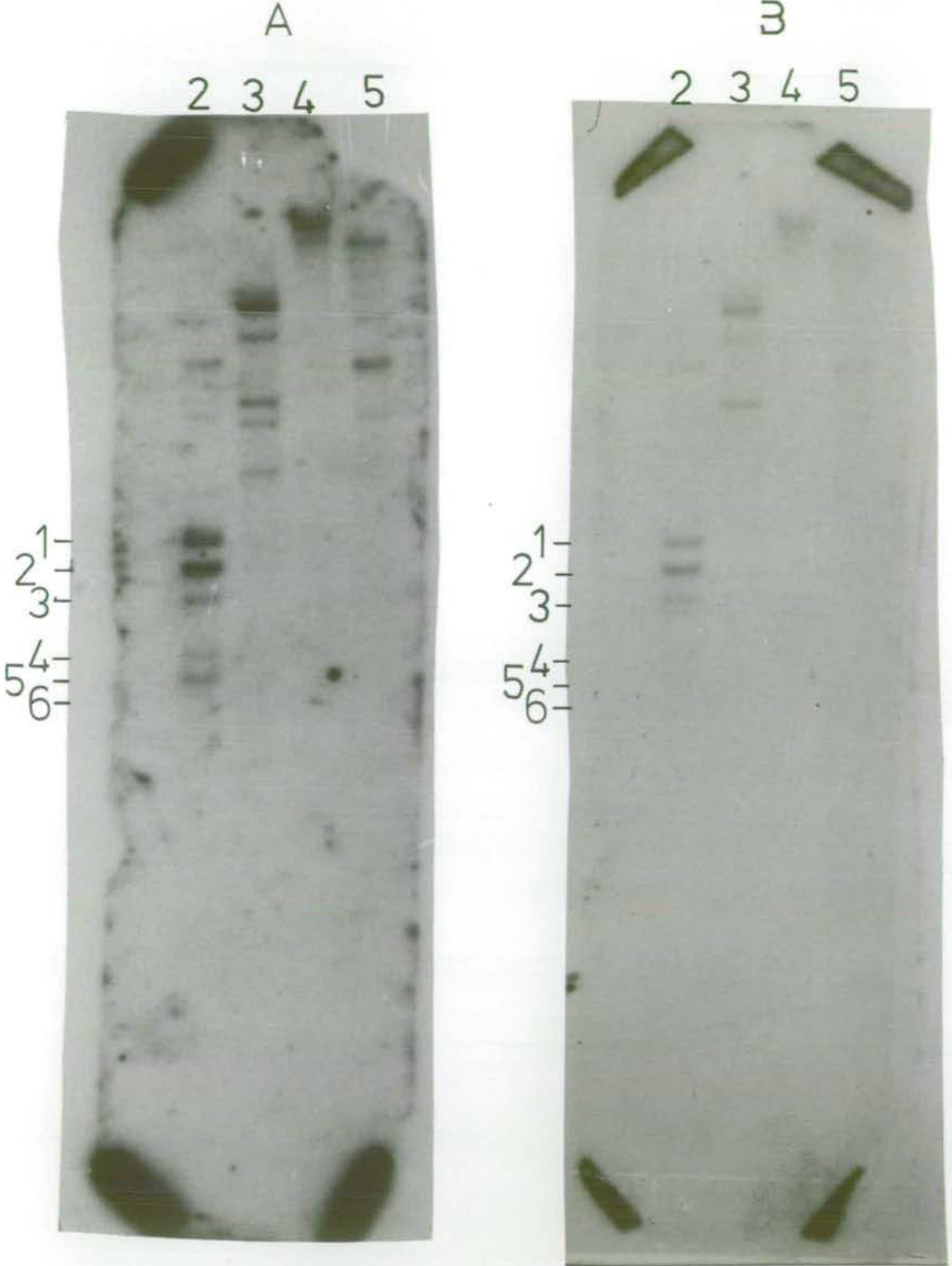


Fig. 4.4

Autoradiographs of gel presented in Fig. 4.3. Southern transfer dehybridised and then probed with total rRNA (Panel A) and then 28S rRNA (panel B).

FIG 4.4



Restriction of genomic DNA with HindIII generates 6 homologous bands. The 7th highest molecular weight fragment observed is the product of a partial restriction digest, as it is reproducibly absent in other HindIII restriction digests (Fig. 4.2, track 2). Secondly, the majority of the rRNA coding sequence is located in the fragments labelled 1 and 2. Thirdly, the DNA homologous to the rRNA (rDNA) contains EcoRl sites; 6 fragments are generated by this enzyme (Fig. 4.3b, track 2). The fragments sizes are approximately 6.2, 7.4, 8.0, 11, 13.5 and 16.0 kb. The strongest homology to pDm 238/29 is shown by the 13.5 and 16.0 kb fragments. Only the 13.5 kb fragments shows strong homology to total rRNA, and this as well as the 8kb fragment are the most obvious when 28S rRNA is used as a probe. The differences between pDm 238/29 and total rRNA may be due to the spacer sequences present in pDm 239/29.

Most of the hybridisation in the SalI digest is to the unrestricted DNA, although some homology is seen to two other bands. Either a small proportion of the rDNA contains SalI sites, or they occur near the ends of any putative rDNA transcription unit. By contrast, in the HpaI digest approximately 7 fragments appear. They are all of reasonably high molecular weight (m.w.) and relative intensity, and show homology to pDm 238/29 and total rRNA. Three of these bands are specific for the 28S rRNA probe.

The salient points from these hybridisations are that (a) pDm 238/29 is a suitable probe to screen the HindIII genomic library constructed in phage λ (Goman *et al.* 1982), as it recognises the same fragments as total rRNA. (b) The library can be expected to contain 6 HindIII fragments homologous to pDm 238/29, the strongest

homology should be to clones containing P.f. rDNA fragments of 4.4 kb (1) and 4.2 kb (2). Some of the fragments when isolated will contain EcoRl and HpaI sites.

4.3 Isolation of rDNA clones from λPFH8 library

The λHPF8 library was constructed from genomic DNA isolated from the Gambian isolate HG13 (Butcher, 1981). The HindIII fragments of parasite DNA were sized on a sucrose gradient, and fragments ranging in size from 2-12 kb were ligated into the HindIII site of NM788 (Goman et al. 1982). The recombinant phage were packaged in vitro and plated to form isolated plaques on indicator bacteria. The pooled plaques constitute the genomic library λPFH8.

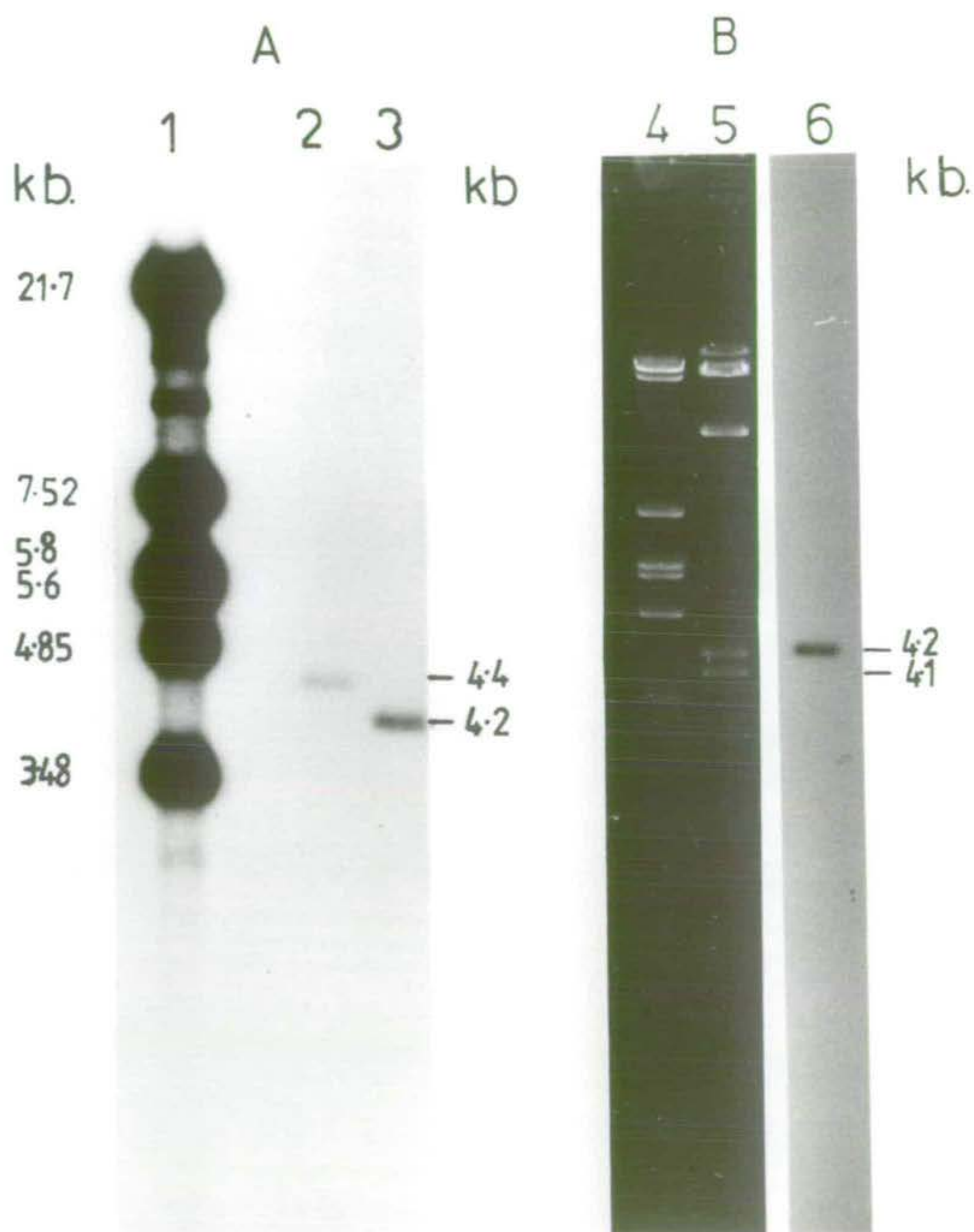
For screening the library was plated out on TGL70 at a dilution yielding approximately 500 plaques per plate. The individual plaques containing rDNA sequence were identified by probing with ^{32}P-labelled pDm 238/29. Of the 25000 plaques screened six positives were obtained, purified and analysed further.

The insert sizes of 5 clones were found to be of two types, 4.4 kb and 4.2 kb. These sizes correspond to the sizes of the 2 major HindIII fragments in the genomic digest (Fig. 4.3). All 5 inserts hybridised to pDm 238/29 and one example of each size is shown in Fig. 4.5a. One clone λHPF8rib2 contained 2 HindIII fragments of 4.2 kb and 4.1 kb (Fig. 4.5b track 5). The 4.1 kb fragment showed no homology to pDm 238/29 (track 6). λHPF8rib2 could be the result of cloning 2 unrelated HindIII fragments, or the cloning of a HindIII partial genomic fragment. For this reason λHPF8rib2 and a clone containing the 4.4 kb fragment λHPF8rib1, were further characterised.

Fig. 4.5

Ethidium bromide stained gel and Southern blot of λPFH8 rDNA clones Inserts of two sizes 4.4 kb (track 2) and 4.2 kb (track 3) show homology to pDm238/29. A third recombinant contains two HindIII fragments of 4.2 kb and 4.1 kb (Track 5), only the 4.2 kb fragment is homologous to pDm238/29 (track 6). Size markers are EcoR1 fragments of λ^{+}DNA, radiolabelled (track 1) and stained with ethidium bromide (track 4).

FIG4.5



4.4 Estimation of A+T rich regions by partial denaturation

The average A+T content (81%) of *P.f.* genomic DNA is unusually high (Bone *et al.* 1983; Goman *et al.* 1982), quite different from that of ribosomal RNA (Hyde *et al.* 1981). These observations would be reconciled if the coding regions (of more average nucleotide distribution) were interspersed with regions very rich in A+T. Perhaps such region(s) could be detected in the rDNA clones. There is a simple way to detect A+T rich sequences of DNA cloned into phage λ . Under appropriate conditions (partial denaturation) such regions form loops visible in the electron microscope (Highton *et al.* 1975).

The two rDNA clones λHPF8*rib1* and λHPF8*rib2* were partially denatured under mild conditions where the arms of phage λ remain entirely native. Only sequences with an A+T content greater than 64% will be denatured (Highton *et al.* 1975; Bone *et al.* 1983). These denaturation loops were measured under the electron microscope, and an example of each is shown (Fig. 4.6). Nineteen λHPF8*rib2* molecules, and 16 λHPF8*rib1* molecules were measured, and their sizes with standard deviations are given in Tables 4.1 and 4.2. The A+T rich sequence (corresponding to the denaturation loop) can be presented graphically: preliminary restriction maps of both inserts were also constructed (data not shown), and the combined data are summarised in Fig. 4.7. Detailed restriction mapping of both the 4.4 kb and 4.2 kb fragments appears in Section 4.5.

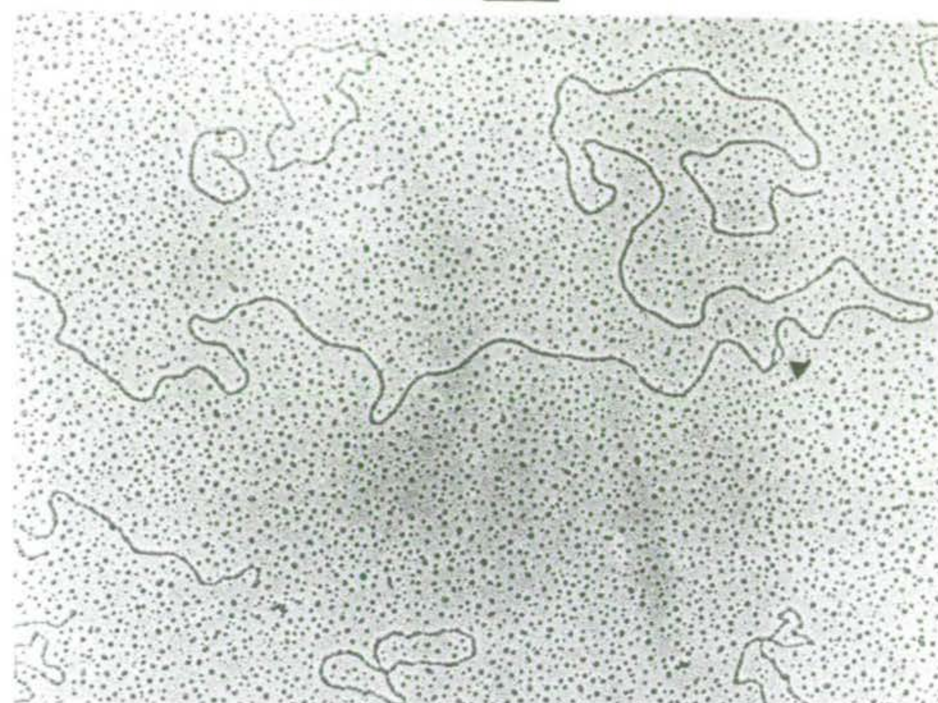
Note that λPFH8*rib1* and the 4.2 kb fragments from λPFH8*rib2* contain only a small (0.5 kb) stretch of high A+T content sequence. Evidence will be presented elsewhere (Section 4.6) to show that this A+T rich sequence occurs between the 18S and 28S genes, and corresponds

Fig. 4.6

Electron micrographs of partially denatured molecules of PFH8 *ribl* and PFH8 *rib2*. The molecules were denatured under mild conditions where only sequence with an A+T content greater than 64% is denatured (Highton, *et al*. 1975, Bone *et al*. 1983). The molecules were spread with both single strand DNA (phage M13) and double strand DNA (PAT 153) as size markers (Denhardt, *et al.* 1978; Twigg and Sherratt, 1980). Nineteen HPF8 *rib2* and sixteen HPF8 *ribl* molecules were measured and the sizes of the denaturation loops are given in Tables 4.1 and 4.2.

FIG4.6

λPFH8 <u>rib1</u>



λPFH8 <u>rib2</u>

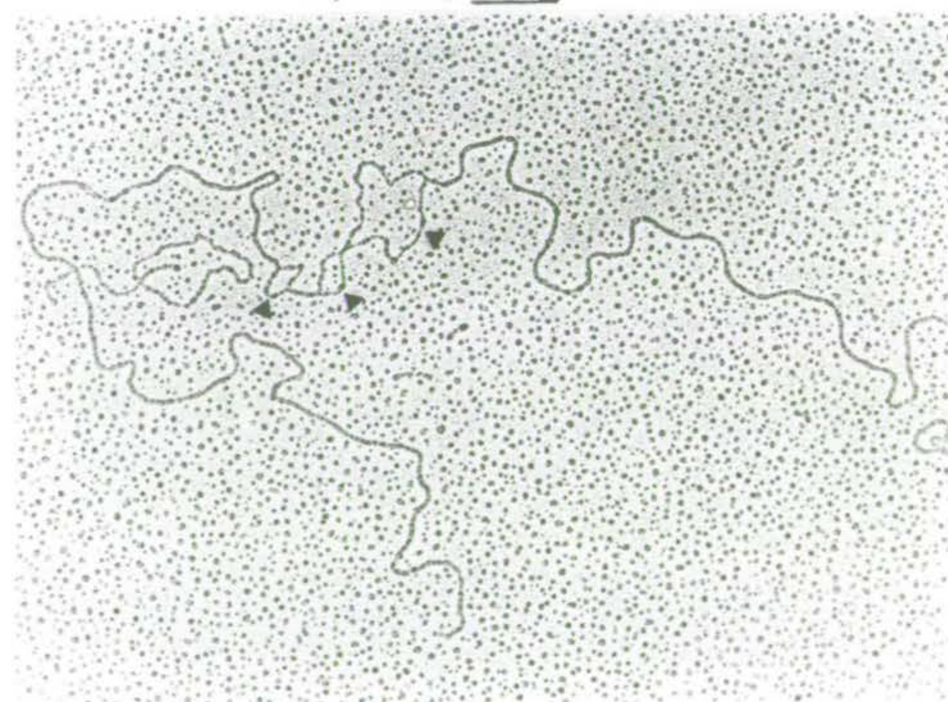


Table 4.1 Partial Denaturation Mapping of λHPF8rib2

	Left arm of λ in kb	Denaturation loop kb	Double-stranded region kb	Denaturation loop kb	Double-stranded region kb	Denaturation loop kb	Right arm of kb	Time denatured in secs.
	24.8	2.8	0.4	1.2	4.0	0.4	13.0	
	26.1	3.3	0.3	1.2	3.7	0.5	12.6	
	26.4	3.0	0.8	1.2	4.0	0.4	14.0	90
	23.4	2.8	0.3	1.0	3.6	0.5	12.5	
	27.1	2.2	0.8	1.5	4.1	0.5	14.5	
	24.0	4.5			3.6	0.3	13.0	
	25.0	3.6			4.0	0.4	13.3	
	23.5	4.0			3.8	0.4	12.7	
	25.6	1.9	1.2	1.0	4.0	0.4	13.4	
	25.3	1.1	2.3	0.4	3.0	0.3	13.4	
	24.3	1.8	1.4	1.0	3.9	0.3	13.4	
	25.1	1.1	2.5	0.8	4.4	0.3	13.8	
	27.0	1.6	1.2	0.8	4.5	0.3	14.2	20
	22.5	2.5	0.4	0.5	3.6	0.3	12.8	
	24.1	1.3	0.7	0.8	2.6	0.2	12.6	
	22.9	1.3	1.2	1.2	4.3	0.3	11.1	
	24.0	1.7	1.2	1.0	4.2	0.4	13.0	
	26.1	1.9	0.8	1.7	3.6	0.4	13.6	
	25.2	2.0	1.0	1.3	3.6	0.4	13.0	
mean	24.8	2.3	1.0	1.0	3.8	0.4	13.2	
S.D.	1.3	1.0	0.6	0.3	0.4	0.1	0.7	

Table 4.1 (continued)

Note: The high standard deviations for the denaturation loops in the insert proximal to the left arm of λ, are due to the sizes of the loops varying considerably with the time of denaturation. In three molecules the small ds region has denatured completely, and is a consequence of it occurring between large stretches of A+T rich sequence.

Table 4.2 Partial Denaturation mapping of λHPF8*ribl*

	Left arm of λ and double-stranded region kb	Denaturation loop kb	Right arm of λ and double-stranded region kb	Time of denaturation in secs.
	24.9	0.3	17.1	
	23.2	0.4	15.6	
	22.9	0.4	15.1	20
	27.1	0.8	17.1	
	24.3	0.8	15.6	
	22.1	0.5	15.0	
	24.1	0.5	17.5	
	22.6	0.8	15.0	
	23.7	0.5	14.5	
	24.1	0.9	16.6	
	25.2	0.6	17.1	90
	24.5	0.9	16.13	
	25.5	0.5	17.8	
	24.8	9.4	15.3	
	23.2	0.4	14.8	
	24.0	0.3	14.4	
mean	24.1	0.6	15.9	
s.D.	1.2	0.2	1.1	

Note: Left arm of λNM788 is 22.6 kb. The denaturation loop therefore occurs 1.5 kb in from the *Hin*dIII site at the end of the left arm.

Fig. 4.7

Graphical representation of Denaturation-loop analysis. Shown graphically are the positions of the denatured regions measured under the electron microscope. Also indicated are the positions of some restriction sites.

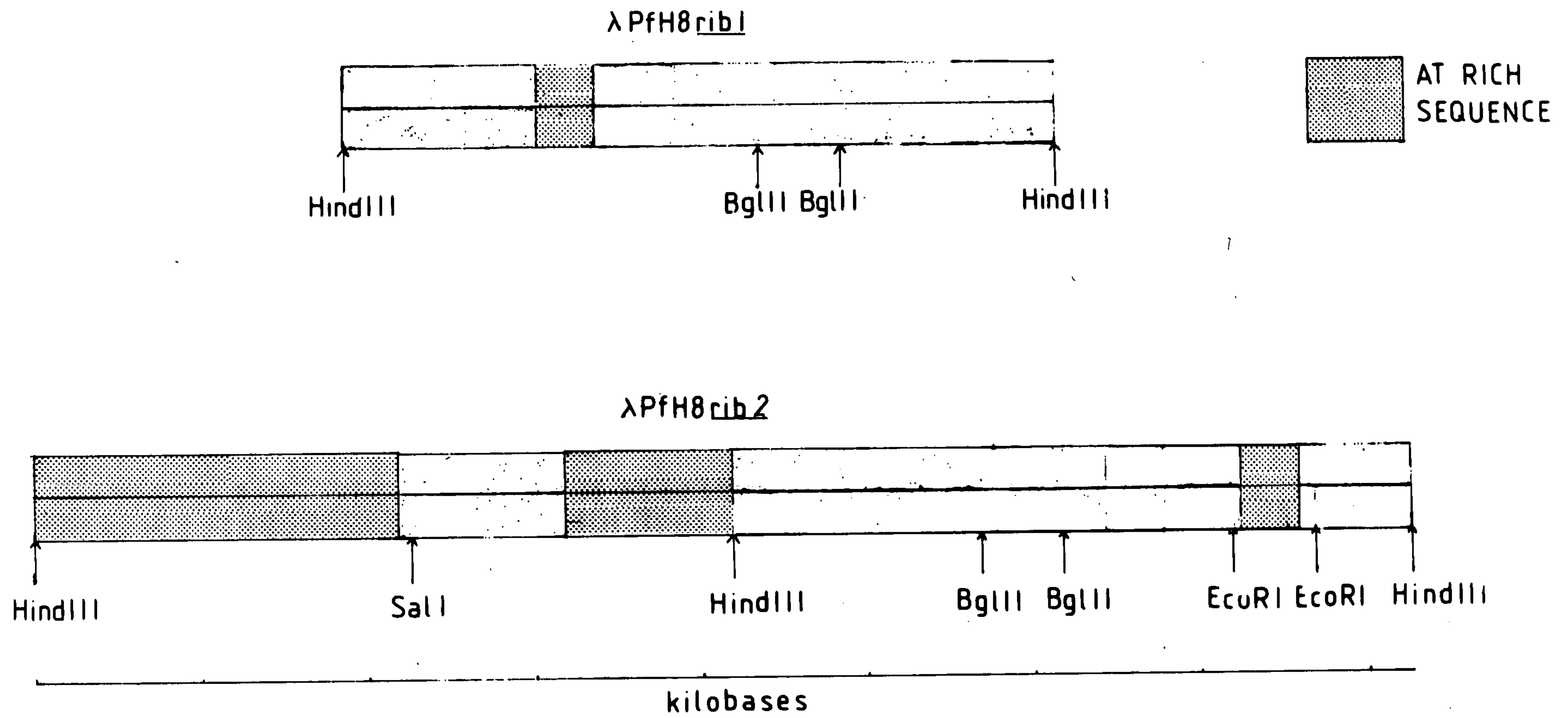


FIG 4.7

to the internal transcribed spacer. Note also that the 4.1 kb fragment present in λPFH8*rib2*, which shows no homology to the rDNA probe, contains two A+T rich segments of approximately 2.3 kb and 1.0 kb.

4.5 Physical characterisation of the 4.4 kb and 4.2 kb rDNA fragments

The λPFH8 library was constructed from the West African (Gambian) strain of *P.f.* HG13 (See Section 4.3). After its construction the HG13 strain became unavailable in this laboratory, due to its dying in culture and the failure to revive frozen stocks. As a consequence a Thai ioslate K1 (Thaithong *et al.* 1981) became the standard laboratory strain. This strain was originally chosen because it is highly resistant to the anti-malarial drug pyrimethamine (Thaithong *et al.* 1981). The genomic digests were therefore performed with DNA from the Thai isolate, and the rDNA clones isolated from the λ PFH8 library are from the Gambian isolate.

The observation that the rDNA fragments in λPFH8*rib1* and λPFH8*rib2* have the same size as fragments in the *Hind*III digest of total *P.f.* DNA, suggest strongly that these fragments are common to both isolates and have been cloned unaltered into the phage vector. Confirmation for this conclusion was obtained by restriction mapping and comparison of the cloned and uncloned fragments.

Firstly, *Bgl*II sites were mapped in λPFH8*rib1* and λPFH8*rib2* (see Fig. 4.7). The genomic DNA was then restricted with *Hind*III and *Bgl*II. This DNA was transferred to nitrocellulose and hybridised with λPFH8*rib1* and λPFH8*rib2*. The fragments homologous to the rDNA clones are listed in Table 4.3 Note that fragments the same size as those present in the rDNA clones are identified in the genomic digests.

Table 4.3 Identification of *HindIII/BglII* genomic fragments homologous to λPFH8*rib1* and λPFH8*rib2*

Sizes in kb. of *HindIII/BglII* restriction fragments

Genomic DNA	λPFH8*rib1*	Genomic DNA	λPFH8*rib2*
		3.9	
2.5	2.5	2.5	
2.2		2.2	2.2
1.9		1.9	
1.4	1.4	1.5	1.5
1.0		1.4	
0.6		0.6	0.55
0.5	0.5	0.5	

Note: λPFH8*rib1* and λPFH8*rib2* cross hybridise to essentially the same set of genomic fragments. Some of the genomic fragments identified have the same size as those cloned in λPFH8*rib1* and λPFH8*rib2*.

Secondly, to facilitate the analysis the HindIII genomic fragments present in λPFH8rib1 and λPFH8rib2 were subcloned into the HindIII site of pBR322 (Bolivar et al. 1977). Restriction digests of the 3 recombinants pPFrib1 (4.4 kb), pPFrib2 (4.2 kb) and pPF4.1 (4.1 kb) are shown in Fig. 4.8. The 4.4 kb fragment from PFH8rib1 runs as a doublet with pBR322 (track 2). The 0.5 kb BglII fragment is difficult to see on this particular gel, the 2.5 and 1.4 kb fragments are clearly visible (track 3). The SalI site in the 4.1 kb fragment (pPF4.1) is obvious in track 5. The presence of BglII and EcoRl sites in pPFrib2 is shown by tracks 7 and 8.

The following observation is particularly significant. Both pPRrib2 and pPF4.1 show good homology to a 7 kb SalI/BglII fragment strongly suggesting that they are related. Note also pPF4.1 shows some cross homology to a 2.4 kb fragment, readily identified by pPFrib2 (Fig. 4.9 tracks 1 and 2). These observations can be explained by supposing that the SalI site present in pPF4.1, which is derived from the Gambian isolate, is absent in the Thai isolate. This results in a 7 kb fragment rather than a 3.5 kb fragment (see Fig. 4.7) being identified. Polymorphism for a restriction site in non-coding DNA is a well known phenomena in rDNA (Petes et al. 1977, Cory et al. 1977), and could be a reasonable explanation here considering the difference in the two strains.

Further evidence that the 2 fragments are related is shown in Fig. 4.10. Note that pPF4.1 and pPFrib2 hybridise to the same EcoRl HpaI and BglII fragments. Taken together these data indicate that present in λPFH8rib2 is a genomic fragment generated by partial digestion with HindIII. Evidence will be presented later (Section 4.7)

Fig. 4.8

Ethidium bromide stained gel of restriction digests of pPF*ribl*, pPF*rib2* and pPF4.1

Track

2 *Hind*III digests of pPF*ribl*, the 4.4 kb insert runs as a doublet with the 4.36 kb pBR322.

3 *Hind*III/BglII digest of pPF*ribl* generates insert fragments of 2.5, 1.3, 0.5 kb.

4 *Hind*III digests of pPF4.1 showing the 4.1 kb insert.

5 *Hind*III/*Sal*I digest of pPF4.1 restricts insert to give 2.1 and 2.0 kb fragments.

6 *Hind*III digest of pPF*rib2* showing 4.2 kb insert.

7 *Hind*III/*Bgl*II digests of pF*rib2* restricts insert to give 2.2, 1.4 and 0.55 kb fragments.

8 *Hind*III/*EcoRl* digests of pPF*rib2* generates insert fragments of 3.0, 0.7 and 0.5 kb

9 *Hind*III fragments of λ^{+} DNA as size markers.

Note: 2 partials can be observed in tracks 3 and 7. Also the definition of the low molecular weight fragments has been lost on photography. They can be seen on original negative.

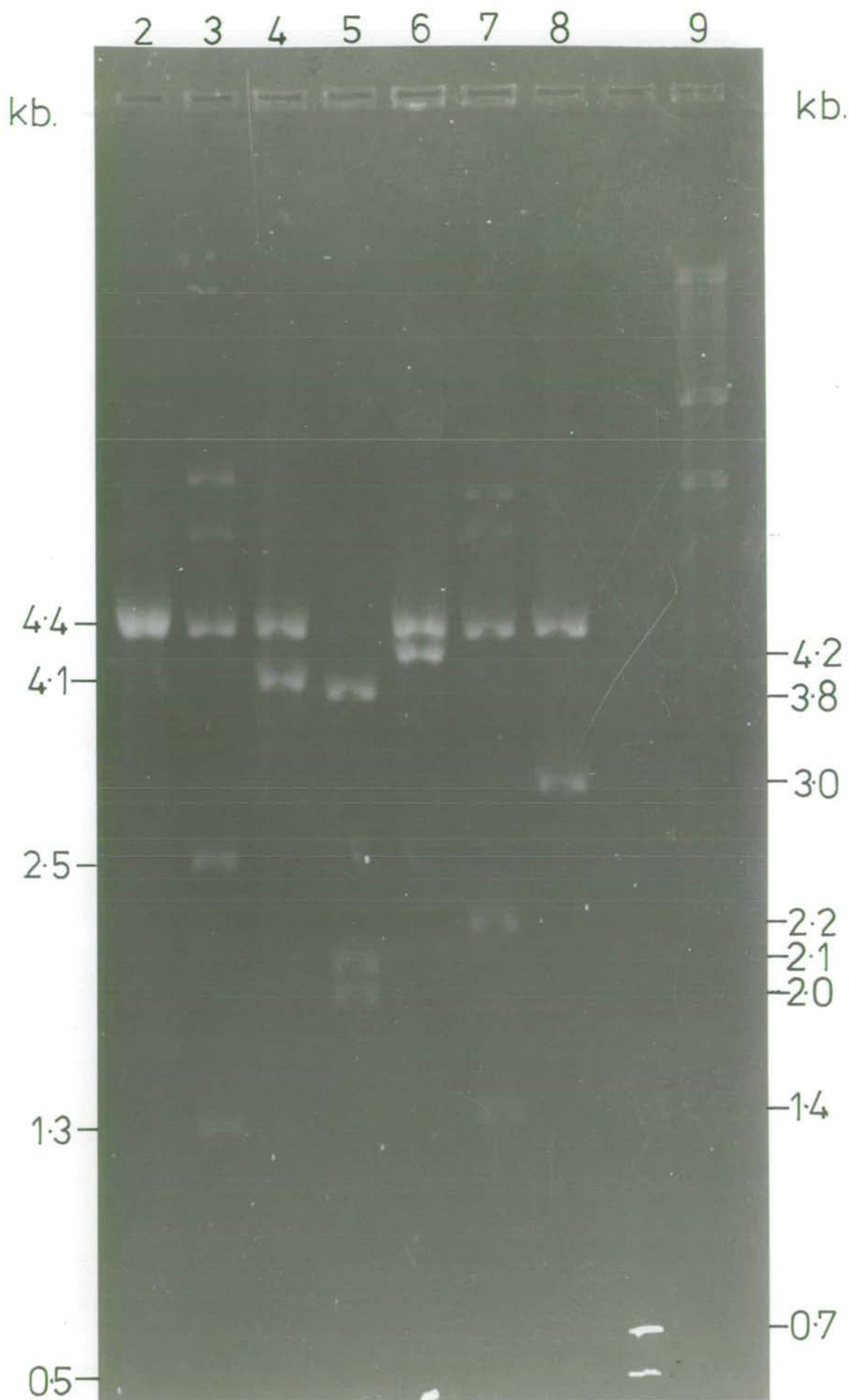


FIG 4.8

Fig. 4.9

Autoradiographs of K1 DNA.

K1 DNA was digested with SalI/BglII, transferred to nitrocellulose and hybridised with pPFrib2 (track 1), and pPF4.1 (track 2). Both show homology to a 7.0 and a 2.4 kb SalI/BglII fragment, indicating that the sequences present in the two probes are related. The extra fragments identified in track 1 are probably due to cross homology to other rDNA sequences.

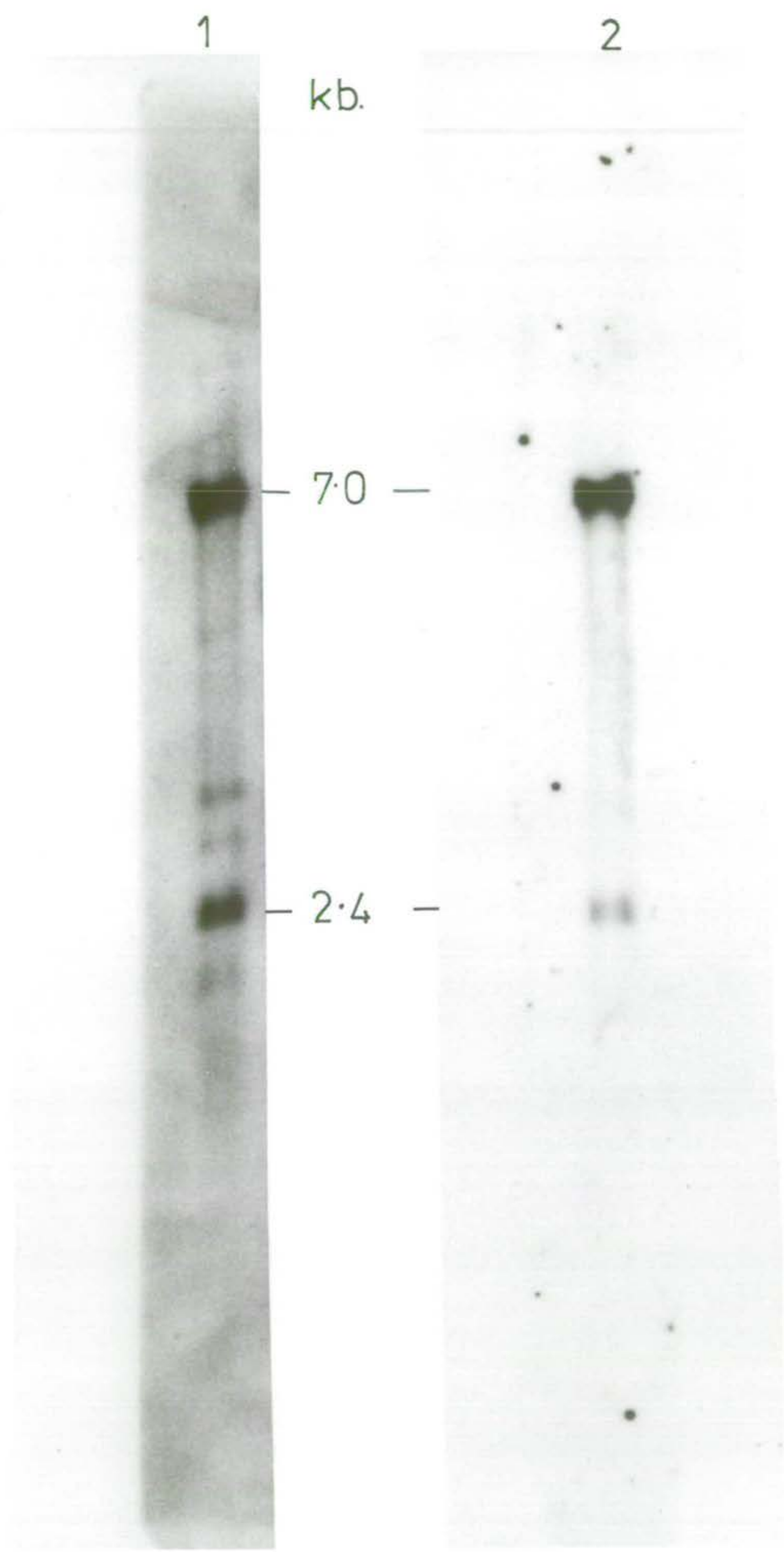


FIG 4.9

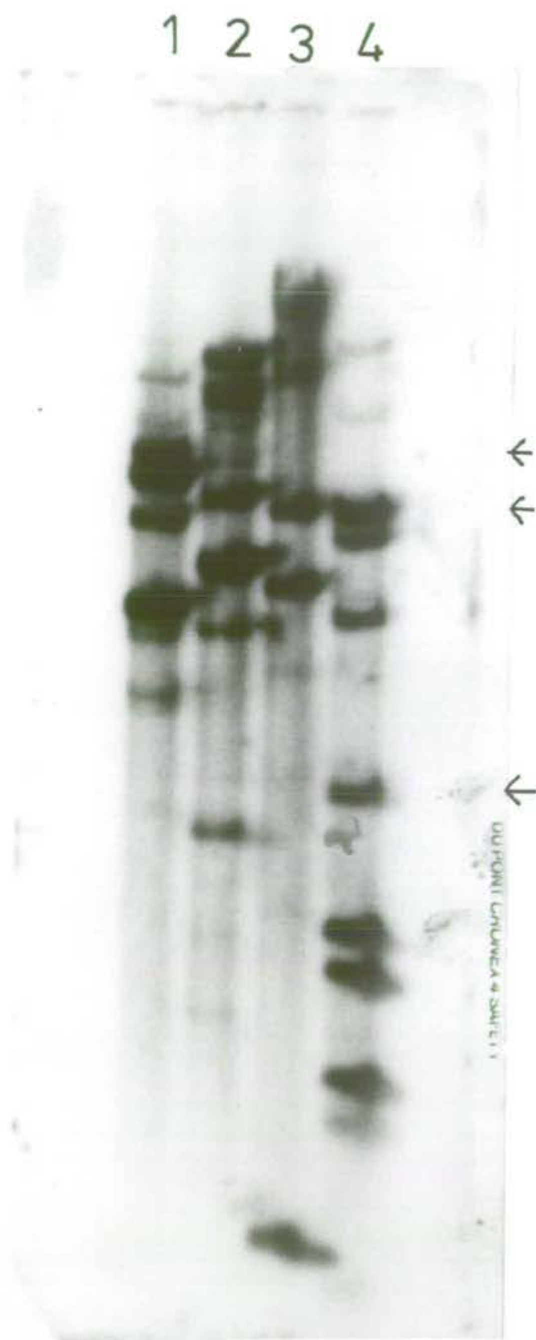
Fig. 4.10

Autoradiographs of K1 genomic DNA.

The DNA was digested with EcoR1 (track 1), HpaI (track 2), AvaI (track 3) and Bgl/II (track 4). After transfer to nitrocellulose the above digested DNA was hybridised to pPFrib2 (Panel A), and pPF4.1 (Panel b). pPFrib2 hybridises to a number of fragments due to cross homology with rDNA sequences. pPF4.1 hybridises to fewer fragments. In 3 of the digests both pPF4.1 and pPFrib2 hybridise to a common fragment (arrowed). The poor homology in track 3 Panel B is most likely due to poor transfer of high molecular weight DNA.

FIG 4.10

A



B

indicating that pPF4.1 may represent an intervening sequence (IVS) in the 28S rRNA gene.

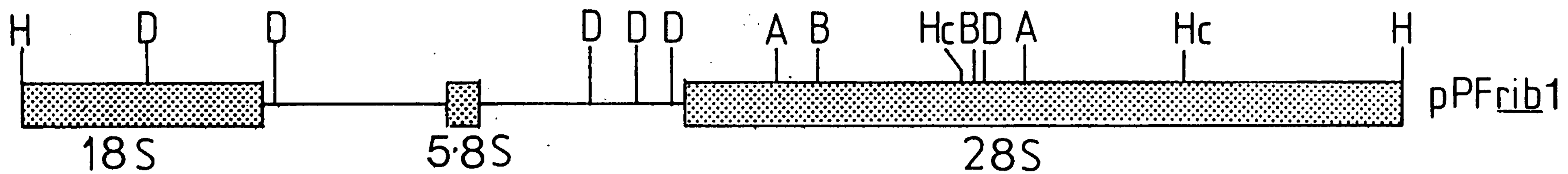
Single or multiple enzyme digestions of pPF*rib1* and pPF*rib2* gave the restriction maps Fig. 4.11. Examples of 2 gels used in this analysis are shown in Fig. 4.12 and 4.13. Although several restriction enzymes were tried, there is a general paucity of sites in the 2.4 kb *HindIII*/*AvaI* fragment of pPF*rib1*. Since the genomic DNA is known to have an A+T content of 81% (Pollock *et al.* 1982; Goman *et al.* 1982), the restriction endonuclease *DraI* (isozyme *AhaIII*) with the recognition sequence (TTT▼AAA) was tried, and found to have 6 sites in pPF*rib1* Fig. 4.14. As the *DraI* sites in pBR322 are 3233, 3252 and 3944 (Sutcliffe 1978), 2 of the 6 sites in pPF*rib1* could be mapped. The remaining 4 sites in pPF*rib1* were ordered using the method of Smith and Birnstiel (1976) (Methods, Section 2.29). pPF*rib1* was partially digested with *DraI*, and the products fractionated by agarose gel electrophoresis. Only those fragments with the common end labelled terminus are detected by autoradiography (Fig. 4.15).

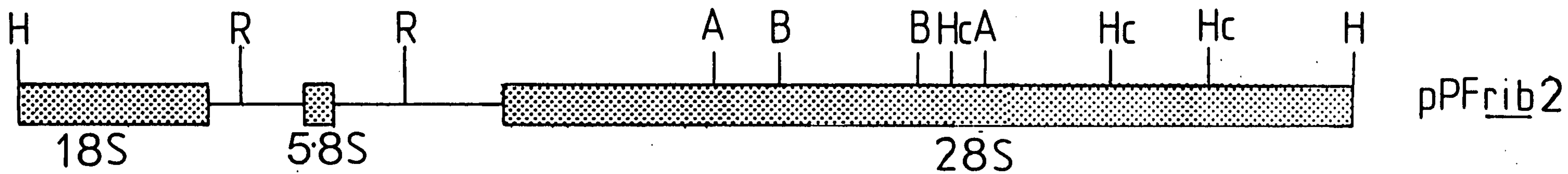
4.6 Mapping of rRNA coding sequences in pPF*rib1* and pPF*rib2*

Hot phenol extraction of total RNA from *P.f.*, and its subsequent analysis reveal it to be comprised of 3 components, *ca.* 81% is rRNA (this figure includes a minor fraction of polyA$^-$ mRNA), *ca.* 14% is tRNA and *ca.* 5% is polyA$^+$ RNA (Hyde *et al.* 1983). The polyA$^+$ RNA can be separated from the polyA$^-$ by oligo-dT-cellulose chromatography (Nakazato *et al.* 1974). The polyA$^-$ RNA was fractionated by agarose/formaldehyde gel electrophoresis (Methods, Section 2.18) into its major constituents 28S, 18S, 5.8S and tRNA (Fig. 4.16A). Examination of the rRNA under the electron microscope revealed it to be contaminated

Fig. 4.11

Maps of pPFrib1 and pPFrib2. Maps show positions of restriction sites *HindIII*, (H); *AvaI*, (A); *BglII*, (B); *DraI*, (D); *EcoRI* (R) and *HincII* (Hc). Shown also (hatched) are the positions of the rRNA genes, the limits of which have not been mapped precisely.





scale 1kb.

FIG 4.11

Fig. 4.12

Ethidium bromide stained gel of restriction digest of pPFrib2, pPFrib1 and pPF4.1

Track	
1	HindIII fragments of λ^{+} DNA as size markers
2	BglII/AvaI digest of pPFrib2 5.1, 2.8, 0.55 kb
3	BglII / Pstl digest of PPFrib2 5.2, 3.0, 0.55 kb
4	SalI/AvaI digest of pPF4.1 5.0, 2.75, 0.75 kb
5	SalI/Pstl digest of pPF4.1, 3.0, 2.75 doublet kb
6	BglII/AvaI digest of pPFribl 5.4, 2.8, 0.5 kb
7	BglII/PstI digest of pPFribl 5.0, 3.3, 0.5 kb
8	HindIII fragment of λ^{+} DNA as size markers

The definition of the small fragments has been lost on photography, though they are clear on original negative. Since there is no BglII site, and the AvaI, SalI and PstI sites are known in pBR322 (Sutcliff, 1978). The AvaI, SalI and BglII sites can be mapped in pPFribl, pPFrib2 and pPF4.1. For maps of pPFribl and pPFrib2, see Fig. 4.11

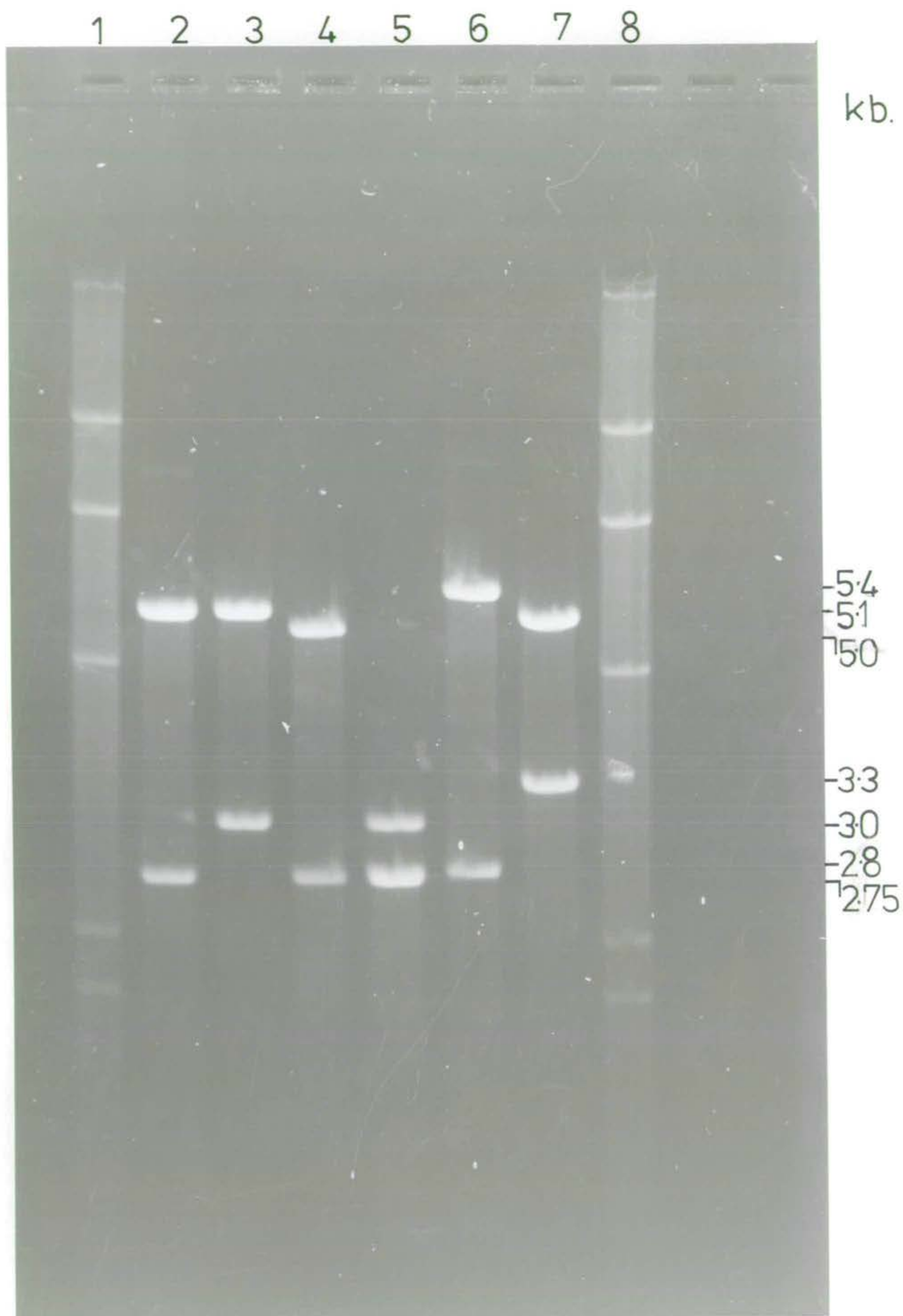


FIG 4.12

Fig. 4.13

Ethidium bromide stained gel of restriction digests of pPFrib1 and pPFrib2.

Track

1 HindIII fragments of λ^+ DNA as size markers

2 HindIII/HincII digests of pPFrib1 giving 3.25*, 3.1, 0.85, 0.7, 0.65, 0.45* kb fragments.

3 HincII/BglII digest of pPFrib1 giving 3.25*, 2.9, 1.3, 0.8, 0.4 kb fragments

4 HindIII/HincII/BglII digest of pPHrib1 giving 3.25*, 2.5, 0.7 doublet, 0.65*, 0.45, 0.4 kb fragments

5 HindIII/HincII digest of pPFrib2 giving 3.25*, 2.8, 0.65*, 0.5 0.45* doublet, 0.36 kb fragments.

6 HincII/BglII digest of pPFrib2 giving 3.25*, 2.65, 1.1, 0.55, 0.5, 0.3 kb fragments

7 HindIII/HincII/BglII digest of pPFrib2 giving 3.25*, 2.2, 0.65* 0.55, 0.5, 0.45*, 0.3 kb fragments

Since the HincII sites are known in pBR322 (Sutcliff, 1978) and the BglII sites in pPFrib1 and pPFrib2 have already been mapped, the HincII sites in the inserts can be positioned. The fragments solely derived from pBR322 are shown *.

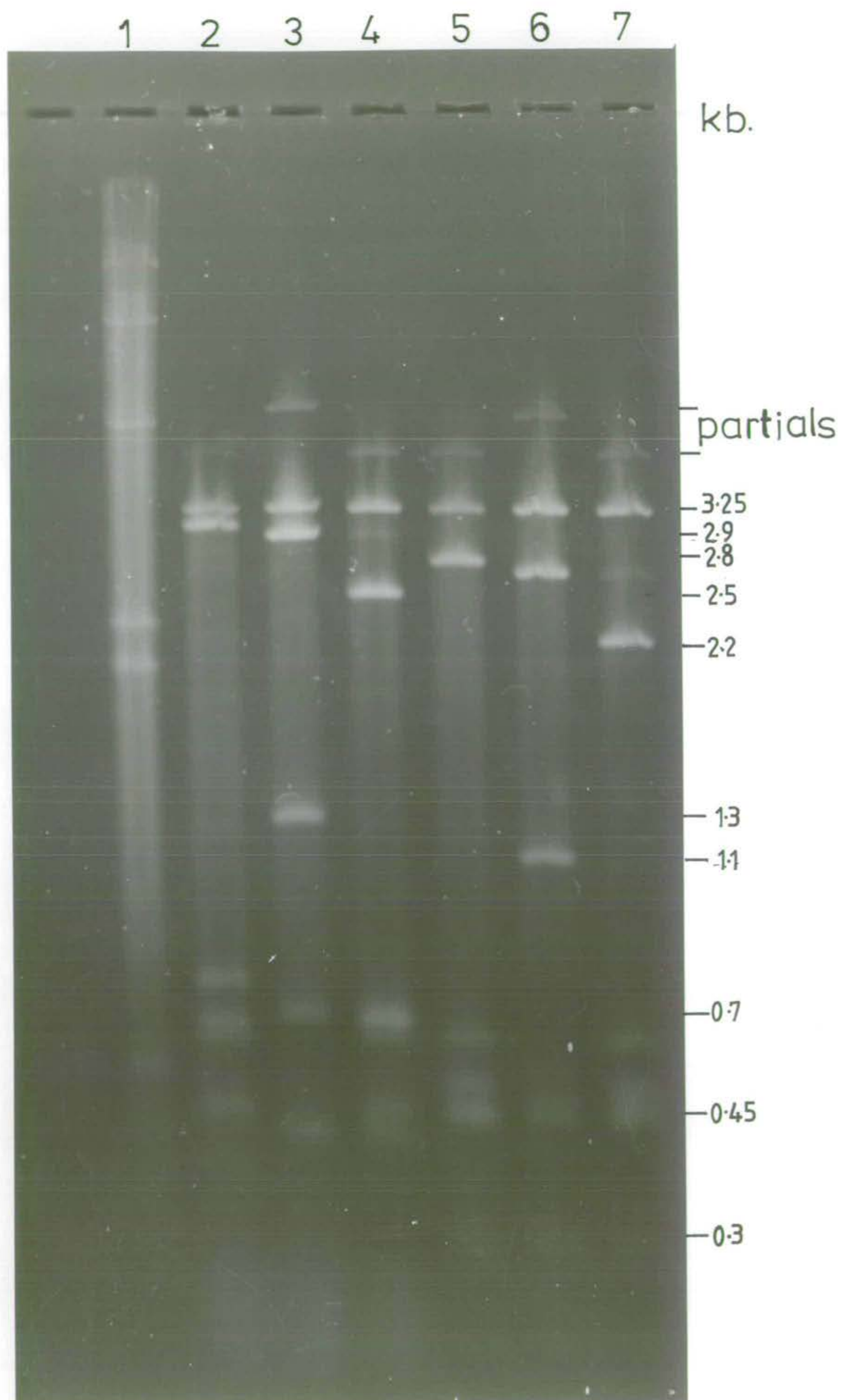


FIG 4.13

Fig. 4.14

Ethidium bromide stained gel of pPFrib1 digested with DraI (track 2). The size markers are HindIII/EcoR1 fragments of λ^{+} DNA (track 1). A DraI digest of pPFrib2 is included for comparison, but the restriction sites were not mapped. A DraI (D) digest of pPFrib1 yields fragments of 4.6*, 1.0 doublet, 0.84*, 0.7, 0.4, 0.15, and 0.1 kb. Since the DraI (AhaIII) sites in pBR322 are known (Sutcliff, 1978), two (*) of the six DraI sites in pPFrib1 could be positioned (see bottom), the others were ordered by the method of Smith and Birnsteil (see Fig. 4.15).

FIG 4.14

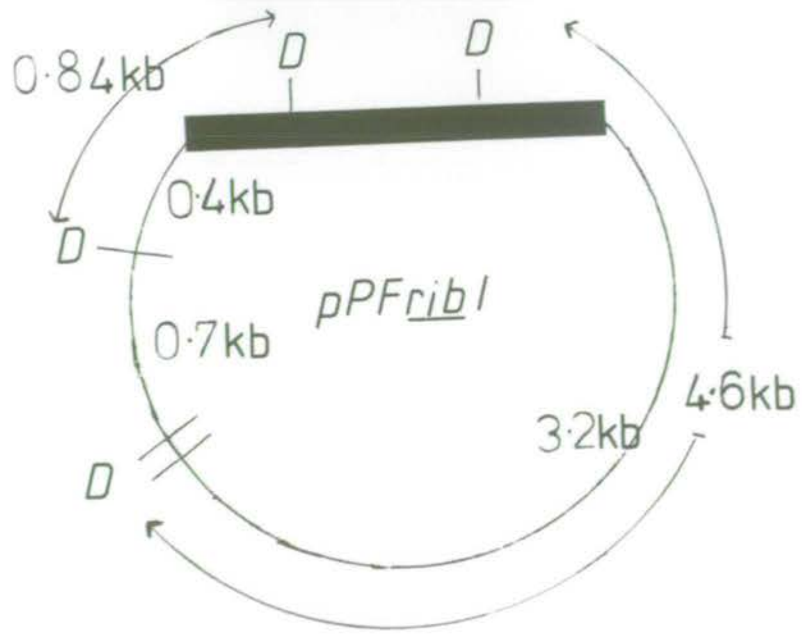
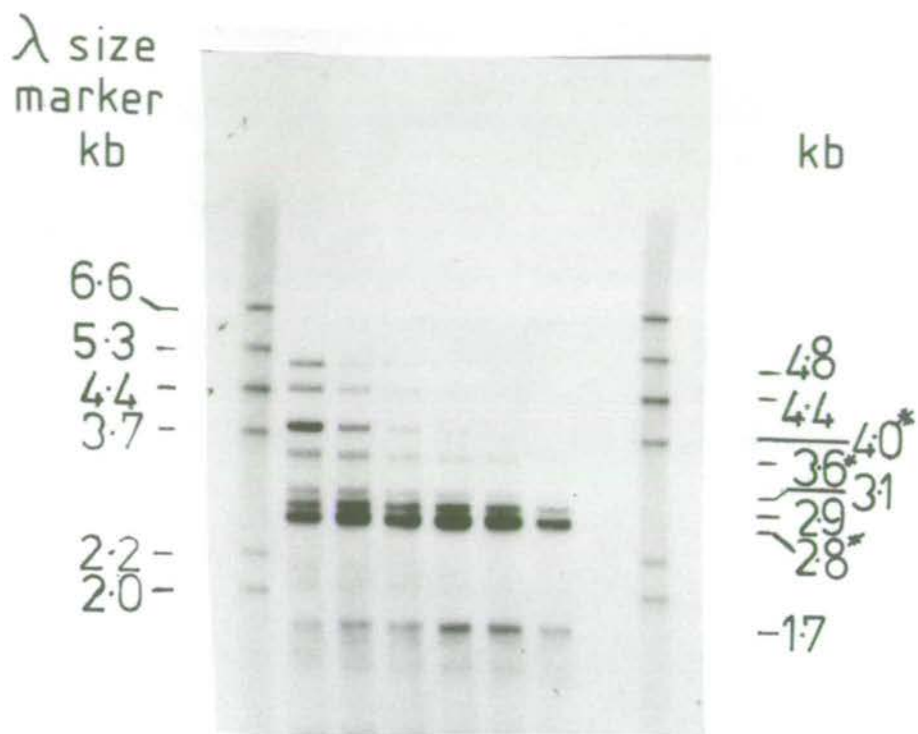


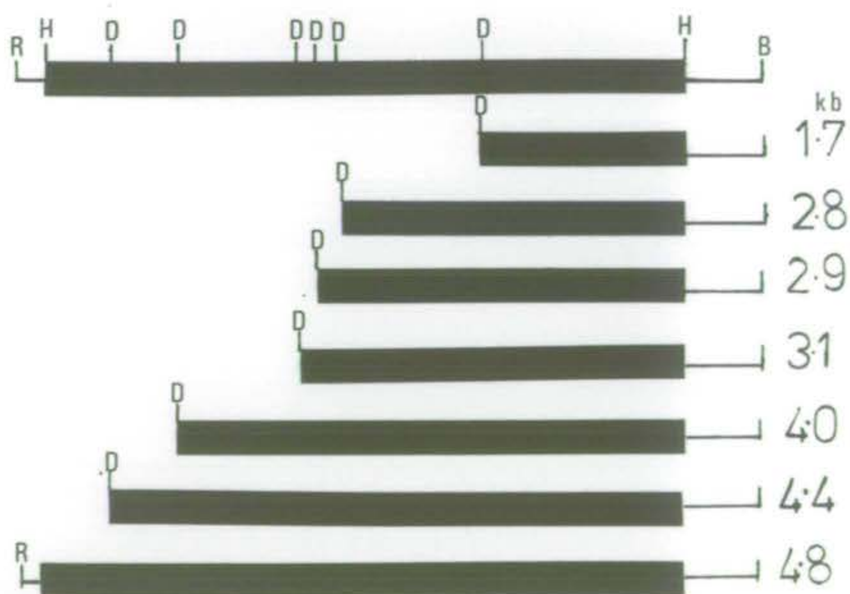
Fig. 4.15

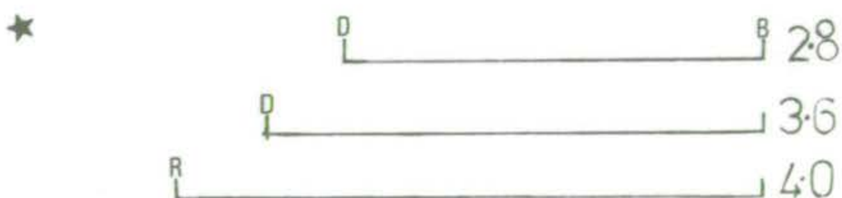
Smith and Birnsteil mapping of *DraI* sites in pPF*ribl*: the recombinant plasmid was first digested with *BamHl*(B), end-labelled, and then digested with *EcoRl* (R). *DraI* (D) was added to give partial digestion, and aliquots were removed over a time course. In this way the *DraI* sites were ordered, and are represented graphically below. The size markers are *Hind*III/*Sal*I fragments of λ^{+} DNA.

*Note the partial at 3.6 kb, and the doublets at 2.8 and 4.0 kb are derived from pBR322.

FIG 4.15







by a small amount of DNA (data not shown). This contaminating DNA was removed by passing the RNA through a 5.7M CsCl cushion (Glisin et al. 1979), the tRNA was also removed by this centrifugation step (Fig. 4.16B). Examination of Fig. 4.16B shows that the rRNA was largely intact with only a small amount of degradation.

The individual rRNA species were removed from the gel, end-labelled with polynucleotide kinase, and hybridised against restricted pPFribl and pPFrib2 (Methods, 2.25). The pattern of hybridisation of 28S rRNA to HindIII/AvaI digested pPFribl and pPFrib2 is shown in Fig. 4.17, tracks 2 and 3. Note that the 28S sequence is confined to the 1.2 kb and 0.86 kb fragments of pPFribl (track 4), whereas all 3 fragments of pPFrib2 show homology (track 5). An EcoRl digest of pPFrib2 (track 6) defines the limit of the 28S rRNA sequence in PFrib2, as neither the 0.7 kb or 0.5 kb fragments show homology (even after long exposure) track 7.

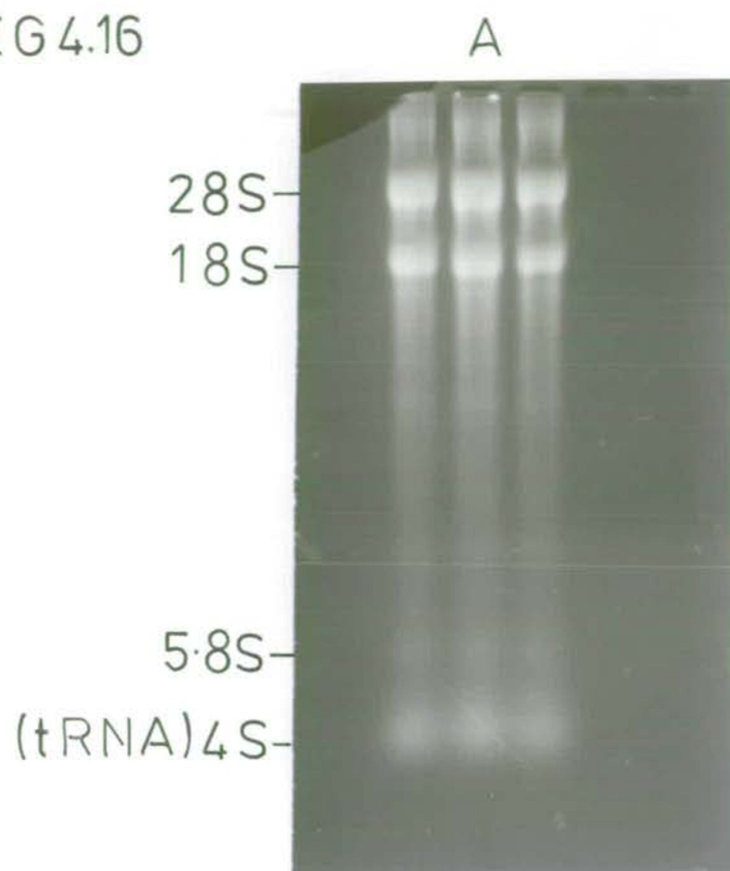
When the process was repeated using ^{32}P- 18S rRNA the 0.7 kb fragment from pPFrib2 and the 2.4 kb fragment from pPFribl are identified (Fig. 4.18). Some hybridisation is seen to other fragments, because the 18S rRNA probe was contaminated with degraded 28S rRNA. Finally 5.8S rRNA shows homology to the 2.4 kb fragment of pPFribl and the 0.5 kb EcoRl fragment of pPFrib2 (Fig. 4.19).

To identify the regions encoding the 28S, 18S and 5.8S rRNA within pPFribl, the plasmid was digested with HindIII, AvaI, and DraI, and the hybridisations repeated (Fig. 4.20). The fragments homologous to the 18S rRNA were identified (track 3), as they showed no homology to 28S rRNA (track 4). The DraI fragment homologous to 5.8S rRNA was identified by its homology to a fragment appearing after digestion with DraI alone (data not shown). In this way the coding

Fig. 4.16

Ethidium bromide stained agarose/formaldehyde polyA$^-$RNA gel. Fractionation of RNA reveals it to be comprised of 4 major class sizes which operationally are termed 28S, 18S, 5.8S and 4S (tRNA). Shown in Panel A are 3 samples of parasite polyA$^-$RNA. The RNA can be purified by a CsCl centrifugation step, this also removes the 4S RNA. Panel B show 2 samples of RNA treated in this way. Note there is some degradation of the RNA.

FIG 4.16



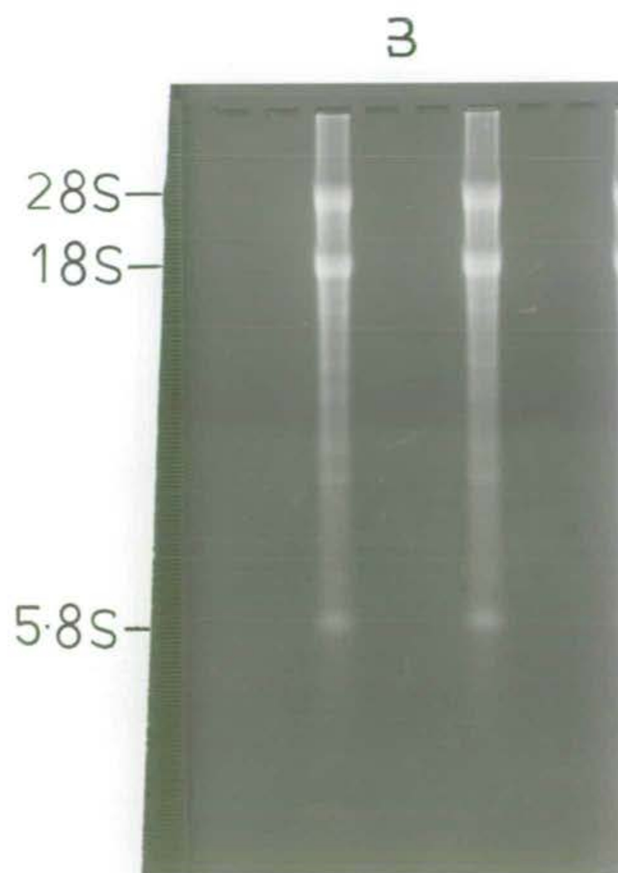


Fig. 4.17

Ethidium bromide stained gels and autoradiographs of pPF*rib1* and pPF*rib2* hybridised with 28S rRNA. *HindIII/AvaI* digests of pPF*rib1* (track 2), and pPF*rib2* (track 3) were transferred to nitrocellulose and hybridised with 28S rRNA. Homology is observed to the 1.2 kb and 0.86 kb fragments of pPF*rib1* (track 4), and to 3 fragments of pPF*rib2* (track 5). An *EcoR1* digest of pPF*rib2* (track 6) shows that the 28S rRNA homology is confined to a single 7.4 kb fragment (track 7).

Note: the pBR322 derived fragments are shown *;
the fragments sizes are given in kb;
size markers are *HindIII* fragments of λ^+ DNA (track 1)

FIG 4.17

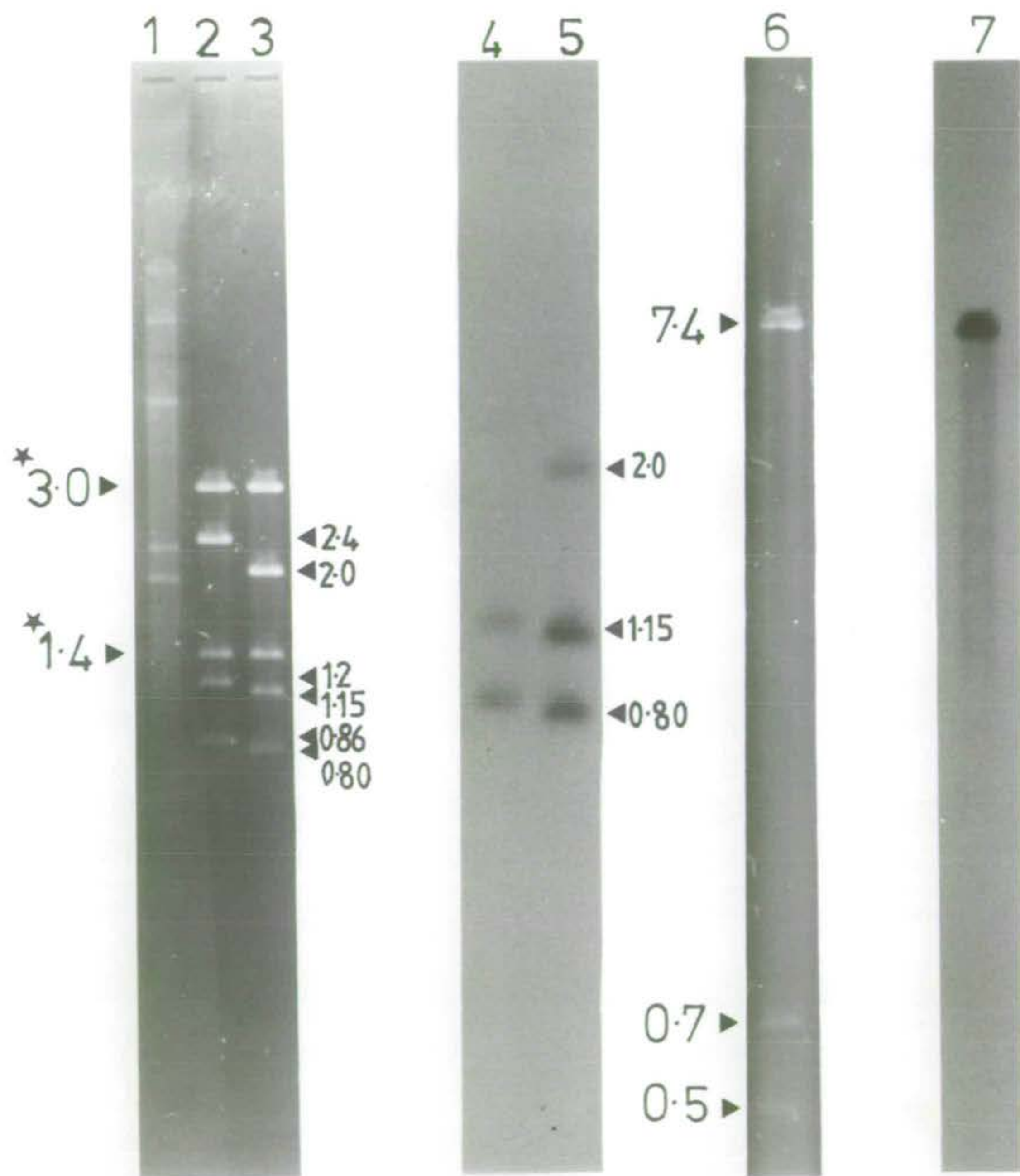


Fig. 4.18

Ethidium bromide stained gels and autoradiographs of pPFrib1 and pPFrib2 hybridised with 18S rRNA. An EcoR1 digest of pPFrib2 (track 1) was transferred to nitrocellulose and probed with 18S rRNA. Homology is to the 0.7 kb fragment (track 2), although some homology is observed to the 7.4 kb fragment, this is due to degraded 28S rRNA in the probe. When pPFrib1 digested with HindIII and AvaI (track 3), is probed with 18S rRNA the strongest homology is to the 2.4 kb fragment (track 4). Again other fragments show homology, due to degraded 28S rRNA in the probe.

Note that the pBR322 derived fragments are shown *

FIG 4.18

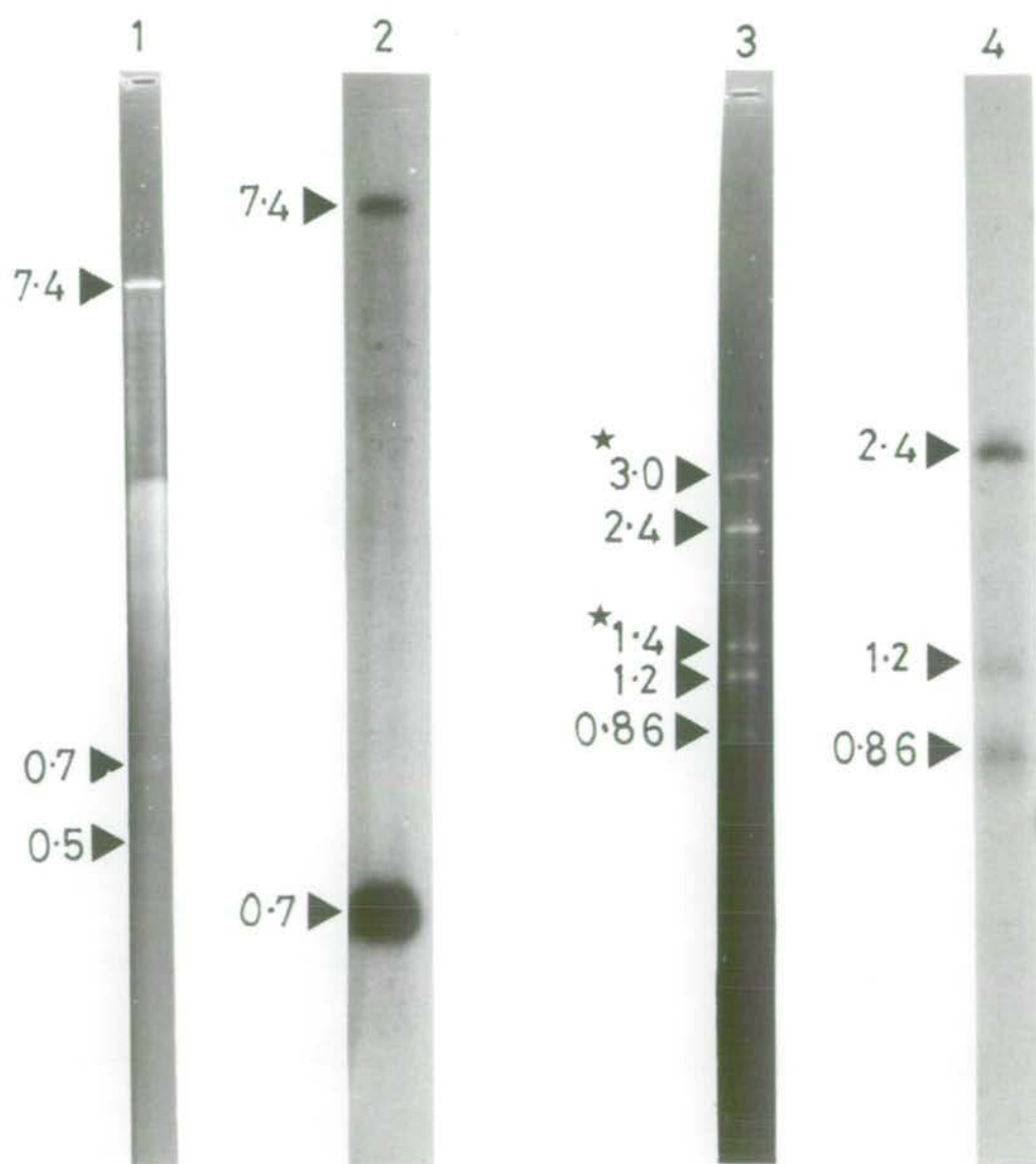


Fig. 4.19

Ethidium bromide stained gel and autoradiograph of pPFrib1 and pPFrib2 hybridised with 5.8S rRNA. When a HindIII/AvaI/EcoR1 digest of pPFrib2 (track 1) is probed with 5.8S rRNA homology is only observed to the 0.5 kb EcoR1 fragment (track 3). On probing a HindIII/AvaI digest of pPFrib1 (track 2), with 5.8S rRNA, the 2.4 kb fragment is identified (track 4).

FIG 4.19

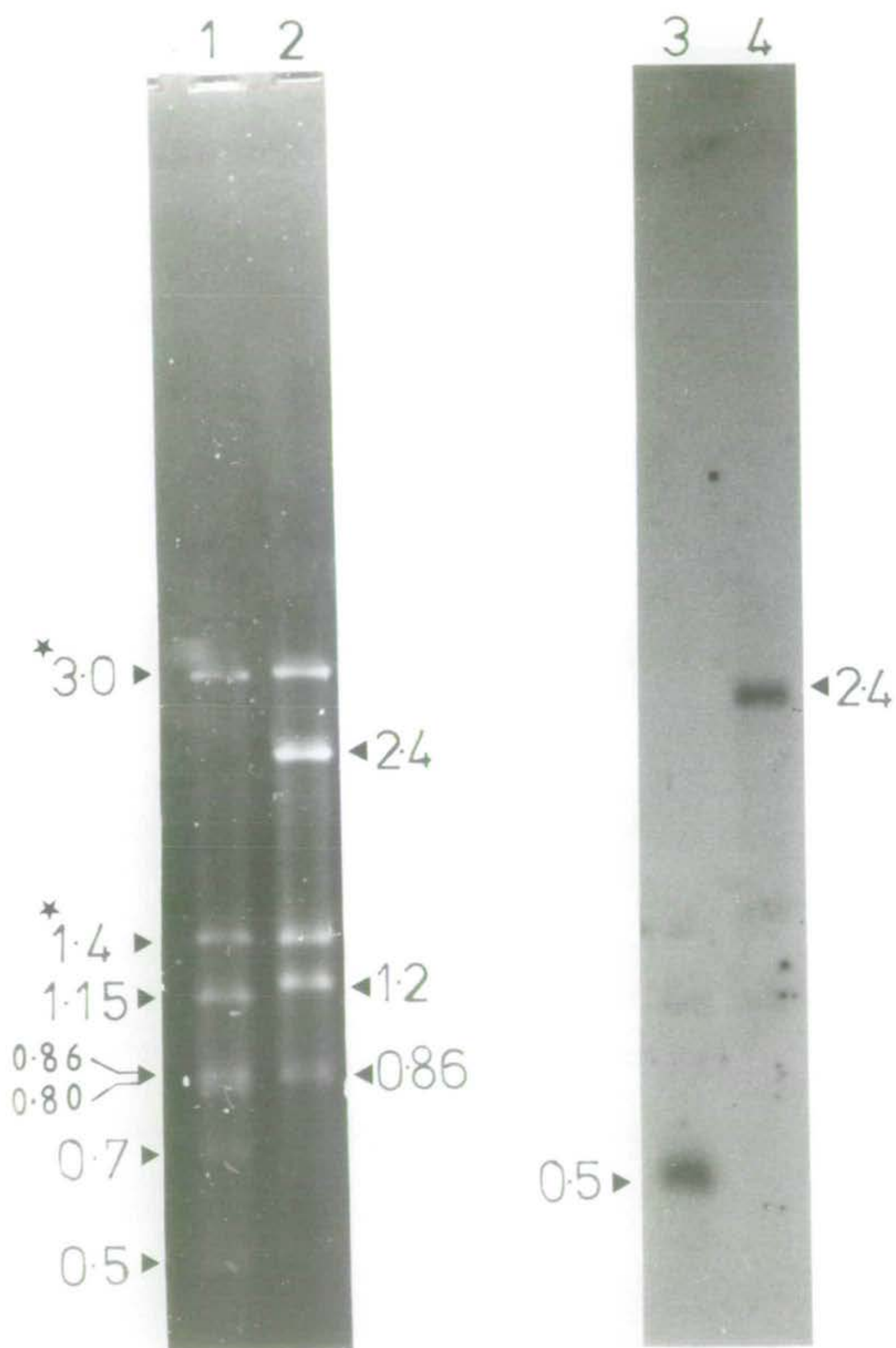
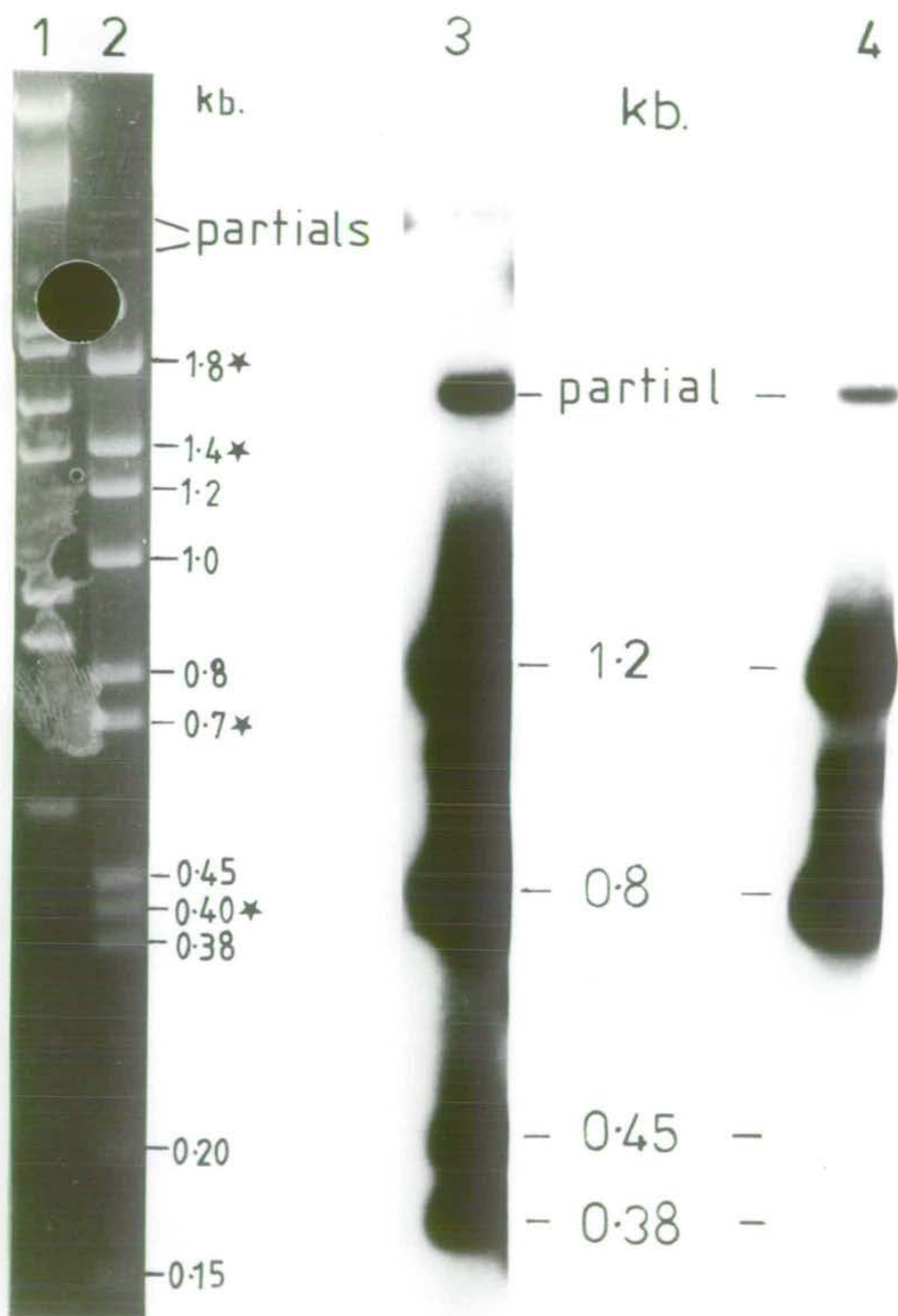


Fig. 4.20

Ethidium bromide stained gel and Southern blot of pPFrib1. When the rDNA recombinant was digested with HindIII/AvaI and DraI (track 2), and hybridised with total rRNA (track 3), the 0.45 kb and 0.38 kb fragments are identified as specific to the 18S rRNA, for they were not identified when 28S rRNA alone was used as a probed (track 4). The size markers (track 1) are HindIII/EcoR1 fragments of λ^{+} DNA. The pBR322 derived fragments are shown (*). Note some partial digestion fragments are observed, one of which shows homology to both probes.

FIG 4.20



sequences homologous to the 3 rRNAs were assigned to the physical maps of pPF*ribl* and pPR*rib2* (Fig. 4.11).

The data presented so far can be summarised as follows. Hybridisation with separated rRNA species demonstrate both pPF*ribl* and pPF*rib2* encode at least part of each of the major rRNA species. These are arranged in the order 18S-5.8S-28S. In this respect I have shown that *P.f.* does not differ from the other eukaryotic systems analysed (Long and Dawid 1980).

However, neither contains the complete gene for either the 18S or the 28S rRNA (see Fig. 4.11). It can be confirmed that the 28S sequences encoded in pPF*ribl* and pPF*rib2* do not together comprise the whole gene from the following 2 observations. Firstly, both contain sequences homologous to the 5.8S rRNA, which is characteristically located 5' to the 28S gene, i.e. both pPF*ribl* and pPF*rib2* encode the 5' end of the 28S rRNA gene. Secondly, if the 28S sequence in pPF*ribl* is contiguous with the 28S rRNA sequence in pPF*rib2*, then an *AvaI* fragment of 2.4 kb should be readily identified in the genome. No such fragment is observed (Fig. 4.21). The remainder of each gene must lie on one of the other *HindIII* fragments identified by pDm238/29 and total rRNA. Although the library was screened further with both pDM238/29 and pPF*rib2* these *HindIII* fragments were not found. A likely explanation for the failure to detect the smaller fragments could be that they are not represented in the λHPF8 library; supporting this view is the fact that the minimum capacity of NM788 is 3-4 kb (Murray *et al*. 1977).

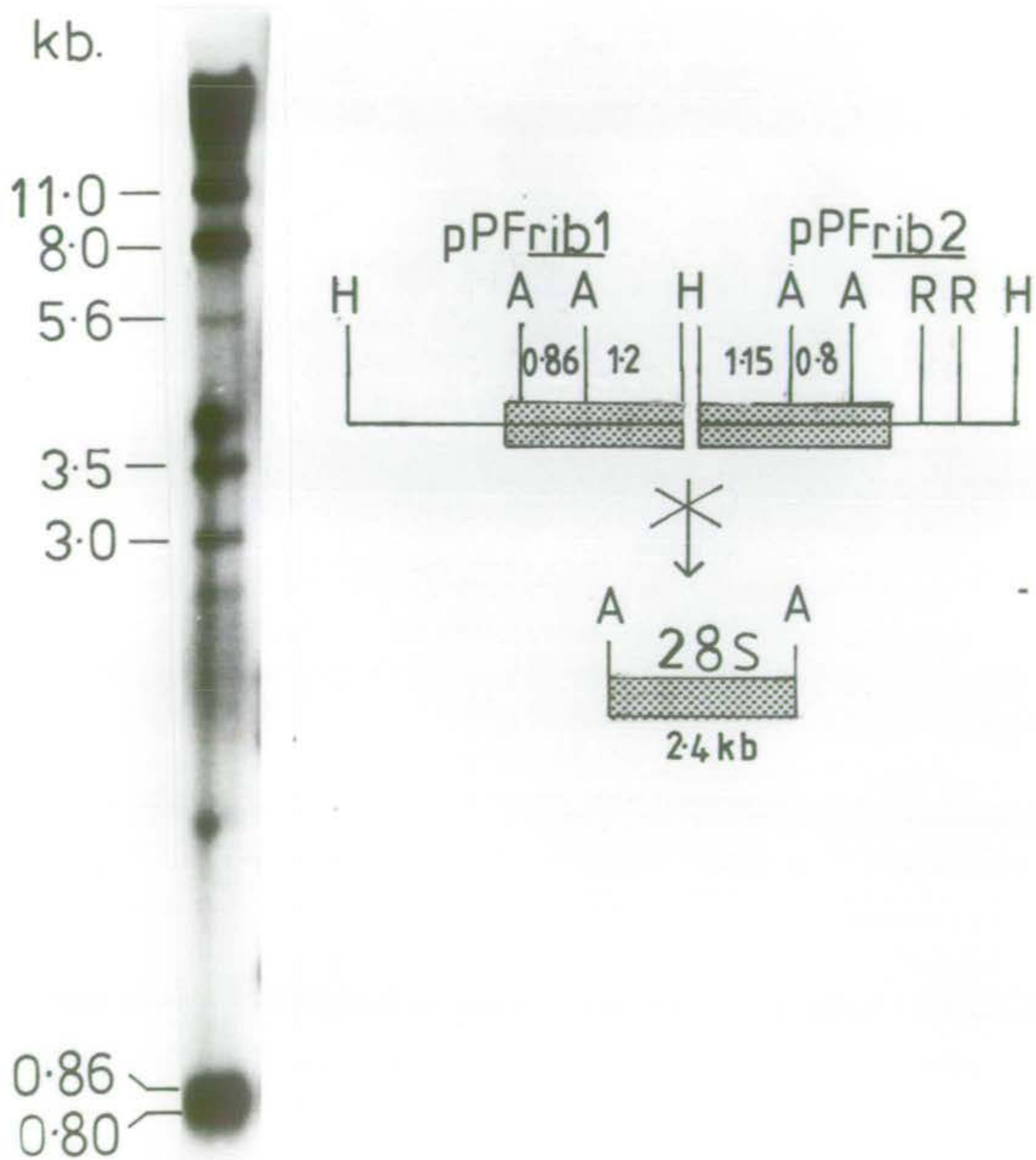
Since both pPF*ribl* and pPF*rib2* encode the 3 rRNA species and yet clearly differ in

(1) overall size 4.2 kb vs. 4.4 kb,

Fig. 4.21

Southern blot of *AvaI* restricted K1DNA hybridised with 28S rRNA. Strong homology is observed to genomic fragments of 11.0, 8.0, 3.5, 3.0, 0.86 and 0.80 kb. There is no homology to a fragment of *ca*. 2.4 kb. This strongly suggests that the fragments present in pPF*rib1* and pPF*rib2* were not contiguous in the genome.

FIG 4.21



(2) distribution of AvaI, BgIII, EcoRl and HincII sites

(3) location of AT rich sequence

see Fig. 4.11. It seems reasonable to conclude that there must be at least two transcription units, which differ slightly, and that pPRribl and pPFrib2 are the analogous HindIII fragments from each unit.

The mapping of the 28S rRNA to the 3 kb HindIII/EcoRl fragment in pPFrib2, and the fact that no 28S homology is detected in 4.1 kb fragment of λPFH8rib2 (see Fig. 4.7) can be explained if the 28S gene is interrupted by a stretch of non-coding DNA (intervening sequence, ivs). The 4.1 kb fragment may be part of such an ivs. From the size of the 28S rRNA the 28S rRNA gene must be at least 4.3 kb (Hyde et al. 1981). At most only 3 kb of this gene is present in pPFrib2, the remaining 1.3 kb being encoded within one of the smaller HindIII fragments, which appear to be absent in the λPFH8 library.

4.7 Isolation and characterisation of a rRNA fragment homologous to 28S rRNA, from an EcoRl genomic library λRlKl

Evidence presented in the previous section suggests that the 28S gene of the Gambian isolate HG13 (Butcher 1981), may possess an ivs. Hybridisation of the 28S rRNA to EcoRl digested genomic DNA of the Thai isolate Kl (Thaithong et al. 1981), identified 3 fragments, the smallest being 8 kb (Fig. 4.4b, track 2). To clone these fragments a genomic library was constructed from EcoRl digested Kl DNA. The restricted Kl DNA was ligated into the EcoRl site of NM1149 and package in vitro (See Methods, 2.16).

Approximately 10^7 recombinant phage per μg of Kl DNA were obtained, and 5000 screened with ^{32}P-labelled 28S rRNA. Three clones identified in this way were all found to contain an 8 kb EcoRl fragment

homologous to the 28S rRNA. One of these, λR1K1*rib2* was analysed further.

Restriction analysis revealed that the 8 kb *EcoR1* fragment (Fig. 4.22, track 3) contains a *HindIII* site, such that a *HindIII*/*EcoR1* digest generates a 5.0 kb and a 3.0 kb fragment (track 4). This 8 kb fragment overlaps the 3 kb *HindIII*/*EcoR1* fragment cloned in pPF*rib2*, since they generate fragments of the same size when further digested with *BglII* (tracks 5 and 6). When *HindIII*/*EcoR1* digested λR1K1*rib2* was probed with labelled 28S rRNA, 90% of the homology was to the smaller fragment (Fig. 4.23 track 3), as determined by densitometry. By this criterion only 300 bp of the 5 kb *HindIII*/*EcoR1* fragment is homologous to the 28S rRNA. As pPF*rib2* encodes at most 3 kb of the 28S rRNA gene, 1 kb or more must be encoded elsewhere, since the 28S rRNA is 4.3 kb (Hyde *et al.* 1981). It therefore follows that the 28S rRNA gene is interrupted by the 5 kb *HindIII*/*EcoR1* fragment of λR1K1*rib2*.

To determine the size of the *HindIII* fragment, of which the 5 kb fragment is a part, the 5 kb *HindIII*/*EcoR1* fragment was labelled and hybridised against restricted genomic DNA. Strong homology is seen to a *HindIII* fragment of 5.4 kb (Fig. 4.24, track 6). Interestingly homology is also observed to the 2.6 kb *HindIII* fragment identified by pDm 238/29 and total rRNA. In addition other bands of higher molecular weight are identified, presumably due to their homology to the non-coding sequence in λR1K1*rib2*, since they are not seen when 28S rRNA is used as a probe (Fig. 4.3). Note also that strong homology is seen to the 8 kb *EcoR1* genomic fragment (track 8) from which the 5 kb *HindIII*/*EcoR1* probe is derived.

These data permit the construction of a genomic map for this

Fig. 4.22

Physical map of λRIKI rib2

λRIKI rib2 is an 8.0 kb EcoR1 genomic fragment of rDNA (track 3), which contains a HindIII site such that a HindIII/EcoR1 double digest gives 5.0 and 3.0 kb fragments (track 4). Cloned in λRIKIrib2 is the overlapping genomic fragment to that present in pPFrib2. Since a HindIII/BglII/EcoR1 triple digest generates the same restriction pattern (tracks 5 and 6). The vector derived fragments are shown (*). Track 2 shows a HindIII only digest of λRIKI rib2. This generates a 3.2 kb fragment which includes ca. 0.2 kb of vector. The size markers are HindIII fragments of λ^{+} DNA.

FIG4.22

1 2 3 4

kb. kb.

3·2 — 8·0 — 5·0 — 3·0

5 6

kb. kb.

5·0 — 4·36*

1·4 — 1·0

*0·9 — 0·7

0·55 — 0·55

*0·47 — 0·50

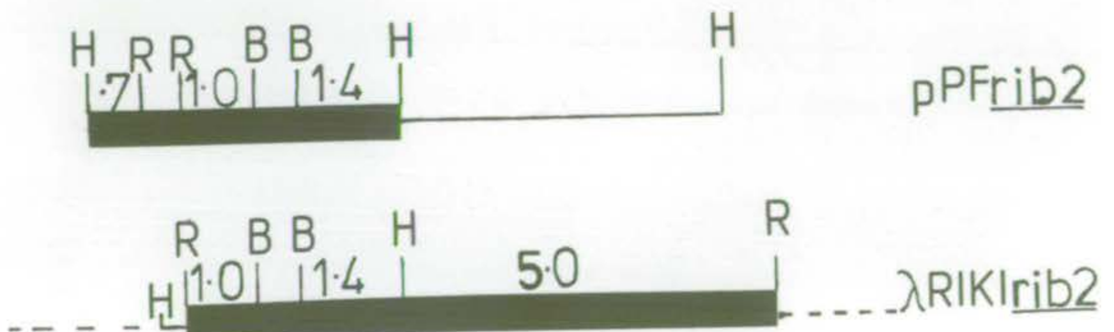


Fig. 4.23

Ethidium bromide stained gel and autoradiograph of restriction digests of λRIKIrib2 and pPFrib2. A HindIII/EcoR1 digest of λRIKI rib2 restricts the insert into 5.0 and 3.0 kb fragments (track 1). A similar digest of pPFrib2 generates insert fragments of 0.5, 0.7 and 3.0 kb (track 2). When the above DNAs are hybridised with 28S rRNA, 90% of the homology is to the 3.0 kb fragment of λ RIKIrib2, and only 10% to the 5.0 kb fragment (track 3). All of the 28S rRNA homology is to the 3.0 kb fragment of pPHrib2 (track 4).

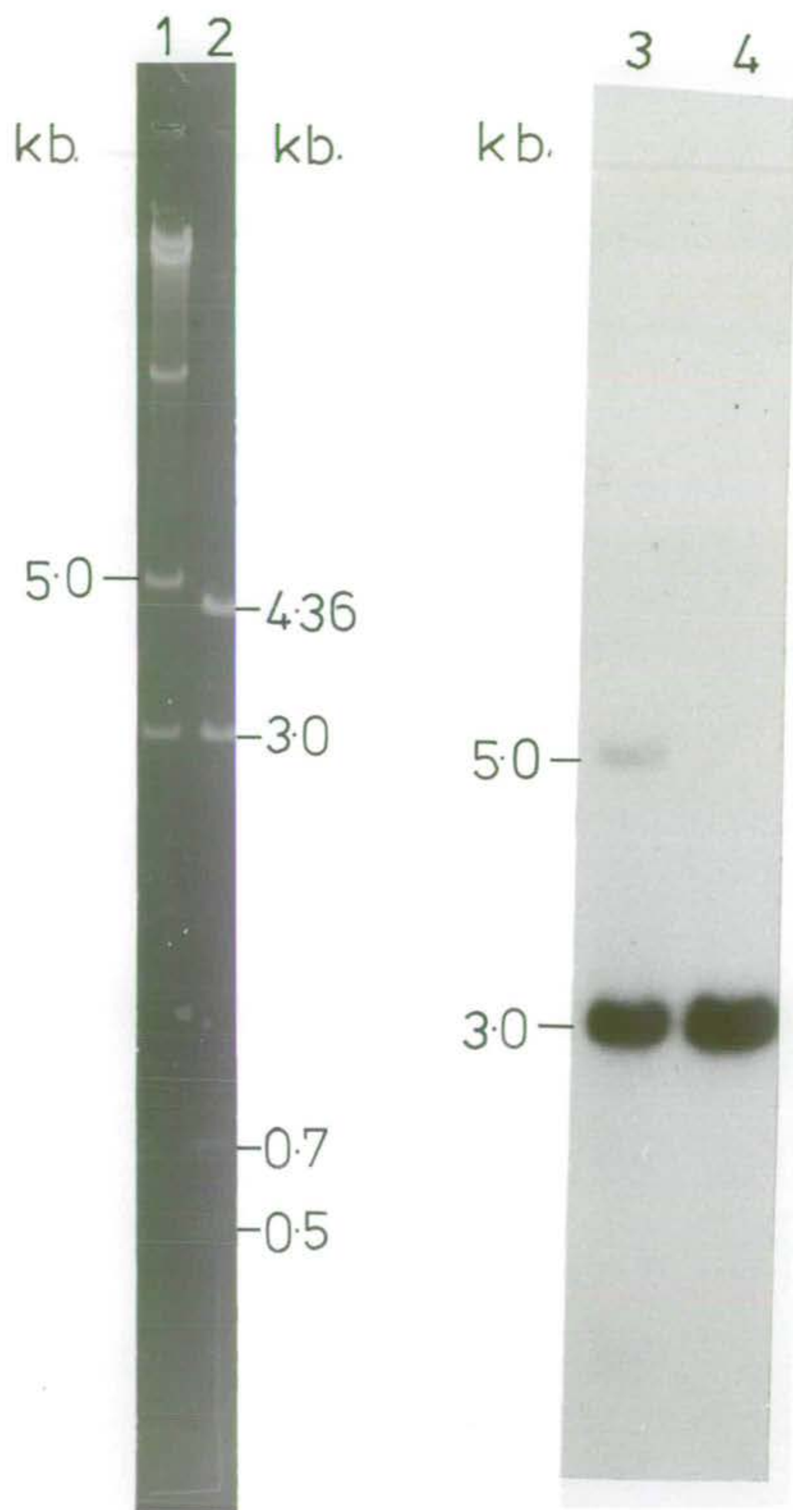


FIG4.23

Fig. 4.24

Ethidium bromide stained gel and autoradiograph of K1DNA. Genomic K1DNA digested with *Hind*III (tracks 2 and 4), and *EcoR*1 (tracks 3 and 5), was hybridised against the 5.0 kb *Hind*III/*EcoR*1 fragment of λRIKI*rib2* (tracks 6 and 8), and pPF4.1 (tracks 7 and 9). The hybridisation patterns are different, compare tracks 6 and 8 with tracks 7 and 9. pPF4.1 hybridises to a 4.1 kb *Hind*III fragment (track 7) and a 13.5 kb *EcoR*1 fragment (track 9); whereas the 5.0 kb fragment of λRIKI*rib2* hybridises to at least 5 *Hind*III fragments. Note the 2.6 kb *Hind*III fragment identified here, is also identified when total rRNA is used as a probe. The size markers are *Hind*III fragments of λ^+ DNA (track 1).

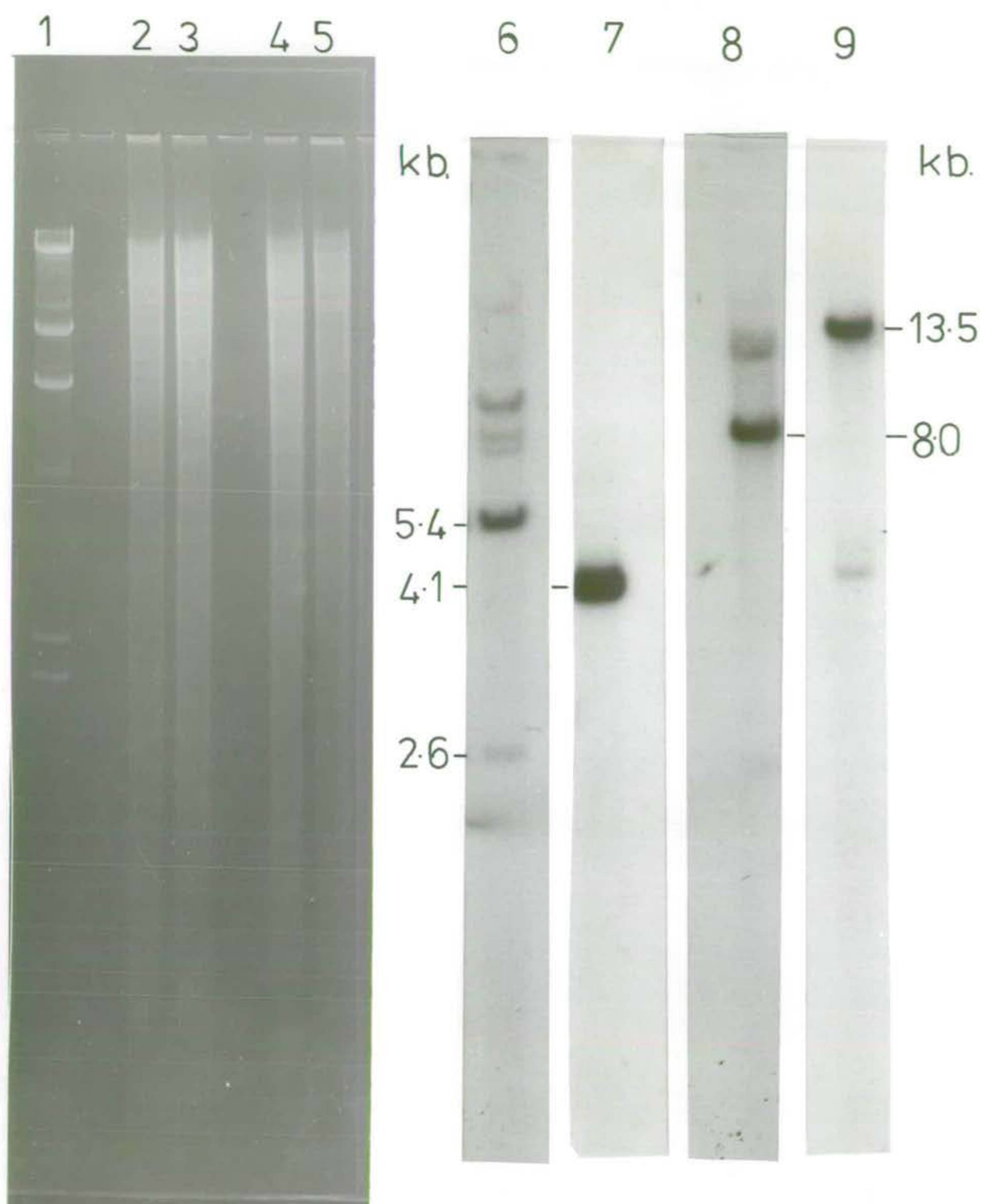


FIG 4.24

region (Fig. 4.25). Present in λR1K1*rib2* may be the analogous genomic fragment to that cloned in λPFH8*rib2*. The 28S rRNA coding sequence common to both appears identical by restriction analysis. In each example the coding sequence is interrupted by a stretch of non-coding DNA. However the non-coding DNA present in λPFH8*rib2* and λR1K1*rib2* is not the same. Firstly, they are of different size, 4.1 kb in λPFH8*rib2* as opposed to 5.0 kb in λR1K1*rib2*. Secondly, pPF4.1 shows a different pattern of hybridisation to K1 genomic DNA. It hybridises strongly to a 4.1 kb *Hind*III fragment (track 9), some homology is also seen to a 4 kb *EcoR1* fragment. The 13.5 kb fragment which shows strong homology to pPF4.1 also cross hybridises weakly to the 5.0 kb *Hind*III/*EcoR1* fragment of λR1K1*rib2* (track 8), presumably due to 28S rRNA sequence homology as these fragments are identified when 28S rRNA is used as a probe. Thirdly, there is no *SalI* site in the 5.0 kb *Hind*III/*EcoR1* fragment (Fig. 4.26 track 3) and finally pPF4.1 does not cross hybridise to λR1K1*rib2* under the stringency conditions used (0.5 x SSC, 37°C) track 5.

The observation that pPF4.1 and the *Hind*III/*EcoR1* fragment are different has at least two possible explanations. One is that pPF4.1 is not related to pPF*rib2* (i.e. it was a coligation product and not a *Hind*III partial cloned in λPFH8*rib2*), and the fact that they hybridise to common *SalI*/*Bgl*II, *EcoR1* and *HpaI* fragments in the genome of K1 was merely fortuitous. The second more exciting possibility is that pPF 4.1 and the 5.0 kb *Hind*III/*EcoR1* fragment of λR1K1*rib2* are representatives of 2 types of ivs which are unrelated; a situation which exists in *D.melanogaster* between Type I and Type II ivs (Wellauer and Dawid 1978).

Fig. 4.25.

Genomic map of rDNA transcription unit. Portions of the genomic DNA that have been cloned and analysed in detail are indicated. The 5.4 and 2.6 kb fragments have only been identified by genomic hybridisation experiments, and as yet no overlapping fragment joining them has been isolated. The cross hybridisation of λR1K1rib2 to the 2.6 kb fragment is probably due to the 28S rRNA sequence. Two types of intervening sequence are shown although a complete 28S rRNA gene containing ivs 2 has not yet been cloned. Some restriction sites are shown and are HindIII, (H); EcoR1 (R); and SalI (S). For more detailed restriction map see Fig. 4.11.

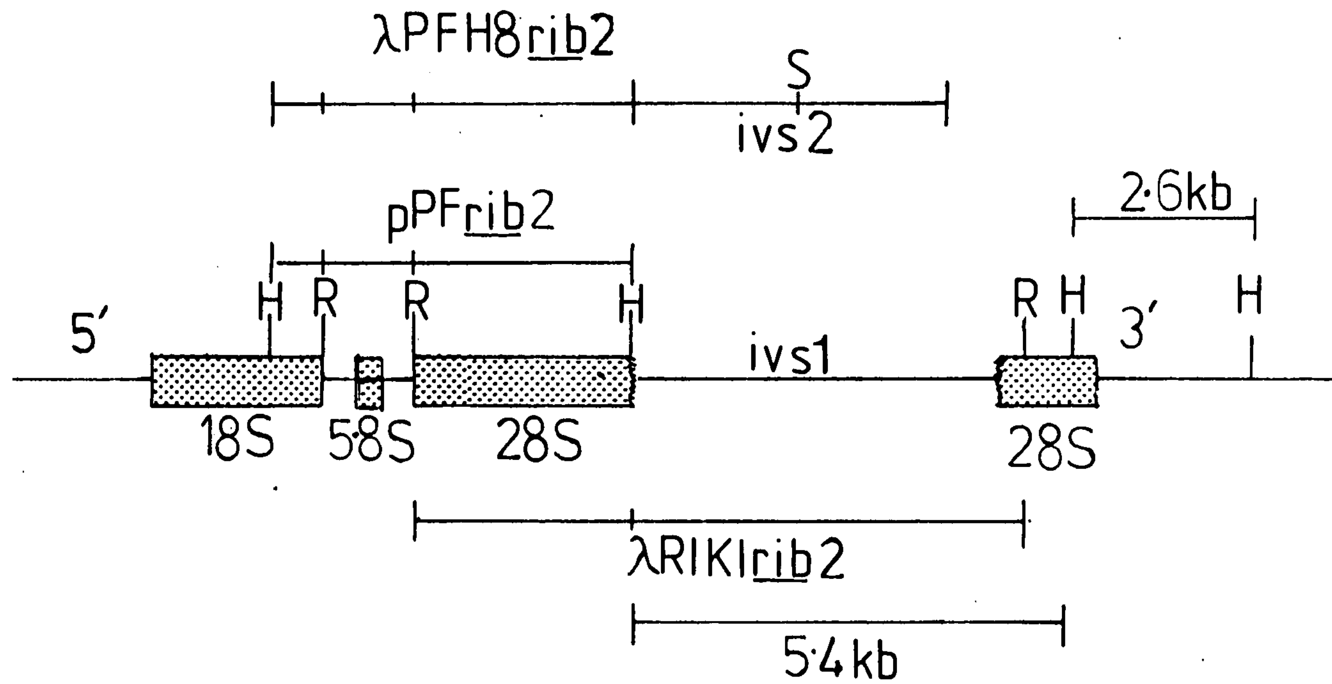


FIG 4.25

Fig. 4.26

Ethidium bromide stained gel and autoradiograph of λRIKIrib2. An EcoR1/HindIII digest generates fragments of 5.0 and 3.0 kb (track 2) 90% of the 28S rRNA homology is to the 3.0 kb fragment (track 4). The insert lacks a SalI site (Track 2), the 3.2 kb fragment derives from the vector NM1149. There is no homology to pPF4.1 (track 5), therefore 5.0 kb and 4.1 kb fragment of pPF4.1 appear to be unrelated. The size markers are EcoR1 fragments of λ^{+}DNA (track 1).

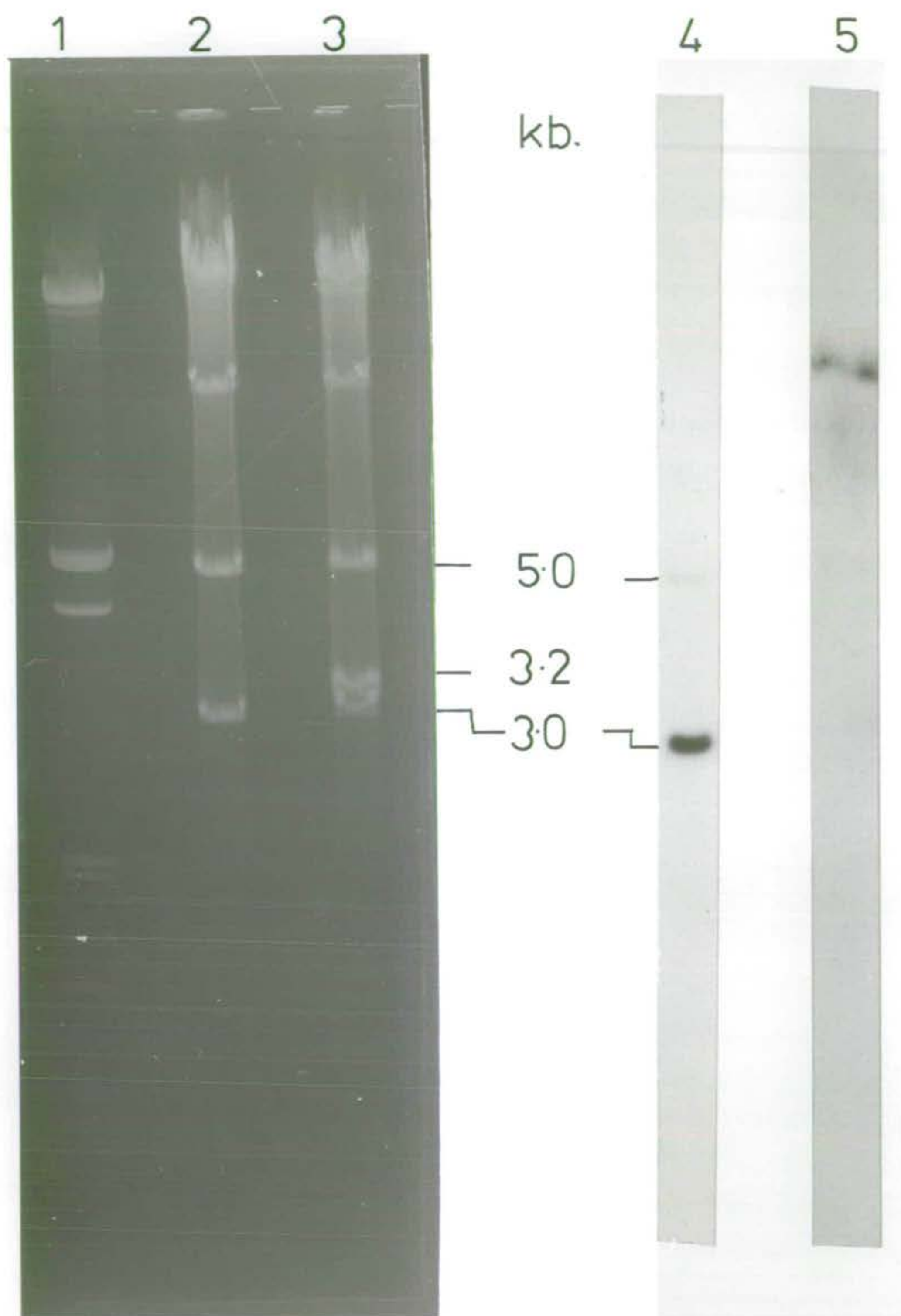


FIG 4.26

Confirmation that pPF4.1 is an ivs awaits the isolation of a 28S rRNA gene whose coding sequence is interrupted by pPF4.1. However the data concerning λR1K1*rib2* argues that some of the 28S rRNA genes in *P.f.* are interrupted by the 5.0 kb *Hind*III/*Eco*R1 fragment. Although the complete 28S rRNA gene has not been cloned, preliminary genomic hybridisation analysis suggests that the 3' end is encoded by the 2.6 *Hind*III fragment. It remains to be seen whether the 28S rRNA gene in the other transcription unit, of which pPF*rib1* is a part, is also interrupted. As yet no overlapping fragment containing pPF*rib1* has been cloned.

4.8 Estimation of the number of copies of the 4.4 kb and 4.2 kb fragments in the genome of *P.falciparum*

Repeated genes are generally ubiquitous, but are more frequent in eukaryotic than in prokaryotic cells. Consideration of structural, functional and evolutionary aspects, has led to the repeated genes being classified in two broad groups of 'dosage' and 'variant' repetition. Dosage repetition means that the genes are present in many copies because the demand for their product is too great for a single gene. This phenomenon is exhibited primarily by the rRNA genes. Protein coding genes frequently occur in families of sequences in a pattern called variant repetition. These repeated genes are not identical, but related.

The rRNA genes are the best studied examples of dosage repetition, and their redundancy has been determined in a large number of eukaryotic species using saturation hybridisation with labelled rRNA in solution and nuclear DNA bound to nitrocellulose filters (Gillespie *et al.* 1965). Another method for estimating the frequency of a particular sequence

in a genome employs Southern hybridisations (Lis et al. 1978). Here the degree of hybridisation is an indication of copy number, and can be estimated by comparison with known standards. This protocol was employed to estimate the number of copies of pPFribl and pPfrib2 present in the genome of P.f.

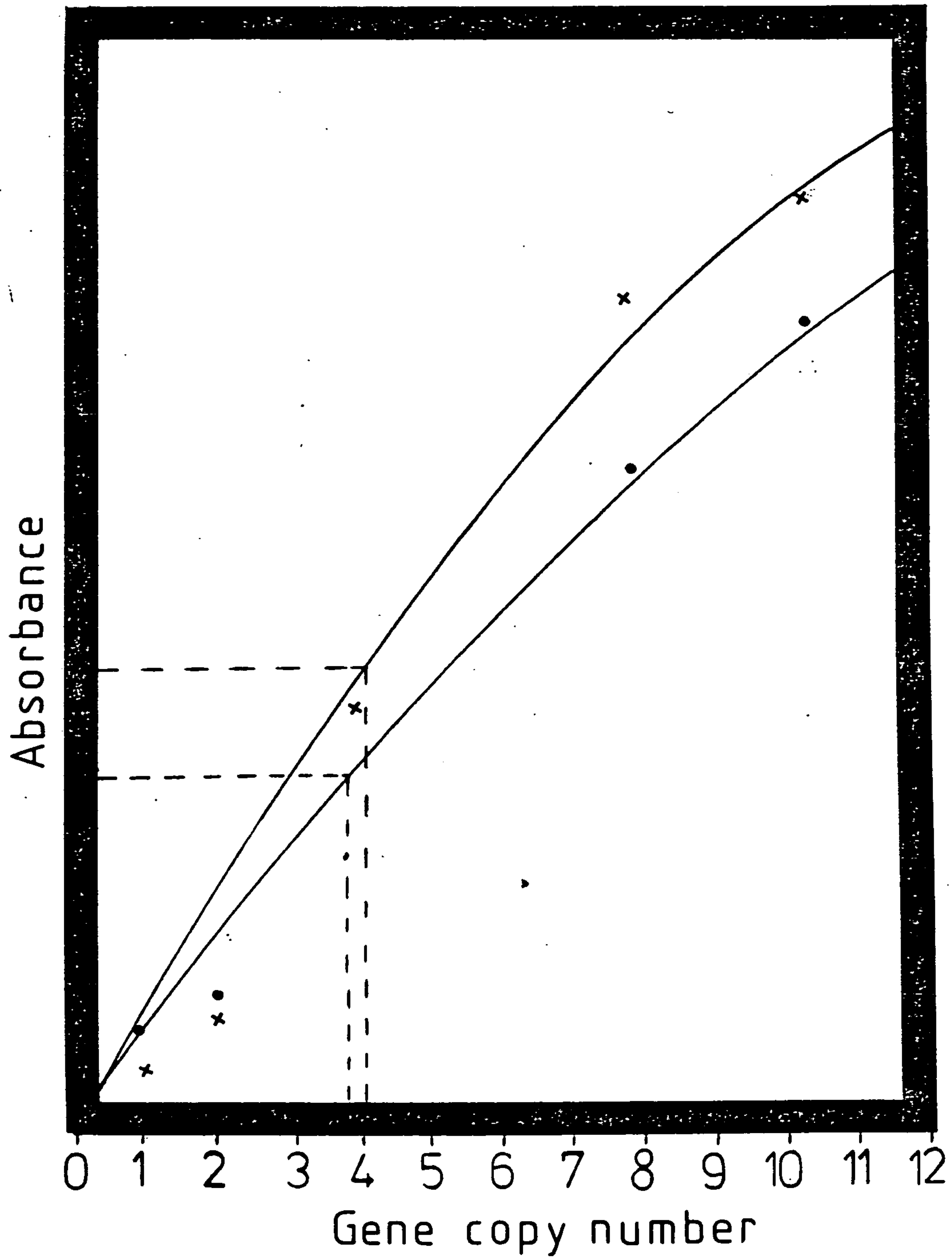
Knowing the DNA content of the genome of P.falciparum is 2.2×10^4 kb (Goman et al. 1982), HindIII restricted pPFribl and pPFrib2 were diluted in salmon sperm DNA such that their concentrations were equivalent to 1,2,4,8,10 and 12 copies per genome of P.f. Genomic DNA from two different isolates of P.f. Kl and Tak 9 (Rosario V. 1981), was restricted with HindIII, and together with the diluted standards fractionated on a 0.7% agarose gel. The DNA was transferred to nitrocellulose and hybridised with ^{32}P-labelled pPFribl and pPFrib2. The degree of hybridisation to the 4.1 kb and 4.2 kb fragments in the genome, and to the diluted standards was quantified by scanning densitometry. The results are represented graphically in Fig. 4.27. It can be seen that there are approximately 4 copies of both pPFribl and pPFrib2, in the genome of P.f. isolated Kl and Tak 9. The equal number for both pPFribl and pPFrib2 is consistent with the observation that they show equal degrees of hybridisation to both pPm 238/29 and total rRNA (Figs. 4.3 and 4.4). Taken together they suggest that there are 8 rRNA genes in total, and these are organised into 2 classes of 4 each.

This number was obtained using an estimate for the genome size of P.f. of 2.2×10^4 kb (Goman et al. 1982). This figure was derived from the amount of DNA extracted from a known number of parasites. Other workers have estimated the genome size of P.f. to

Fig. 4.27

Quantitation of rDNA copy number by Southern blot analysis. The rDNA clones pPFrib1 and pPFrib2 were diluted in 4 µg of salmon sperm DNA and digested with *Hind*III. Aliquots equivalent to a range of copy numbers from 0-12 were fractionated on an agarose gel, together with 4 µgs of *Hind*III digest Tak 9 and K1 genomic DNA. After transfer to nitrocellulose the blot was hybridised with either pPFrib1 or pPFrib2 radiolabelled by nick translation. The degree of hybridisation was estimated by scanning densitometry, the absorbance is equivalent to the area under peak of the densitometer scan. The copy number of pPFrib1 (X) and pPFrib2 (●) was estimated by comparison with the amount of hybridisation observed to the Tak 9 and K1 genomic DNA.

FIG 4.27



be 3.8 x 10^5 kb, by comparing the kinetics of reassociation of the non-repetitive DNA with that of E.coli (Hough-Evans and Howard 1982). This estimate is approximately 17 times higher than that obtained for P.f. and other Plasmodia (see below), and if used would result in an estimate of 136 copies of the rRNA genes in P.f. This higher copy number would indicate that the rRNA genes in P.f. are moderately repetitive as has been described for other eukaryotes (Long and Dawid 1980).

The low copy number of 8 rRNA genes is in good agreement with the 4 copies recently obtained for the rRNA genes of P.berghei (P.b.) (Dame and McCutchan 1983a). The number they obtained is based on a genome content of 0.05 pg/nucleus (Bahr et al. 1972), estimates by other workers put the DNA content of P.b. at 0.025 pg/nucleus (Dore et al. 1980, Cornelissen et al. 1983), which is consistent with 0.018 pg/nucleus for P.knowlesi (Gutteridge et al. 1970) and 0.02 pg/nucleus for P.f. (Goman et al. 1982). A DNA content of 0.025 pg/nucleus for P.b. would imply that there are 8 rRNA genes, rather than the 4 reported (Dame and McCutchan 1983a).

The 17 times higher genome size reported for P.f. by Hough-Evans and Howard (1983), is therefore in conflict with that reported for the genome size of P.f. (Goman et al. 1982), and other Plasmodia (Bahr et al. 1972, Cornelissen et al. 1983, Dore et al. 1980, Gutteridge et al. 1970), and gives P.f. a genome size 2-3 times that of D.m. (Rasch et al. 1971). This would also result in a DNA:RNA ratio of 20:1, since Hyde et al. (1983), estimate the amount of RNA in P.f. to be equivalent to the DNA content reported by Goman et al. (1982). The excess DNA:RNA could be explained if a large proportion of the genome of P.f. was repetitive and non-coding, however Hough-Evans

and Howard report that only 10% of the genome of *P.f.* is repetitive. This is lower than that reported for *P.f.* (Goman *et al.* 1982), and for *P.b.* (Dore *et al.* 1980).

Until other estimates for the genome size of *P.f.* are published, it seems reasonable to take the genome size of *P.f.* as 2.2×10^4 kb (Goman *et al.* 1982), and the number of rRNA genes as 8. *P.f.* and *P.b.* therefore are exceptional eukaryotes in regard to the low number of rRNA genes per haploid genome.

4.7 Conclusions

Two genomic rDNA fragments from a Gambian isolate HG13 (Butcher 1981) have been cloned and analysed in detail. Sequences homologous to the 28S, 18S and 5.8S rRNA have been mapped to these fragments, and are arranged 18S-5.8S-28S. This arrangement is the same as that reported for most eukaryotes (for review see Long and Dawid 1980), where the 18S rRNA gene is at the 5' end. In *P.berghei* although no 5.8S rRNA gene was mapped, it has been shown that the 18S rRNA gene is 5' to the 28S rRNA gene in the rRNA transcription unit (Dame and McCutchan 1983b). It is reasonable to assume therefore that in *P.falciparum* the transcription unit is also ordered 5'-18S-5.8S-28S-3'. In *P.falciparum* I have also shown that the 5.8S rRNA gene is located in a stretch of A+T rich DNA, as has been reported for *D.melanogaster* and proposed as a processing point (Pavlakis *et al.* 1979).

When the rRNA was fractionated under denaturing conditions the 5.8S rRNA appeared as a discrete band, implying that unlike the *Trypanosomatids*, *Plasmodia* do not possess the numerous small rRNAs (Cordingley and Turner 1980, Hernandez *et al.* 1983). Furthermore no 5S rRNA was detected leading to the assumption that 5S rRNA may be absent in *P.falciparum*. In support of this view is the

observation that no homology was detected to cloned Xenopus laevis 5S RNA in the genome of P.berghei (Dame and McCutchan 1983b).

Demonstration that the two rDNA fragments contain sequences homologous to all 3 rRNA species, but are clearly different, allows one to conclude that they are derived from separate transcription units. The high molecular weight RNA homologous to the rDNA clones (Hyde et al. 1981) is presumably the rRNA precursor. Each rDNA fragment is represented 4 times in the genome, this number did not vary between the two isolates compared, and implies that there are 8 rRNA transcription units of two distinct classes. The low copy number and the observation of at least two classes of transcription unit has also been reported for P.berghei (Dame and McCutchan 1983a).

The low copy number obtained for P.falciparum and P.berghei sets them apart from other parasitic blood protists such as Leishmania donovani and Trypanosoma cruz, which have 166 and 144 nuclear rRNA genes respectively (Leon et al. 1978, Castro et al. 1981). In other protozoa the rRNA genes are amplified on an extrachromosomal element, for example although Tetrahymena has only a single rRNA gene in its micronucleus (Yao and Gall, 1977), in the macronucleus the rRNA genes are amplified to approximately 600 copies on an extrachromosomal palindrome (Yao et al. 1979). The absence of satellite DNA in P.falciparum (Goman et al. 1982, Pollock et al. 1982) precludes the amplification of the rRNA genes on an extrachromosomal element. The 8 nuclear copies therefore represent the full complement of rRNA genes.

However these data were obtained from in vitro cultures of the asexual stages of P.falciparum, and it would be of interest to ascertain

if the copy number of rRNA genes varies with changes in the life cycle. Intererythrocytic asexual parasites give rise to micro (male) and macro (female) gametocytes, and several lines of evidence show that macro-gametocytes have many more ribosomes than microgametocytes (for review see Sinden 1983). This would agree with the general idea that female germinal cells must be ready to sustain more abundant and diverse protein synthesis. Can the increase in the number of ribosomes be correlated with an amplification in the number of rRNA genes? Preliminary evidence emerging from the study of infective (i) strains of P.berghei suggests so. It has been previously demonstrated that gametocyte infectivity (to the mosquito) requires an amplification and rearrangement of repetitive DNA (Birago et al. 1982, Dore et al. in press). Genomic hybridisations of i and ni DNA with pPFrib2 indicates that the rRNA genes are amplified to approximately 100 copies, and undergo a rearrangement, as part of this process (C. Frontali pers. comm.). This evidence suggests that the mechanism allowing ribosome over-production in macrogametes is direct amplification of the ribosomal genes. Breakdown of this mechanism would thus result in loss of infectivity of syringe passaged cultures without blocking asexual growth of the parasite.

Finally, analysis of a rDNA fragment cloned from the genome of a Thai isolate K1 (Thaithong et al. 1981) indicates that the 28S rRNA gene in P.falciparum may contain an intervening sequence. Intervening sequences in the large subunit rRNA gene have been reported for D.m. (Glover et al. 1977), Tetrahymena (Wild and Gall 1979), and Physarum (Campbell et al. 1979). The presence of an inserted sequence in the large s.u. gene of T.brucei has been noted (Hasan, et al. 1982), with the characteristics of a mobile element. In P.falciparum

although sequences homologous to the inserted DNA are detected elsewhere they are not repetitive or amplified.

In P.berghei the possibility of an intervening sequence has been reported for the 28S rRNA gene of one of the transcription units (Dame and McCutchan 1983a), due to a variation in size of a HindIII genomic rDNA fragment. Another explanation for this observation is a variation in size of the internal transcribed spacer. Since neither fragment has been cloned, it is impossible at this point to delineate between these two possibilities.

In P.falciparum the interrupted 28S rRNA gene was cloned from the transcription unit characterised by having 2 EcoR1 sites in the internal transcribed spacer. It remains to be seen if both classes of transcription unit have interrupted 28S rRNA genes, and whether more than one type of ivs exists in each strain. Moreover, are all strains of P.falciparum the same, or do they vary like Tetrahymena pigmentosa where the 25S rRNA genes is interrupted in strain GUM, but intact in 8 ALP (Wild and Gall, 1979)? Preliminary evidence suggests that P.f. strains K1 and HG13 both have 28S rRNA genes with an ivs. In addition does the possession of an ivs render the gene transcriptionally inactive as in D.m. (Long et al. 1981), or have no effect as in Tetrahymena (Wild and Gall 1979)? If the genes are transcriptionally active, how are they processed?

Evidence is emerging that the number of copies of the rRNA genes vary with changes in the life cycle, and it will be of interest to determine whether or not expression of the rDNA is also differentially controlled throughout the life cycle; bearing in mind P.falciparum's parasitic existence and alternation between two quite markedly different hosts (man and mosquito).

CHAPTER 5

ISOLATION OF PUTATIVE ACTIN GENES FROM GENOMIC AND cDNA LIBRARIES OF *P.FALCIPARUM*

5.1 Introduction

5.2 Results and Discussion

5.2.1 Sequences in the genome of *P.falciparum* homologous to the *Dictyostelium* actin probes pcDdB1 and Eco 13.4.9

5.2.2 Isolation of putative actin genes from a *Hind*III genomic library

5.2.3 Identification of a mRNA transcript, and isolation from a *P.falciparum* cDNA library, of sequences homologous to a *Dictyostelium* actin gene

5.3 Conclusions

5.1 Introduction

There are three fundamental mechanisms which have evolved for cell motility. These are ultrastructurally, biochemically and mechanically distinct. Two of these, bacterial flagella and eukaryotic flagella serve to move cells through fluid media, but they function differently (for review see. Adler, J., 1975; Clarke, M. and Spudich, J. A., 1977). The third motility mechanism, that based on microfilaments is responsible for amoeboid-type movement of a cell on a solid substratem. The major component of the microfilaments of non-muscle cells is actin; it is a major component of the cytoskeleton, as well as being involved in many forms of cellular motility (for a review of the biochemistry of actomyosin-dependent cell motility, see Korn, E. D., 1978).

Analysis of mammalian actin protein by high resolution two dimensional gel electrophoresis shows 3 iso-electric forms. Two forms, cytoplasmic β and γ, are found in all cell types while the third, the most acidic form, α actin, is found exclusively in muscle cells (Korn, _loc cit_). A striking feature of actins is their very high sequence conservation, even between different species. The amino acid (a.a.) sequence of the cellular slime mold _Physarum_ actin, has been determined and compared to mammalian actins. _Physarum_ and mammalian cytoplasmic (β and γ) actin differ in only 4% of their a.a. residues, i.e. lower eukaryotic actin is more similar to mammalian cytoplasmic actin, than cytoplasmic actin is to skeletal actin (Vandekerckhove, _et al_. 1978). Comparison of _Physarum_ and _Dictyostelium_ (cellular slime mould) a.a. sequences with those of mammalian cytoplasmic actin shows that the a.a. differences are clustered in residues 1-6

and after 159. The a.a.s 7-159 are conserved and therefore appear important for non-muscle actins. Consistent with this view of conservation for function, is the observation that skeletal actins show 7 changes in this region (Vanderkerckhove, et al. 1980).

The evolutionary requirement for the functionally conserved region of non-muscle actins, is thought to be a consequence of the interaction of actin with the numerous actin-binding proteins (for review, see Craig and Pollard, 1982). A major protein with which actin interacts is myosin (which is an ATPase that forms an enzymatically active complex with actin), and presumably provides the energy for the motility events. The two other major proteins known to bind with actin in microfilaments are tropomyosin and α actinin, and all of these proteins are widely distributed; for example the small soil amoeba Acanthamoeba contains one actin, 2 myosins and 4 gelactins (Korn, loc cit). In the red blood cell, actin forms a complex with spectrin and ankyrin, and together these anchor glycophorin and band 3 proteins in the plasma membrane (Pasvol and Wilson, 1982). Clearly the links between actin, spectrin, ankyrin, band 3 and glycophorin afford a pathway that the invading malaria parasite must deform during entry.

The a.a. sequence conservation of actin is reflected in the nucleotide sequence of the organisms compared. Although during evolution there has apparently been pressure to maintain a specific sequence in parts of this protein, the codon usage has changed resulting in the nucleic acid sequences being more divergent. However nucleotide conservation is also reflected between disparate organisms; Hanukoglu et al. (1983), compared the sequence of actin cDNAs from human, Dictyostelium and yeast. Human actin shows 98

and 85% homology to the Dictyostelium and yeast actins respectively.

The conservation of sequence for actin coding regions, has permitted the isolation of actin genes from a number of organisms, using as a probe, actin sequences previously cloned from a different species. For example, the actin genes in both Drosophila and the nematode Caenorhabditis elegans were isolated using the Dictyostelium cDNA actin clone pcDdB1 (Bender et al. 1978; Fyrberg et al. 1970; Files, et al. 1983). The actin genes from sea urchin were isolated using cloned Drosophila actin (Durica et al. 1980), and actin genes from the ciliated protozoan Oytricha fallax by homology to yeast actin (Kaine and Spear, 1980).

A more comprehensive review of actin genes is presented in Chapter 1. Also shown in Chapter 1 is the life cycle of Plasmodium falciparum, and a brief review of it is presented below.

When a female Anopheles mosquito ingests the blood of a human host, the infected erythrocytes set the parasites free in the mosquito's stomach. In the mosquito the mature sexual cells (gametocytes) commence sexual reproduction. The male gametocytes nucleus divides into 4 to 8 nuclei each of which forms a long flagellum; they shoot out from the original cell, lash about and then break free. This process is called exflagellation. The female (macrogamete) attracts the male (microgamete), and they fuse to form a zygote. This is at first a motionless globular body, but it elongates and becomes motile; this worm like stage is known as an ookinete. The ookinete forces its way out of the stomach and into the body cavity, where it forms an oocyst. The sporozoites develop within the oocyst, which later bursts liberating thousands of mobile sporozoites. These move to the

salivary gland and the female mosquito is now infective.

Following inoculation by the mosquito bite, the sporozoites move to the liver and enter the hepatocyte cells, where they undergo pre-erythrocytic schizogony. When this stage is complete, the schizont bursts, the liver cells rupture, releasing merozoites into the blood circulation. The merozoites invade the red blood cells, and undergo erythrocytic schizogony. This erythrocytic cycle is repeated until some of the merozoites develop into gametocytes, which are ingested by another mosquito, so completing the life cycle.

It is obvious that motility is involved throughout the life cycle. This observation as well as the development of flagella, suggests that *P.falciparum* might use actin as part of its motility machinery. The isolation of *P.falciparum* actin genes would allow the study of a gene(s) coding for a structural protein, which in many cases is known to be developmentally regulated (for review see Firtel, 1981), and could play an important role in the life cycle of the parasite. The conservation in actin nucleotide sequences, permits the isolation of this gene by cross homology to a previously cloned actin sequence; thus obviating the need to employ sophisticated techniques such as, cloning into an expression vector, hybrid selection and *in vitro* translation of specific mRNA.

The experiments described here are preliminary and they detail the isolation from *P.falciparum* of putative actin genes, due to their homology to actin genes cloned from the cellular slime mould *Dictyostelium discoideum*.

5.2. Results and Discussion

5.2.1 Sequences in the genome of P.falciparum homologous to the two Dictyostelium actin probes pcDdBl and Eco. 13.4.9.

The two D.discoideum actin genes used as probes in this study are pcDdBl (Bender, et al. 1978), and Eco.13.4.9. (a kind gift of Dr. J. Williams). Maps of the recombinant plasmids are given in Fig. 5.1. Note that pcDdBl is a 1050 bp cDNA copy of Dictyostelium cytoplasmic polyA$^+$ mRNA, and therefore represents over 95% of the coding sequence. It shows homology to an actin gene cloned in recombinant plasmid M6, which contains 6 kb. of genomic DNA (Bender, et al. 1978). However, subsequent analysis indicates that the mRNA copy present in pcDdBl is not transcribed from the M6 actin gene (McKeown, et al. 1978; McKeown, et al. 1981), but could be the transcript of either pDd actin 2-sub-1, pDd actin 5 or 7 as they all code for the same NH_2 terminal peptides (Firtel, et al. 1979). The restriction map of pcDdBl given in Fig. 5.1 was taken from Firtel, et al. (1979).

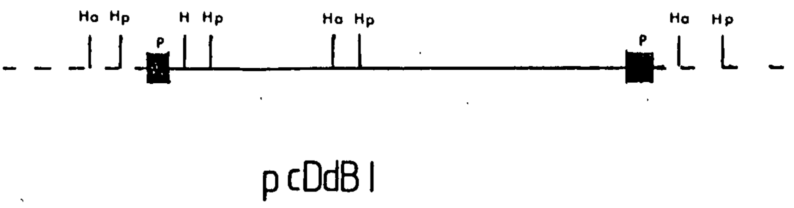
By contrast the plasmid Eco 13.4.9 contains a 4.4 kb EcoRl genomic DNA fragment, which includes a complete actin gene plus flanking sequences. The map presented is that kindly provided by Dr. J. Williams.

DNA extracted from the Kl isolate (Thaithong et al. 1981) of P.falciparum, was digested with HindIII, fractionated on a 0.7% agarose gel, and transferred to nitro cellulose. The filter was hybridised with pcDdBl radio-labelled by nick translation (Methods). An autoradiograph of the filter is presented in Fig. 5.2 (track 1). Five HindIII fragments were identified in this way and their sizes

Fig. 5.1

Restriction map of pcDdBl, the cDNA was cloned by G+C tailing into the Pstl (P) site of pBR322 (-----; Bender, *et al*. 1978). The *Hind*III (H) site is at the 5' end of the structural gene. The *Hae*III (Ha), and *Hap*II (Hp) sites are shown.

Restriction map of Eco 13.4.9, the 4.4 kb genomic DNA was cloned into the *EcoRl* site of PMB9 (----). The structural gene is shown hatched, along with the direction of transcription.



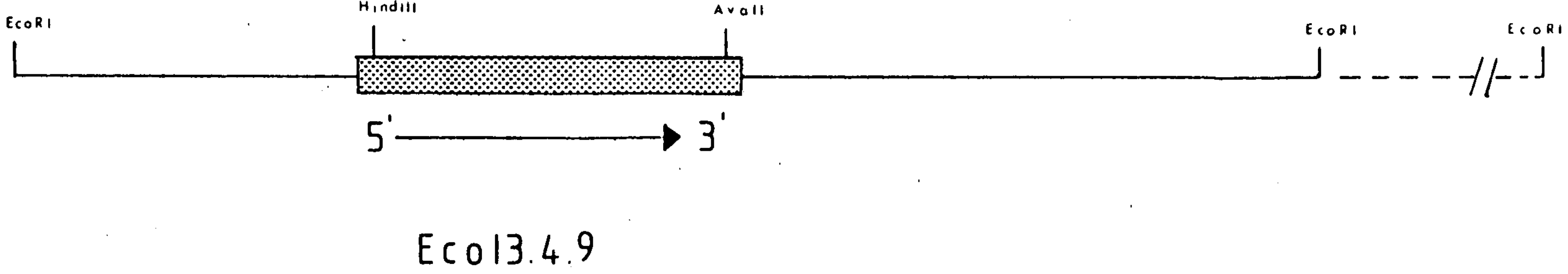


FIG 5.1

are 5.5, 4.8, 3.8, 2.8 and 0.6 kb.

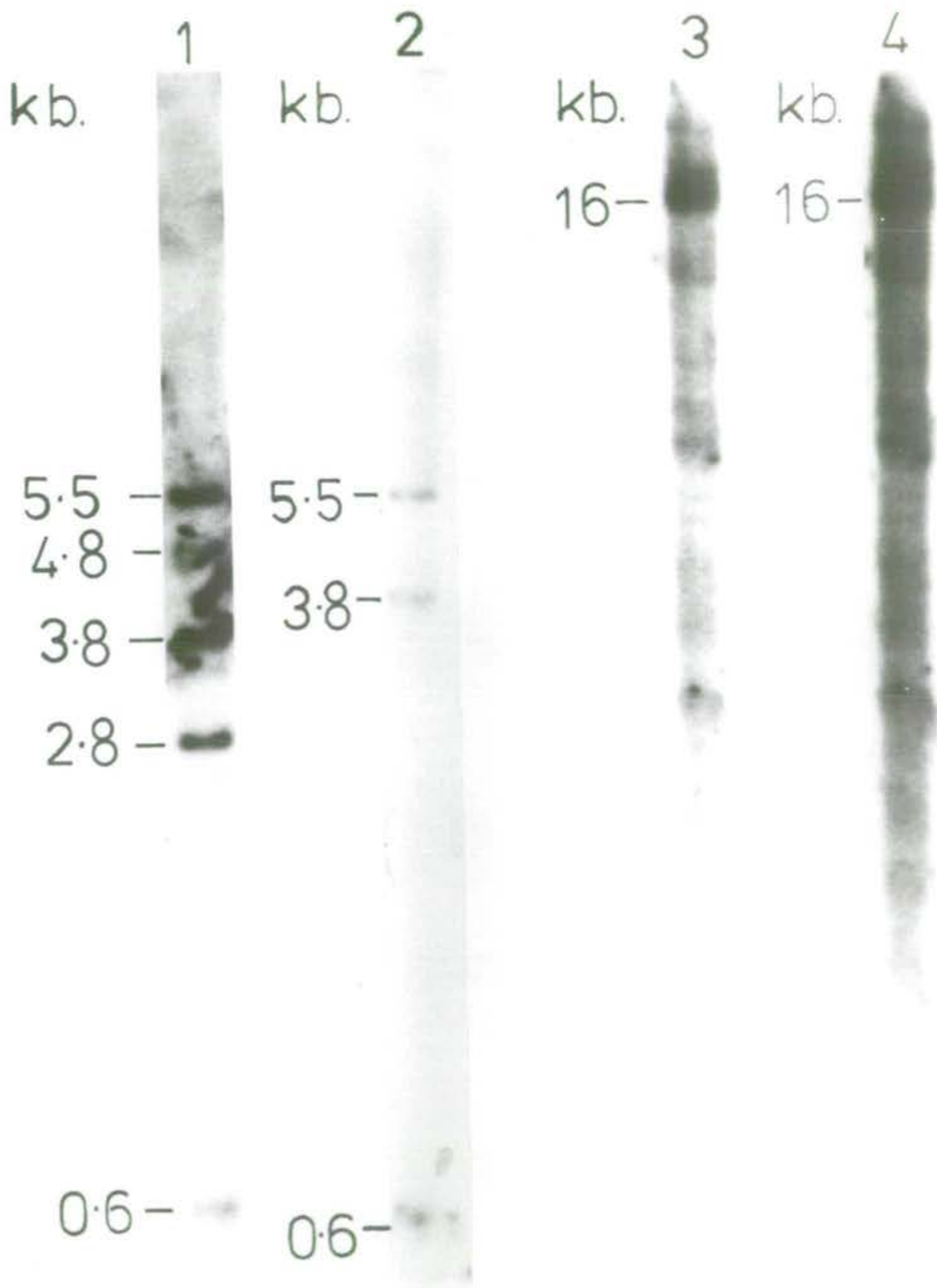
McKeown, et al. (1978) reported that when different actin genes were used as probes, differing degrees of homology were observed, due to heterogeneity between actin genes. To test this possibility in P.falciparum, a HindIII digest of genomic DNA was hybridised to the plasmid Eco 13.4.19 (track 2). Now homology is only detected to fragments of 5.5, 3.8 and 0.6 kb. Although homology might be detected to other fragments with a longer exposure of the autoradiograph, the 2.8 kb fragment appears to specifically hybridise to pcDdBl. It is noteworthy that the extra fragments are identified by the cDNA actin clone (pcDdBl), and it therefore seems unlikely that the extra bands are identified due to some spurious homology. When the genomic DNA is digested with EcoRl and hybridised to the plasmid Eco 13.4.9, the majority of the homology is to a 16 kb fragment (track 3), although some faint homology is detected to other fragments after a long exposure (track 4).

These data indicate that there are sequences in the genome of P.falciparum homologous to Dictyostelium actin genes, and a number of tentative conclusions can be drawn from the hybridisation patterns. First, the differential hybridisation pattern suggest that P.falciparum has more than one actin-like sequence, and that these are different. Interestingly there is a HindIII site near the 5' end of all actin genes so far examined (McKeown, et al. 1978; Firtel, et al. 1979) The identification of 5 HindIII fragments raises the possibility that there are as many as 5 actin genes in P.falciparum. The majority of homology displayed by the plasmid Eco 13.4.9. to the 16 kb EcoRl fragment could be interpreted as an indication that some of the putative

Fig. 5.2

Autoradiographs of K1 genomic DNA, hybridised with pcDdB1 (track 1), Eco 13.4.9 (tracks 2, 3 and 4). The DNA was digested with HindIII (tracks 1 and 2), and EcoR1 (tracks 3 and 4). The sizes of the fragments are given in kb. Note that Eco 13.4.9 (track 2) hybridises to only 3 of the 5 fragments identified by pcDdB1 (track 1). Note also that the majority of homology shown to the plasmid Eco 13.4.9 is located on a 16 kb EcoR1 fragment (track 3), although some homology is detected to other smaller fragments after a long exposure (track 4).

FIG 5.2



actin genes are clustered. Actin genes are known to be clustered in Dictyostelium, sea urchins and nematodes (McKeowan, et al. 1978); Durica, et al. 1980; Files, et al. 1983); the dispersed actin genes in Drosophila appear to be exceptional in this respect (Fyrberg, et al. 1980; Tobin et al. 1980).

To analyse further the identified putative actin sequences, clones homologous to the D.discoideum actin probes were isolated from genomic and cDNA libraries of P.falciparum. These studies are detailed in sections 5.2.2 and 5.2.3

5.2.2 Isolation of putative actin genes from a HindIII genomic library

The HindIII fragments identified in the genome of the isolate K1, were sought in the genomic library constructed from the Gambian isolate HG13, cloned into phage lambda (Goman et al, 1982). The λPFH8 library was screened and out of the 4000 individual clones two were homologous to pcDdB1. The two phage recombinants identified by pcDdB1 were purified, and their homology to D.discoideum actin checked against the plasmid Eco. 13.4.9 (Fig. 5.3). Note that Eco 13.4.9 also hybridises to a nematode actin sequence cloned into phage λ. The nematode actin sequence (a kind gift of Dr. M. Salvato) is a BamH1 genomic fragment (ca. 10 kb) inserted into the phage vector 1059 (Korn, et al. 1980).

The two putative P.falciparum actin clones were amplified, the DNA extracted and digested with HindIII. Both were found to contain fragments of 5.5, 4.8, and 0.6 kg, and an example of such a digest is shown in Fig. 5.4 (track 2). This is a particularly satisfying result as fragments of these sizes were identified in the genome of the Thai isolate K1, and the clones are derived from the

Fig. 5.3

Actin gene containing recombinant phage hybridised with Eco 13.4.9. Homology is observed to the nematode recombinant (A), and the two *P.falciparum* recombinants (B and C).

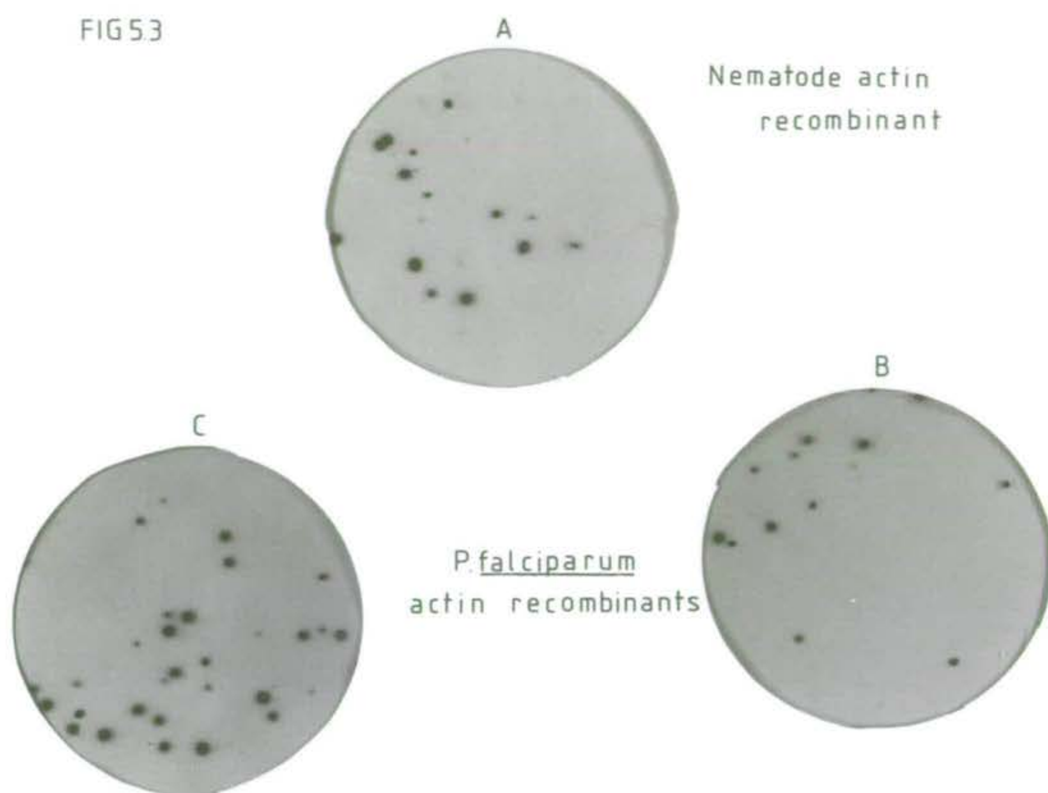
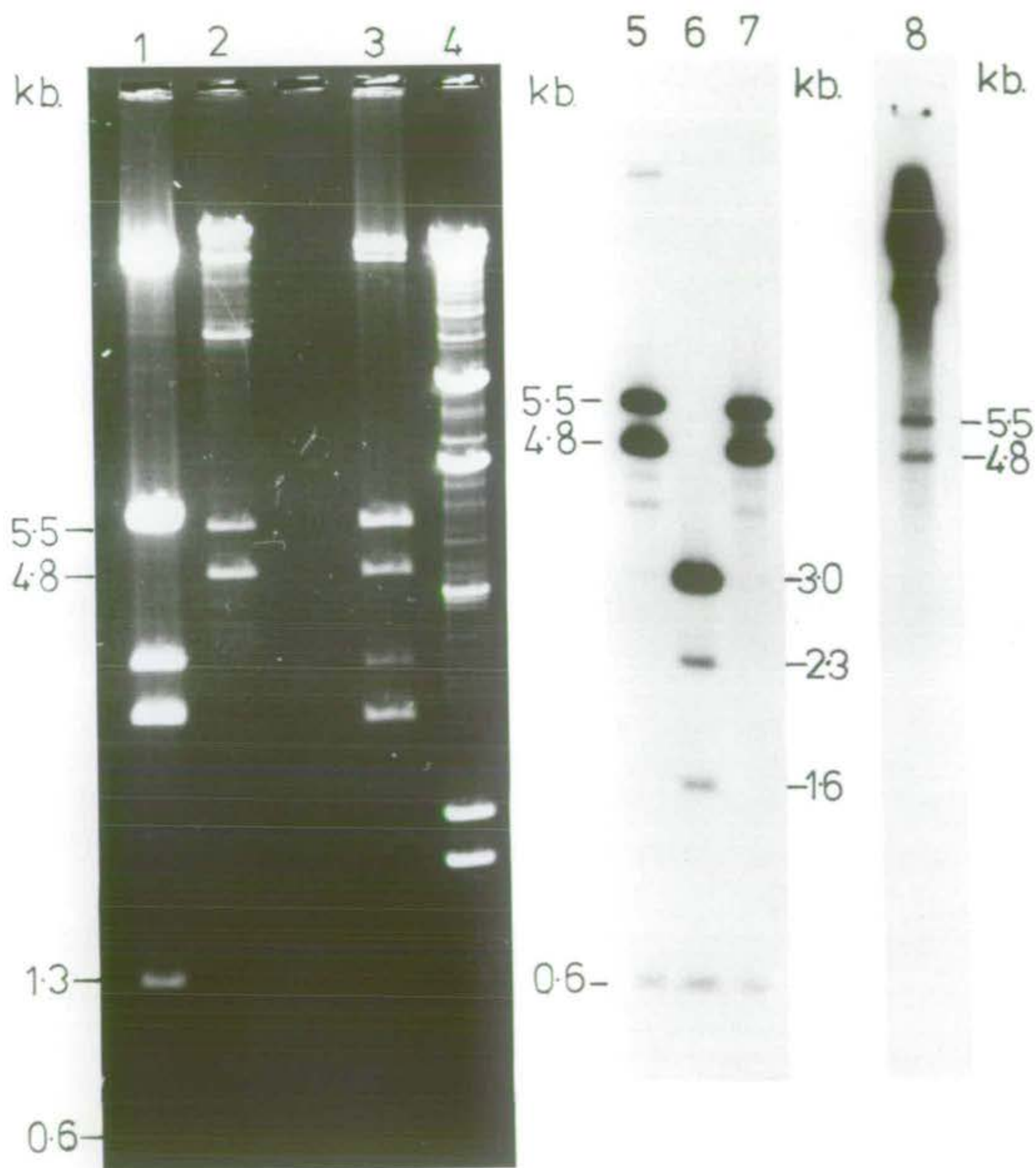


FIG 5.3
A
Nematode actin
recombinant
B
C
P. falciparum
actin recombinants

Fig. 5.4

Ethidium bromide stained gel and autoradiograph of restriction digests of NM788 and λPFH8 actl. The actin recombinant contains 3 HindIII fragments of 5.5, 4.8 and 0.6 kb homologous to the pcDdB1 probe (tracks 2 and 5). The insert lacks EcoR1 and BglII sites, since a HindIII/EcoR1 and a HindIII/BglII double digest (tracks 3 and 7) yields the same 5.5, 4.8 and 0.6 kb fragments. An EcoR1/HindIII digest of NM788 is included for comparison (track 1). Note the ability to see the 1.3 kb vector fragment, and the 0.6 kb actin fragment (tracks 2 and 3) has been lost upon photography, although they were clear on original negative. The 0.6 kb fragment is clearly visible in tracks 5, 6 and 7. A HindIII/SalI digest generates four fragments of 3.0, 2.3, 1.6 and 0.6 kb homologous to pcDdB1 (track 6), the 3.0 kb fragment showing the strongest homology. When λPFH8actl is hybridised to a nematode actin phage recombinant, both the 5.5 and 4.8 kb fragments display homology. The strong hybridisation to the high molecular weight fragment is due to homology between vectors. At this exposure the 0.6 kb fragment is not visible. Size markers were HindIII fragment of λ^{+} DNA (track 4). Note some homology is seen in tracks 5 and 7 to fragments which are the result of partial restriction.

FIG 5.4



Gambian isolate HG13 (Butcher, 1981). All 3 fragments are homologous to pcDdB1 (track 5), and in total comprise ca. 10 kb of genomic DNA. Note that some partials homologous to the probe are observed in tracks 5 and 7. The isolation of a single recombinant phage of 3 HindIII fragments all homologous to pcDdB1, which together account for 10 kb of the genome is consistent with the view that some of the putative actin genes are clustered in P.falciparum.

The Plasmodium DNA insert is digested by SalI and SalI/HindIII fragments of 3.0, 2.3, 1.6, 0.6 kb are found to be homologous to pcDdB1 (track 6). Since the insert is approximately 10 kb, the SalI/HindIII digest reveals that only 7 kb codes for actin and furthermore the majority of the actin sequence is encoded by the 3.0 kb fragment, as it shows the strongest homology to pcDdB1. Note that the insert does not contain sites for EcoR1 (track 3) or BglII (track 7).

Finally, the putative P.falciparum actin clone, λPFH8 act1 also shows homology to the Nematode actin gene cloned in lambda (track 8). Note that the small 0.6 kb HindIII fragment is not detectable at the exposure presented, however both the 5.5 and 4.8 kb can be seen to be homologous. Note also the strong homology observed to the high molecular weight fragments is because the probe is also a phage lambda recombinant. Taken together with the observation that λPFH8 act1 is homologous to the two Dictyostelium probes, it seems reasonable to conclude that λPFH8 act1 contains P.falciparum actin genes.

5.2.3. Identification of a mRNA transcript, and isolation from a P.falciparum cDNA library, of sequences homologous to a cloned Dictyostelium actin gene

Of the 17 actin genes identified in the genome of Dictyostelium,

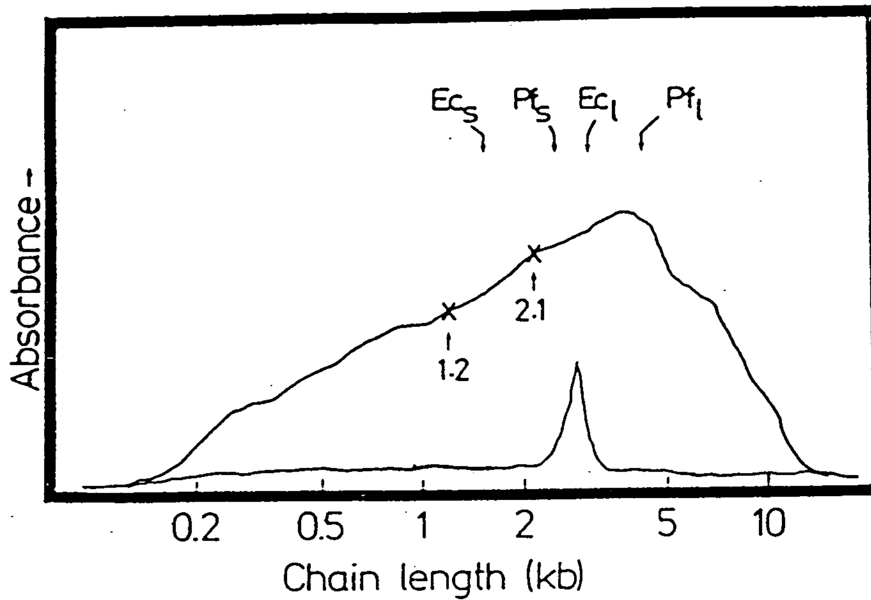
only six are known to be expressed (McKeown, _et al._ 1981). In order to discover whether the actin-like sequences cloned from _P.falciparum_ are expressed, I sought a transcript homologous to the _Dictyostelium_ actin probe, in _P.falciparum_ polyA$^+$ mRNA.

To identify and estimate the size of mRNA homologous to _Dictyostelium_ actin genes, total polyA$^+$ RNA extracted from an asynchronous culture of the asexual blood forms of _P.falciparum_ (Methods, 2.3.b) was fractionated on an agarose/formaldehyde gel (2.18). The RNA was transferred to nitrocellulose, and hybridised against the plasmid Eco 13.4.9. A single band of RNA with size _ca._ 2.8 kb hybridised to the probe (Fig. 5.5, lower trace). At 2.8 kb this is larger than most actin mRNAs, although a 2.5 kb actin transcript has been reported for _Drosophila_ (Zulaf, _et al._ 1981); however most are generally 1.2-2.0 kb (Firtel, 1981).

Having established that the 'actin' sequences in _P.falciparum_ are expressed, cDNA copies of the 'actin' mRNA were sought in a cDNA library of _P.falciparum_. The construction of this cDNA library has been described (Hyde, _et al._ 1983). Purified polyA$^+$ RNA from an asynchronous culture of K1 parasites, was copied into double stranded cDNA. To this _EcoR1_ linkers were added, and the cDNA fragments cloned into the _EcoR1_ site of the vector NM1149. The recombinant phage were packaged _in vitro_, and plated on indicator bacteria. All the plaques were harvested, and stored at 4°C, this step resulted in an amplification of 2×10^6.

A proportion of this amplified cDNA library (_ca._ 25,000 plaques) was screened with radio-labelled pcDdP1 and Eco.13.4.9. Twelve recombinant phage were identified in this way. Phage stocks were made

FIG 5.5



Upper trace: Size distribution of total P.falciparum polyA$^+$ from K1 on 0.8% denaturing gel showing number average (1.2 kb) and weight average (2.1 kb) chain lengths. Markers are the small and large rRNA chains from E.coli (Ec_s = 1.54 kb, Ec_l = 3.24 kb) and those from P.falciparum (Pf_s = 2.27 kb, Pf_l = 4.34 kb, Hyde, et al. 1981).

Lower trace: Total polyA$^+$ RNA from above probed with a Dictyostelium actin clone; scan of autoradiogram.

and the DNA extracted from 3 of the recombinants (Fig. 5.6 tracks 2, 3 and 4). All 3 contained a cDNA insert of 3.5 kb, which is homologous to pcDdB1 (tracks 5, 6 and 7). The discrepancy between the cDNA (3.5 kb) and transcript (2.8 kb) is difficult to explain. A number of possibilities exist, though none of these has been tested at this time. One possibility is that the estimation for the size of the transcript is low due to the RNA not running true on the gel. Although a denaturing gel was used, the transcript may still have possessed secondary structure, and this imparted an anomalous mobility to the RNA. Another possibility is that the cDNA derives from a rare 3.5 kb message, which is more readily cloned, or that the discrepancy in size results from a cloning artifact. In conclusion, the extreme possibility exists that the mRNA and the cDNA clones identified by the Dictyostelium probe are unrelated.

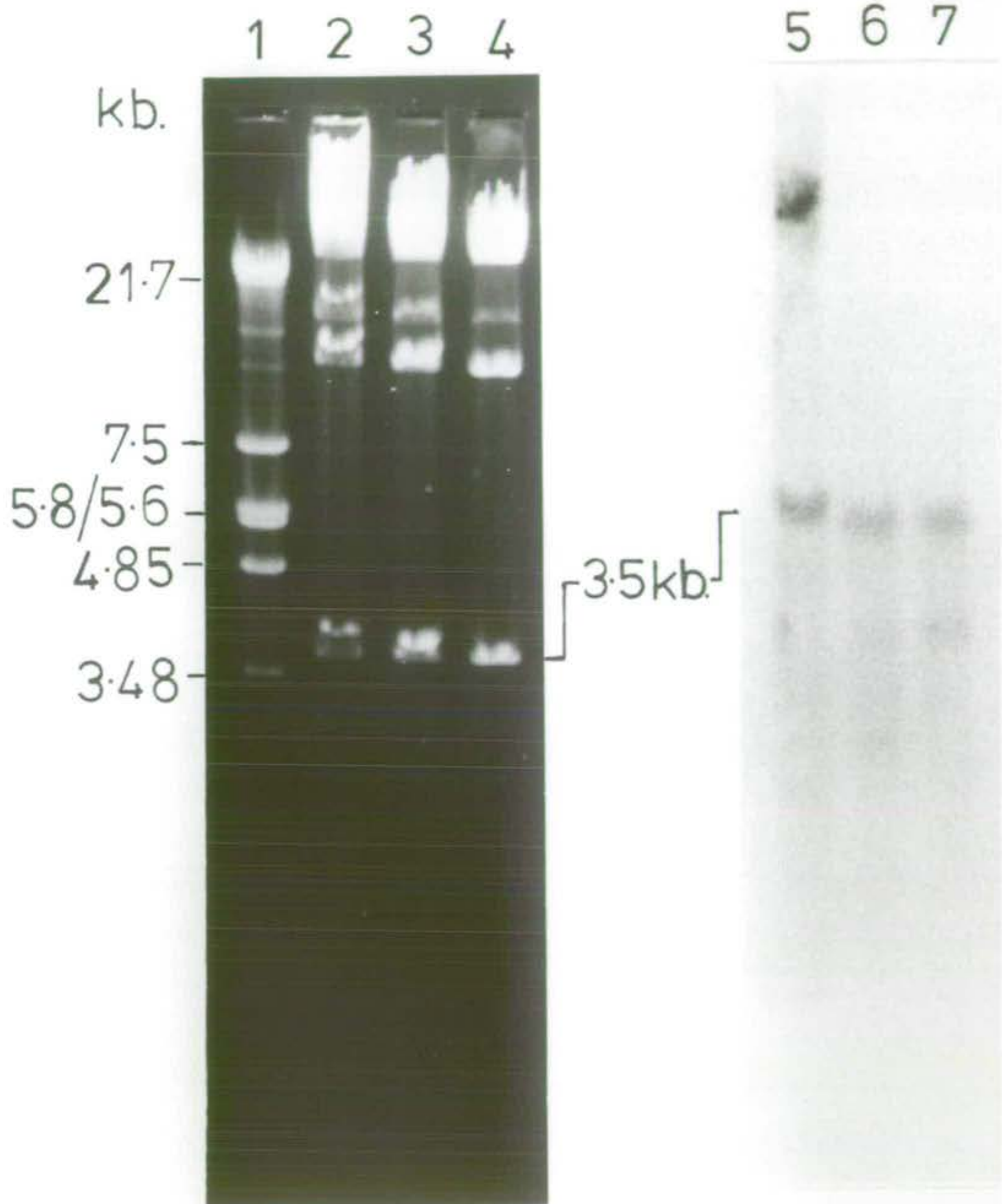
Confirmation that a cDNA clone actually codes for actin, is usually obtained either by nucleic acid sequencing or hybrid selection and translation to give a protein recognised as actin (Gallwitz, et al. 1980; Durica, et al. 1980; Zulaf, et al. 1981; Files, et al. 1983). The latter approach was initiated here for the following reason. In this laboratory the major emphasis is the recognition and cloning of protective antigen genes, as a prelude to the development of a vaccine against malaria. An integral step in this process is the identification of particular antigen genes by hybrid selection. The hybrid selection of actin mRNA by the putative DNA copy provides a model system, and the preliminary experiments towards this goal are presented below.

Rat antibodies raised against yeast actin and tubulin have been shown to react with mammalian actin and tubulin (Chang et al. 1981,

Fig. 5.6

Ethidium bromide stained gel and autoradiograph of EcoRl digests of P.falciparum cDNA clones homologous to pcDdBl. Track 1 shows EcoRl size marker fragments of λ^{+}DNA. Tracks 2, 3 and 4 show EcoRl digests of P.falciparum cDNA clones. Tracks 5 , 6 7 show above digested DNA hybridised with pcDdBl. Note P.falciparum insert specifically hybridises with pcDdBl.

FIG 5.6



and John Kilmartin, pers. comm.). It is believed that the conserved nature of actin (Vanderkerckhove, et al. 1978 and 1980) allows some actins from disparate organisms to share a common epitope, however the cross reaction is not general. As the rat α-yeast actin antibodies (a kind gift of Dr. J. Kilmartin) is known to cross-react, it was used here in an attempt to identify P.falciparum actin. To date actin has not been identified or purified in Plasmodia.

Total parasite proteins extracted by detergent (NP40) from an asynchronous population of K1 parasites, were fractionated on a SDS-polyacrylamide gel (Methods, 2.19, and Fig. 5.7, track 3). Included for comparison were proteins extracted from uninfected red blood cells (track 2) and yeast (track 4). The proteins were transferred to nitro-cellulose and screened with rat α-yeast actin antibody. The precipitated antibody was recognised by the ELISA test (Methods 2.27). Yeast actin is identified (track 7) of the correct molecular weight (ca. 42000). No cross reaction to P.falciparum (track 6) or red blood cell (track 5) was observed, even when affinity purified rat α-yeast actin antibody was used (data not shown).

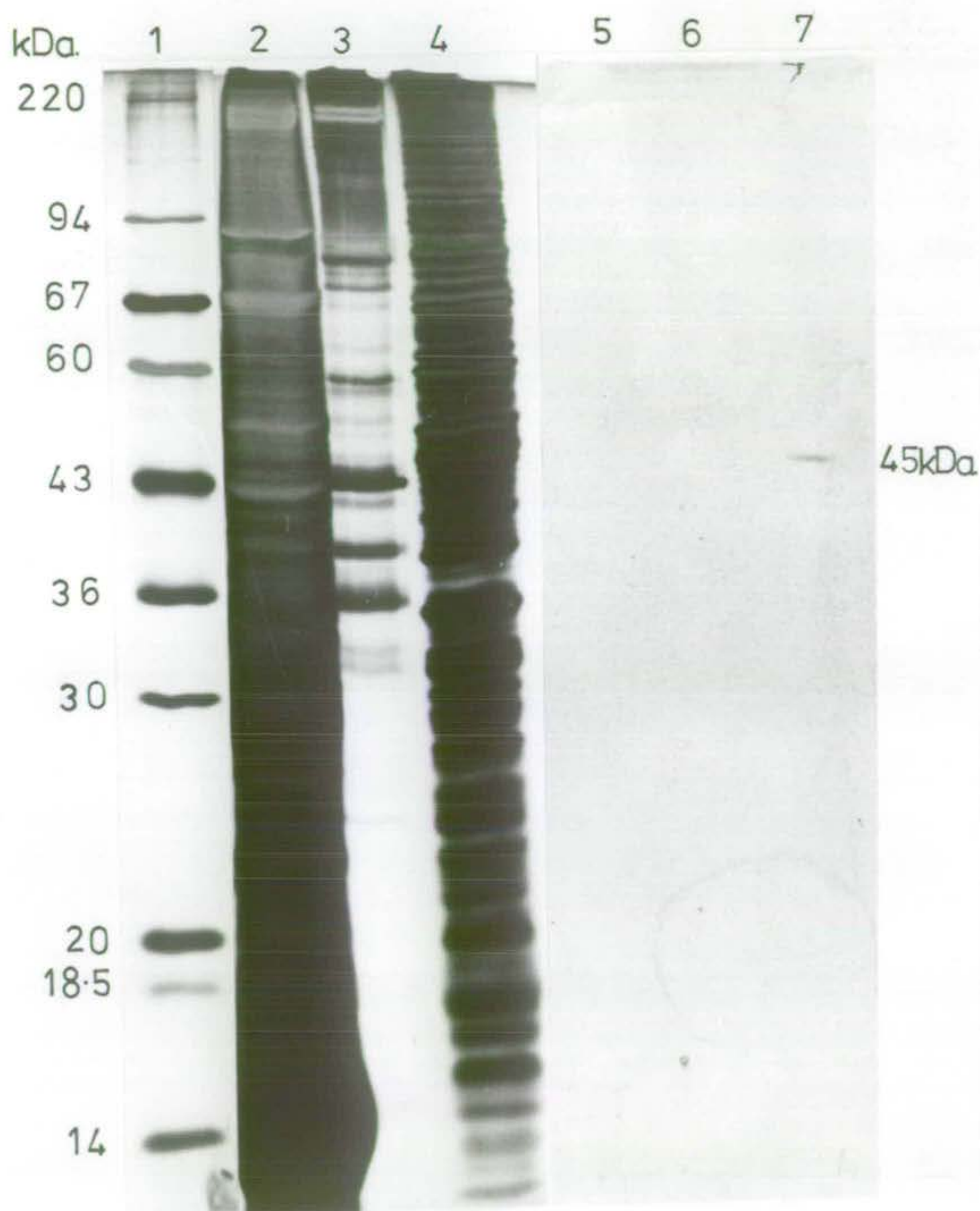
Although I failed to detect any cross-reaction between yeast-actin antibody and P.falciparum, I do not think this approach has been exhausted. Antibody raised against a different actin may well cross-react, and a number of different antisera could be tried; antisera to cytoskeletal microfilaments and chicken muscle actin are now commercially available (Miles Scientific).

Clearly a more systematic approach is to purify actin from P.falciparum, and raise anti-P.falciparum actin antibodies. Actin is usually purified by a procedure involving chromatography on DEAE-cellulose,

Fig. 5.7

Silver stained SDS-polyacrylamide (Laemmli) gel and western blot. Total proteins extracted from yeast (track 4), *P.falciparum* (track 3), and uninfected red blood cells (track 2). Molecular weight size marker proteins shown in track 1. Above proteins were transferred to nitrocellulose, and probed with rat α-yeast actin antibodies. The precipitated antigen recognised by the ELISA assay (see Methods 2.27). Note no actin precipitated from *P.falciparum* (track 6), or uninfected red blood cells (track 5). Note also that actin of the correct molecular weight (*ca.* 45000) precipitated from total yeast proteins (track 7).

FIG5.7



polymerization and depolymerization, and gel filtration of monomeric G-actin (Korn, 1978). However, a more rapid approach is to take advantage of actins inhibitory binding to DNAse I; a DNAse I-agarose affinity column can be used to purify actin (Lazarides, et al. 1974). The availability of P.falciparum actin would, as mentioned, allow antisera to be raised. Furthermore more determination of its molecular weight would indicate the minimum size of the nucleotide coding region; for example how much of the putative 2.8 kb transcript actually codes for protein, or is this unusually large actin mRNA due to untranslated sequences at the messenger termini? In this context a 500 bp 3' untranslated region has been reported for a human cytoplasmic actin (Hanukoglu, et al. 1983).

5.3 Conclusions

Hybridisation analysis using two Dictyostelium discoideum actin gene containing recombinants (pcDdB1 and Eco 13.4.9) indicate the presence of homologous sequences in the genome of P.falciparum. The different patterns of hybridisation towards the two probes suggests that there may be more than one (and possibly 5) 'actin' genes, and that they are heterogeneous. Furthermore, some of these putative genes could be clustered, and the isolation of a single phage recombinant containing 3 of the 5 HindIII fragments identified by the probes lends support to this view.

The identification of a specific transcript by Eco 13.4.9 indicates that some of these putative genes are expressed, and a cDNA clone homologous to both probes has been isolated. Hybrid selection of a transcript specific to the putative cDNA actin gene is in progress (J. Hyde, pers. comm.), and translation of said message will confirm whether the cDNA clone is an actin gene.

As mentioned previously nucleotide sequence analysis would also confirm the cDNA as an actin clone. The availability of the sequence would also reveal whether *P.falciparum* actin genes contain an intervening sequence(s). With the exception of *Dictyostelium*, all actin genes sequenced so far posses intervening sequences, the position and number of which appear highly variable (Fyrberg, *et al*. 1981; Sanchez *et al*. 1983; Ng and Abelson, 1980; Schuler, *et al*. 1981; Firtel, 1981). Furthermore the nucleotide sequence may reveal the nature of the discrepancy in size between the cDNA and mRNA identified by the probes.

During the life cycle of *P.falciparum* certain stages are highly motile, and at one point the microgamete suddenly produces flagella; whilst at other stages the parasite is sessile (oocyst). It is not unreasonable therefore to expect actin gene expression in *Plasmodia* to be developmentally regulated, and the cDNA actin clone may permit the investigation of developmentally regulated gene expression. There is a precedent as actin genes are developmentally expressed in *Dictyostelium*, *Drosophila* and sea urchin (McKeown, *et al*. 1981; Sanchez, *et al*. 1983; Merlino, *et al*. 1980), and this was determined using the untranslated portions of the cDNA clones as probes for mRNAs expressed at the different developmental stages.

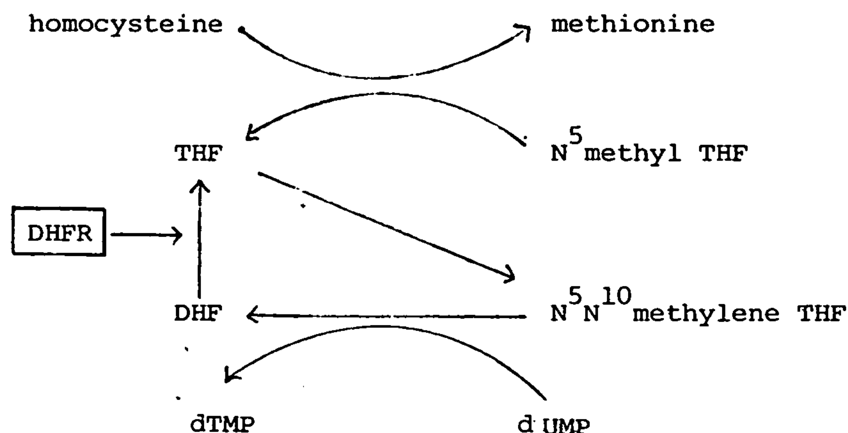
CHAPTER 6

GENERAL DISCUSSION

The work presented in this thesis details the initial chatacterisation of the genome of *P.falciparum*, as characterised by the analysis of three particular classes of DNA. The first class of DNA represented by λPFH8rep20 is highly repeated in the genome, and apparently is not expressed (or is expressed at a very low level), since no homologous mRNA species could be detected on ~~western~~ northern blots (data not shown). The second class of sequence, namely the rRNA genes, have been shown here to be only moderately repetitive (*ca.* 8 gene copies). The rRNA genes are conserved in their sequence (when compared to other eukaryotic rRNA genes), and are transcribed into a product (rRNA which is readily obtained from parasites cultured *in vitro*). Finally the third class, the putative actin genes, are also conserved in their coding sequence, but moreover are translated into a protein product.

The latter two, the rRNA and actin genes, are both conserved in their coding sequence. This was crucial in their isolation, enabling recombinant phage containing them to be identified by cross homology to cloned genes from another species. This constraint (of conservation in sequence) upon the library screening with heterologous probes was illustrated by the failure to detect the gene coding for dihydrofolate reductase (DHFR), using as a probe the DHFR gene cloned from mouse. The DHFR genes are not known to be conserved in their sequence.

The enzyme DHFR is involved in folate metabolism, it reduces dihydrofolate to tetrahydrofolate (see below), and is the site at which many of the antifol drugs act (for review of DHFR in *Plasmodia*, see Sherman, 1979).



The substrate affinity of the plasmodial DHFR is 2-3 times greater than host cell DHFR, and is probably the reason why the antifols (pyrimethamine, trimethoprin, cycloguaril, dihydrotriazimes) are effective as antimalarials.

Methotrexate (MTX) is a specific inhibitor of DHFR, and in rodents and Leishmania tropica, MTX resistance is due to an amplification of the DHFR gene (Kaufman, et al. 1979; Coderre et al. 1983). Since the P.falciparum isolate Kl is resistant to pyrimethamine (Thaithong and Beale, 1981), the Kl DHFR gene was sought using as a probe the cDNA DHFR gene from mouse (Chang, et al. 1978). I was unable to detect any homology between the mouse DHFR sequences and those of P.falciparum, even under conditions where the single DHFR gene was detected in MTX sensitive mice. Using the same cDNA probe Schimke and his co-workers were unable to detect DHFR sequences in L.tropica (Coderre et al. 1983), and the L.tropica DHFR gene does not cross hybridise with malaria DHFR sequences (R. Santi, pers. comm.). Thus illustrating the requirement for conservation in sequence when using heterologous probes.

In *L.tropica* amplified DHFR genes were detected in MTX resistant cells, as a prominant band on an ethidium bromide stained agarose gel. Comparing Kl (pyrimethamine resistant) to Tak 9 (sensitive) genomic DNA on ethidium bromide stained gels, failed to reveal any prominant species even when the DNA was extracted from parasites in the presence of challenge; this is also true for the mRNA (J. Hyde, pers. comm.). The naturally occurring pyrimethamine resistant strains of malaria, appear to be stable mutants which do not revert in the absence of challenge, and the failure to detect any amplification of their DNA, implies that resistance may be due to single (or multiple) lesion, rather than to amplification of DHFR sequences (for review of anti-malarial drug resistance, see Peters 1982).

The data presented in Chapter 3, combined with that of Thaithong *et al*. (1983) demonstrates that there are marked differences between strains of *P.falciparum*. They emphasise that an isolate can be a mixture of several genetically distinct strains, and cannot therefore be considered as genetically pure. This may be relevant to this laboratory, where the *P.falciparum* laboratory standard is the uncloned isolate Kl.

The importance of working with a genetically defined strain cannot be overstressed, and failure to do so has led to some errors in interpretation of data. In the late 1960's Peters and his co-workers reported the selection from the *P.berghei* isolate NK65, of a number of chloroquin resistant lines (NS lines), an example of such is Ramkaran and Peters (1969). It was later shown (by Carter and co-workers in Edinburgh), by isoenzyme typing, that the NS lines were in fact *P.yoelii*, which is naturally resistant to chloroquin, and Peters,

et al. (1978) had to reinterpret their previous data.

Using the same technique of isoenzyme typing D.Walliker and his co-workers have recently shown that the non-infective *P.berghei* NK65 line of Dore, *et al*. (1980, 1983; and Birago, *et al*. 1982) contains *P.yoelii* (D. Walliker pers. comm.). It appears that mechanical passage of NK65 has led to the selection of *P.yoelii* out of the mixed isolate, and in the light of this finding the relationship between infectivity and repetitive DNA content hypothesised by Dore, *et al*. may have to be re-examined.

P.berghei and *P.yoelii* are morphologically very similar, and it is difficult to distinguish between them on that criterion alone. However, this is not the case for the human malarias, for example, *P.falciparum* has characteristically crescent shaped gametocytes (Carter and Miller, 1979). Moreover the monoclonal antibody and *in vitro* translation of mRNA studies to date, show K1 (if composed of more than one strain) to be a stable isolate even over long periods of *in vitro* culture (Hall, *et al*. 1983; Hyde, *et al*. 1983).

In addition to heterologous probing, there are two further methods of identifying a gene of interest from a number of clones (library). Firstly, when the mRNA of the gene in question is relatively abundant, and the gene product is known, the cDNA copy can be identified by hybrid selection and *in vitro* translation of said message. This protocol was used by Merlino *et al*. (1980) to clone to sea urchin actin gene. In sea urchins 10% of the polyA$^+$ mRNA codes for actin. The product of the *in vitro* translation co-migrated with both rabbit muscle actin, and actin purified from sea urchin. However, in *P.falciparum* actin has yet to be purified, nor is the

mRNA known to be abundant. Moreover the gene could be isolated by homology to other actin sequences and for these reasons this approach was not used here.

Secondly, the gene coding for a particular protein can be identified providing that antisera to the protein is available, and the cloned gene expresses sufficient antigen. The gene coding for the sporozoite surface antigen of *P.knowlesi* was identified in this way (Ellis *et al*. 1983). Kemp *et al*. (1983) have recently reported the construction of an expression cDNA library from *P.falciparum*. Using human sera from a malarious area, a number of clones were identified which presumably may code for clinically important antigens. An expression cDNA library is also being screened with immune sera in this laboratory, and the isolation and characterisation of an antigen gene now seems more probable than possible.

However, the data presented here on the rRNA and 'actin' genes represents the first analysis of genes from the human malaria parasite *P.falciparum*. An understanding of the basic biology of the parasite will surely be crucial in long term efforts aimed at its eradication, and this work is the first step in elucidating the genomic organisation of *P.falciparum*. The results from the λPFH8 rep 20 analysis implies that the genome of *P.falciparum* may be very plastic, and by recombination and or transposition capable of a high rate of change (mutation). The notably low copy number for the rRNA genes, and the large 'actin' transcript size, are indicative of the unusual life cycle of *P.falciparum* (from man to mosquito) and may be reflected in the genomic organisation, and these could be the first of many surprises to come.

BIBLIOGRAPHY

ADLER, J. (1975). Chemotaxis in bacteria. Ann.Rev.Biochem. 44: 341-356.

ARNHEIM, N. and KUEHN, M. (1979). The genetic behaviour of a cloned mouse ribosomal DNA segment mimics mouse ribosomal gene evolution. J.Mol.Biol. 134: 743-765.

ARNHEIM, N. and SOUTHERN, E.M. (1977). Heterogeneity of the ribosomal genes in mice and mean. Cell 11: 363-370.

BAHR, G.F. and MIKEL, U. (1972). The arrangement of DNA in the nucleus of rodent malaria parasites. Proc.Helminthol.Soc. Wash. 39: 361-372.

BAKKEN, A., MORGAN, G., SOLLNER-WEBB, B., ROAN, J., BUSBY, S. and REEDER, R.H. (1982). Mapping of transcription initiation and termination signals on Xenopus laevis rDNA. Proc.Natl.Acad. Sci. USA 79: 56-60.

BARKER, D.C. and BUTCHER, J. (1983). The use of DNA probes in the identification of Leishmanias: discrimination between isolates of the Leishmania mexicana and L. braziliensis complexes. Trans.R.Soc.Trop.Med.Hyg. 77: 285-297.

BENDER, W., DAVIDSON, N., KINDLE, K.L., TAYLOR, W.C., SILVERMAN, M. and FIRTEL, R.A. (1978). The structure of M6, a recombinant plasmid containing Dictyostelum DNA homologous to actin messenger RNA. Cell 15: 779-788.

BENTON, W.D. and DAVIS, R.W. (1977). Screening λgt recombinant clones by hybridisation to single plaques in situ. Science 196: 180-182.

BIRAGO, C., BUCCI, A., DORE, E., FRONTALI, C. and ZENOBI, P. (1982). Mosquito infectivity is directly related to the proportion of repetitive DNA in P. bergei. Molecular and Biochemical Parasitology 6: 1-12.

BIRNBOIM, H.C. and DOLY, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acid Res. 1: 1512-1523.

BLACKBURN, E.H. and GALL, J.G. (1978). A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. J.Mol.Biol. 120: 33-53.

BOLIVAR, F., RODRIGEZ, R.L., GREENE, P.J., BETLACK, M.C., HEYNEKER, H.L., BOYER, H.W., CROSA, J.H. and FALKOW, S. (1977). Construction and characterisation of new cloning vehicles. II. A multipurpose cloning system. Gene 2: 95-113.

BONE, N., GIBSON, T., GOMAN, M., HYDE, J.E., LANGSLEY, G.W., SCAIFE, J.G., WALLIKER, D., YANKOVSKY, N.K. and ZOLG, J.W. (1983). Investigation of the DNA of human malaria parasite P. falciparum by in vitro cloning in phage λ. in The Molecular Biology of Parasites, Eds. J. Guardida, L. Luzzatto, W. Trager. Raven Press, New York.

BORST, P. and CROSS, G.A.M. (1982). Molecular basis for Trypanosome antigenic variation. Cell 29: 291-303.

BORST, P., FASE-FOWLER, F., FRASCH, A.C.C., HOEIJMAKERS, J.H.J. and WEIJERS, P.J. (1980). Characterisation of DNA from Trypanosoma brucei and related trypanosomes by restriction endonuclease digestion. Molecular and Biochemical Parasitology 1: 221-246.

BOSCHEK, B.C., JOCKUSCH, B.M., FRIIS, R.R., BACK, R., GRUNDMANN, E. and BAUER, H. (1981). Early changes in the distribution and organisation of microfilament proteins during cell transformation. Cell 24: 175-184.

BOTCHAN, P., REEDER, R.H. and DAWID, I.B. (1977). Restriction analysis of the nontranscribed spacers of Xenopus laevis ribosomal DNA. Cell 11, 599-607.

BRAY, R.S. and GARNHAM, P.C.C. (1982). The life-cycle of primate malaria parasites. British Medical Bulletin 38: 117-112.

BREATHNACH, R. and CHAMBON, P. (1981). Organisation and expression of eukaryotic split genes coding for proteins. Ann.Rev.Biochem. 50: 349-383.

BRITTEN, R.J. and DAVIDSON, E.H. (1969). Gene regulation for higher cells: a theory. Science 165: 349-357.

BRITTEN, R.J. and KOHNE, D.E. (1968). Repeated sequences in DNA. Science 161: 529-540.

BROWN, G.V., ANDERS, R.F., MITCHELL, G.F. and HEYWOOD, P.F. (1982). Target antigens of purified human immunoglobulins which inhibit growth of Plasmodium falciparum in vitro. Nature 297: 591-593.

BROWN, K.N. and BROWN, I.N. (1965). Immunity to malaria: antigenic variation in chronic infections of Plasmodium knowlesi. Nature 208: 1286-1288.

BRUCE-CHWATT, L.B. (1980). Chemotherapy and Chemoprophylaxis. in Essential Malariology. William Heinemann Medical Books Ltd.

BUTCHER, G.A. (1981). A comparison of static thin layer and suspension cultures for the maintenance in vitro of P. falciparum. Ann.Trop.Med.Parasitol. 75: 7-17.

CAMERON, J.R., LOH, E.Y. and DAVIS, R.W. (1979). Evidence for transposition of dispersed repetitive DNA families in yeast. Cell 16: 739-751.

CAMPBELL, G.R., LITTAU, V.C., MELERA, P.W., ALLFREY, V.G. and JOHNSON, E.M. (1979). Unique sequence arrangement of ribosomal genes in the palindromic rDNA molecule of *Physarum polyphalum*. Nucleic Acid Res. 6: 1433-1447.

CARTER, R. and MILLER, L.H. (1979). Evidence for environmental modulation of gametocytogenesis in *Plasmodium falciparum* in continuous culture. Bulletin of the World Health Organisation 57: 37-52.

CARTER, R. and WALLIKER, D. (1975). New observations on the malaria parasites of rodents of the Central Africa Republic. Annals of Tropical Medicine and Parasitology 69: 187-196.

CASTRO, C., HERNANDEZ, R. and CASTANEDA, M. (1981). *Trypanosoma cruzi* rRNA: internal break in the large-molecular-mass species, and number of genes. Molecular and Biochemical Parasitology 2: 219-233.

CECH, T.R. and RIO, D.C. (1979). Localisation of transcribed regions on extrachromosomal ribosomal RNA gene of *Tetrahymena thermophila* by R-loop mapping. Proc.Natl.Acad.Sci. USA 76: 5051-5055.

CECH, T., TANNER, N.K., TINOCO, Jr.I., WEIR, B.R., ZUKER, M. and PERLMAN, P.S. (1983). Secondary structure of the *Tetrahymena* ribosomal RNA intervening sequence. Structural homology with fungal mitochondrial intervening sequences. Proc.Natl.Acad. Sci. USA 80: 3906-3907.

CECH, T.R., ZAUG, A.J., GRABOWSKI, P.J. and BREHM, S.L. (1982). Transcription and splicing of the ribosomal RNA precursor of *Tetrahymena*. *in* The Cell Nucleus, Vol. X rDNA, Part A, pp. 171-201, Eds. Busch, H. and Rothblum, L.A.P.

CHAMBON, P. (1975). Eucaryotic nuclear RNA polymerase. Ann.Rev. Biochem. 44: 613-638.

CHANCE, M.L., MOMEN, H., WARHURST, D.C. and PETERS, W. (1978). The chemotherapy of rodent malaria, XXIX. DNA relationships with the subgenus *P. vinckei*. *in* Annals of Tropical Medicine and Parasitology 72: No. 1.

CHANG, A.C.Y., NUNBERG, J.H., KAUFMAN, R.J., ERLICH, H.A., SCHIMKE, R.T. and COHEN, S.N. (1978). Phenotypic expression in *E. coli* of a DNA sequence coding for mouse dehydrofolate reductase. Nature 275, 617-624.

CHANG, M.T., KILMARTIN, J. and DOVE, W.F. (1981). Monoclonal anti-bodies directed against yeast tubulin cross reaction to tubulin from pig and *Physarum polycephalum*, and subunit specificity. J. Cell. Biol. 91: A333.

CLARKE, M. and SPUDICH, J.A. (1977). Non-muscle contractile proteins: The role of actin and myosin in cell motility and shape determination. Ann.Rev.Biochem. 46: 797-822.

CLEVELAND, D.W., KIRSCHNER, M.W. and COWAN, N.J. (1978). Isolation of separate mRNAs for α- and β-tubulin and characterisation of the corresponding *in vitro* translation products. Cell 15: 1021-1031.

CLEVELAND, D.W., LOPATA, M.A., MacDONALD, R.J., RUTTER, W.J. and KIRSCHNER, M.W. (1980). Number and evolutionary conservation of α- and β-tubulin, and cytoplasmic β- and γ-actin genes using specific cloned cDNA probes. Cell 20: 95-105.

COCKBURN, A.F., NEWKIRK, M.J. and FIRTEL, R.A. (1976). Organisation of the ribosomal RNA genes of Dictyostelium discoideum: mapping of the nontranscribed spacer regions. Cell 9: 605-613.

COCKBURN, A.F., NEWKIRK, M.J. and FIRTEL, R.A. (1978). Dictyostelium rDNA consist of free palindromic dimers containing coding regions for 5S and 36S rRNA. Chromosoma 70: 19-29.

CODERRE, J.A., BEVERLEY, S.M., SCHIMKE, R.T. and SANTI, D.V. (1983). Overproduction of a bifunctional thymidylate synthetase-dihydrofolate reductase, and DNA amplification in methotrexate resistant Leishmania tropica. Proc.Nat.Acad.Sci. USA 80: 2132-2136.

CORDINGLEY, J.S. and TURNER, M.J. (1980). Small rRNAs of Trypanosomes. Molecular and Biochemical Parasitology 1: 91-96.

CORNELISSEN, A.W.C.A., OVERDULVE, J.P. and VAN DER PLOEG, M. (1983). Determination of the nuclear DNA of 5 Eucoccidian parasites, Isospora (Toxoplasma) Gondii, Sarcocystis cruzi, Eimeria tenella, E. acervulina and Plasmodium berghei with special reference to gamontogenesis and meiosis in I. (T.) gondii. Parasitology (in press).

CORY, S. and ADAMS, J.M. (1977). A very large repeating unit of mouse DNA containing 18S, 28S and 5.8S rRNA genes. Cell 11: 795-805.

COUDERC, J.L., SORBRIER, M.L., GIRAUD, G., BECKER, J.L. and DASTUGUE, B. (1983). Actin gene expression is modulated by ecdysterone in a *Drosophila* cell line. J.Mol.Biol. 164: 419-430.

CRAIG, S.W. and POLLARD, T.D. (1982). Actin-binding proteins. Review of many actin binding proteins. Trends.Biochem.Sci.88-92.

CRAIN, W.R., DAVIDSON, E.H. and BRITTEN, R.S. (1976). Contrasting arrangement in *Apis mellifera* (honey bee) and *Musca domestica* (house fly). Chromosoma 59: 1-12.

DAME, J.B. and McCUTCHAN, T.F. (1983a). The four ribosomal DNA units of the malaria parasite *Plasmodium berghei*. Journal of Biological Chem. 258: 6984-6990.

DAME, J.B. and McCUTCHAN, T.F. (1983b). Cloning and characterisation of a ribosomal gene from *Plasmodium berghei*. Molecular and Biochemical Parasitology 8: 263-279.

DA SILVEIRA, J.F. and MERCEREAU-PUIJALON, O. (1983). *Plasmodium chabaudi* antigens synthesised in an mRNA dependent cell-free translation system. Molecular and Biochemical Parasitology 7: 159-172.

DAVIDSON, E.H., HOUGH, B.R., AMENSON, C.S. and BRITTEN, R.J. (1973). General interspersion of repetitive with nonrepetitive sequence elements in the DNA of *Xenopus*. J.Mol.Biol. 77: 1-23.

DAVIDSON, E.H., GALAN, G.A., ANGERER, R.C. and BRITTEN, R.J. (1975). Comparative aspects of DNA organisation in metazoa. Chromosoma 51: 253-259.

DAVIDSON, G. (1982). Developments in malaria vector control. British Medical Bulletin 38: 201-206.

DAWID, I.B., WELLAUER, P.K. and LONG, E.O. (1978). Ribosomal DNA in _Drosophila melanogaster_. 1. Isolation and characterisation of cloned fragments. J.Mol.Biol. 126: 749-768.

DAWID, I.B., LONG, E.O., Di NOCERA, P.P. and PARDUE, M.L. (1981). Ribosomal insertion-like elements in _Drosophila melanogaster_ are interspersed with mobile sequences. Cell 25: 399-408.

DENHARDT, D.J., DRESDER, D. and RAY, D.S. (1978). The single stranded DNA phages. Cold Spring Harb. N.Y.

DIN, N., ENGBERG, J., KAFFENBERGER, W. and ECKERT, W.A. (1979). The intervening sequence in the 26S rRNA coding region of _T. thermophila_ is transcribed within the largest stable precursor for rRNA. Cell 18: 525-532.

DORE, E., BIRAGO, C., FRONTALI, C. and BATTAGLIA, P.A. (1980). Kinetic complexity and repetivity of _Plasmodium berghei_ DNA. Molecular and Biochemical Parasitology 1: 199-208.

DORE, E., FRONTALI, C., BIRAGO, C. and PICCI, L. (1983). Repetitive DNA organisation in _P. berghei_ genome, evidence of changes associated with the production of infective gametes. Molecular and Biochemical Parasitology (in press).

DURICA, D.S., SCHLOSS, J.A. and CRAIN, J.R.,W.R. (1980). Organization of actin gene sequences in the sea urchin: Molecular cloning of an intron-containing DNA sequence coding for a cytoplasmic actin. Proc.Natl.Acad.Sci. USA 77: 5683-5687.

ECKERT, W.A., KAFFENBERGER, W., KROHNE, G. and FRANKE, W.W. (1978). Introduction of hidden breaks during rRNA maturation and ageing in _Tetrahymena pyriformis_. Eur.J.Biochem. 87: 607-616.

ELLIS, J., OZAKI, L.S., GWADZ, R., COCHRANE, A.H., NUSSENZWEIG, V., NUSSENZWEIG, R.S. and GODSON, G.N. (1983). Cloning and expression in E. coli of the malarial sporozoite surface antigen gene from Plasmodium knowlesi. Nature 302: 536-538.

ENDOW, S.A. and GLOVER, D.A. (1979). Differential replication of ribosomal gene repeats in polytene nuclei of Drosophila. Cell 17: 597-605.

ENGBERG, J., ANDERSSON, P., LEICK, V. and COLLINS, J. (1976). Free ribosomal DNA molecules from Tetrahymena pyriformis GL are grant palindromes. J.Mol.Biol. 104: 455-470.

FEDOROFF, N.V. (1979). Review on spacers. Cell 16: 697-710.

FILES, J.G., CARR, S. and HIRSH, D. (1983). Actin gene family of Caenorhabditis elegans. J.Mol.Biol. 164: 355-375.

FINDLAY, R.C. and GALL, J.G. (1978). Free ribosomal RNA genes in Paramecium are tandemly repeated. Proc.Natl.Acad.Sci. USA 75: 3312-3316.

FIRTEL, R.A. (1981). Multigene families encoding actin and tubulin. Cell 24: 6-7.

FIRTEL, R.A., TIMM, R., KIMMEL, T.R. and McKEOWN, M. (1979). Unusual nucleotide sequence at the 5' end of actin genes in Dictyostelium discoideum. Proc.Natl.Acad.Sci. USA 76: 6206-6210.

FORD, P.J. and SOUTHERN, E.M. (1973). Different sequences for 5S RNA in kidney cells and ovaries of Xenopus laevis. Nature New Biol. 241: 7-12.

FYRBERG, E.A., KINDLE, K.L. and DAVIDSON, N. (1980). The actin genes of Drosophila: a dispersed multigene family. Cell 19: 365-378.

FYRBERG, E.A., BOND, B.J., HERSHEY, N.D., MIXTER, K.S. and DAVIDSON, N. (1981). The actin genes of _Drosophila_: protein coding regions are highly conserved but intron positions are not. Cell _24_: 107-116.

GALL, J.G. (1974). Free ribosomal RNA genes in the macro nucleus of _Tetrahymena_. Proc.Natl.Acad.Sci. USA _71_: 3078-3081.

GALLWITZ, D. and SEIDEL, R. (1980). Molecular cloning of the actin gene from yeast _Saccharomyces cerevisiae_. Nucleic Acid Res. _8_: 1043-1059.

GARNHAM, P.C.C. (1966). Malaria Parasites and other Haemosporidia. Oxford Blackwell Scientific Publications.

GILLESPIE, D. and SPIEGELMAN, S. (1965). A quantitative assay for DNA-RNA hybrids with DNA immobilized on a membrane. J.Mol.Biol. _12_: 829-842.

GLISIN, V., CRKVENJAKOV, R. and BJUS, C. (1979). Ribonucleic acid isolated by caesium chloride centrifugation. Biochemistry _13_: 2633-2637.

GLOVER, D.M. and HOGNESS, D.S. (1977). A novel arrangement of the 18S and 28S sequences in a repeating unit of _Drosophila melanogaster_ rDNA. Cell _10_: 167-176.

GOMAN, M., LANGSLEY, G., HYDE, J.E., YANKOVSKY, N.K., ZOLG, J.W. and SCAIFE, J.G. (1982). The establishment of genomic DNA libraries for the human malaria parasite _P. falciparum_ and identification of individual clones by hybridisation. Molecular and Biochemical Parasitology _5_: 391-400.

GRABOWSKI, P.J., ZAUG, A.J. and CECH, T.R. (1981). The intervening sequence of the ribosomal RNA precursor is converted to a circular RNA in isolated nuclei of *Tetrahymena*. Cell *23*: 467-476.

GRAHAM, D.E., NEUFELD, B.R., DAVIDSON, E.H. and BRITTEN, R.J. (1974). Interspersion of repetitive and non-repetitive DNA sequences in the Sea Urchin genome. Cell *1*: 127-137.

GUTTERIDGE, W.E. and TRIGG, P.I. (1970). Incorporation of radioactive precursors into DNA and RNA of *P. knowlesi in vitro*. J.Protozoology *17*: 89-96.

HAHN, U., LAZARUS, C.M., LUNSDORF, I.H. and KUNTZEL, H. (1979). Split gene for mitochondrial 24S ribosomal RNA of *Neurospora crassa*. Cell *17*: 191-200.

HALL, R., McBRIDE, J., MORGAN, G., TAIT, A., ZOLG, J.W., WALLIKER, D. and SCAIFE, J. (1983). Antigens of the erythrocytic stages of the human malaria parasite *Plasmodium falciparum* detected monoclonal antibodies. Molecular and Biochemical Parasitology *7*: 247-265.

HANUKOGLU, I., TANESE, N. and FUCHS, E. (1983). Complementary DNA sequence of a human cytoplasmic actin: interspecies divergence of 3'-non coding regions. J.Mol.Biol. *163*: 673-678.

HASAN, G., TURNER, M.J. and CORDINGLEY, J.S. (1982). Ribosomal genes of *Trypanosoma brucei*. Cloning of a rRNA gene containing a mobile element. Nucleic Acid Res. *10*: 6747-6761.

HAWLEY, R.S. and TARTOF, K.D. (1983). The ribosomal DNA of *Drosophila melanogaster* is organised differently from that of *D. hydei*. J.Mol.Biol. *163*: 499-503.

HENDERSON, A.S. and ATWOOD, K.C. (1976). Satellite association frequency and rDNA content of a double-satellited chromosome. Hum.Genet. 31: 113-115.

HERNANDEZ, R. and CASTANEDA, M. (1983). An endonuclease restriction analysis of the ribosomal RNA genes of *Trypanosoma cruzi*. Molecular and Biochemical Parasitology 8: 305-315.

HERNANDEZ, R., NAVA, G. and CASTANEDA, M. (1983). Small-size ribosomal RNA species in *Trypanosoma cruzi*. Molecular and Biochemical Parasitology 8: 297-304.

HIGHTON, P.J. and WHITFIELD, M. (1975). Similarities between the DNA molecules of bacteriophages 424, λ, and 21. Determined by denaturation and electron microscopy. Virology 63: 438-446.

HOMMEL, M., DAVID, P.H. and OLIGINO, L.D. (1983). Surface alterations of erythrocytes in *Plasmodium falciparum* malaria. I. Antigenic variation, antigenic diversity and the role of the spleen. J. Exp.Med. 157: 1137-1148.

HOUGH-EVANS, B.R. and HOWARD, J. (1982). Genome size and DNA complexity of *P. falciparum*. Biochimica et Biophysica Acta 698: 56-61.

HOWARD, R.J., BARNWELL, J.W. and KAO, V. (1983). Antigenic variation in *Plasmodium knowlesi* malaria: Identification of the variant antigen on infected erythrocytes. Proc.Natl.Acad.Sci. USA 80: 4129-4133.

HUMPHREYS, G.O., WILLSHAW, G.A. and ANDERSON, E.S. (1975). A simple method for the preparation of large quantities of pure plasmid DNA. Biochimica et Biophysica Acta 383: 457-463.

HUMPHRIES, S.E., WHITTALL, R., MINTY, A., BUCKINGHAM, M.E. and WILLIAMSON, R. (1981). There are approximately 20 actin genes in the human genome. Nucleic Acid Res. 9: 4895-4908.

HYDE, J.E., GOMAN, M., HALL, R., OSLAND, A., HOPE, I.A., LANGSLEY, G., ZOLG, W. and SCAIFE, J.G. (1983). Messenger RNA from the human malaria parasite *Plasmodium falciparum*; characterisation, translation studies and construction of a cDNA library. Molecular and Biochemical Parasitology (in press).

HYDE, J.E., ZOLG, J.W. and SCAIFE, J.G. (1981). Isolation and characterization of ribosomal RNA from the human malaria parasite *P. falciparum*. Molecular and Biochemical Parasitology 4: 283-290.

JORDAN, B.R. and GLOVER, D.M. (1977). 5.8S and 2S rDNA is located in the 'transcribed spacer' region between the 18S and 26S rRNA genes in *Drosophila melanogaster*. FEBS Letters 78: 271-274.

KAINE, B.P. and SPEAR, B.B. (1980). Putative actin genes in the macronucleus of *Oxytricha fallax*. Proc.Natl.Acad.Sci. USA 77: 5336-5340.

KARN, S., BRENNER, S., BARNETT, L. and CESARENI, G. (1980). Novel bacteriophage λ cloning vector. Proc.Natl.Acad.Sci. USA 77: 5172-5176.

KAUFMAN, R.J., BROWN, P.C. and SCHIMKE, R.T. (1979). Amplified dihydrofolate reductase genes in unstably methotrexate resistant cells are associated with double minute chromosomes. Proc. Natl.Acad.Sci. USA 76: 5669-5673.

KEMP, D.J., COPPEL, R.L., COWMAN, A.F., SAINT, B.B., BROWN, G.V. and ANDERS, R.F. (1983). Expression of Plasmodium falciparum blood-stage antigens in Escherichia coli: Detection with antibodies from immune humans. Proc.Natl.Acad.Sci. USA 80: 3787-3791.

KIDD, S.J. and GLOVER, D.M. (1981). Drosophila melanogaster ribosomal DNA containing Type II insertions is variably transcribed in different strains and tissues. J.Mol.Biol. 151: 645-662.

KINDLE, K.L. and FIRTEL, R.A. (1978). Identification and analysis of Dictyostelium actin gene, a family of moderately repeated genes. Cell 15: 763-778.

KOHORN, B.B. and RAE, P.M.M. (1982). Nontranscribed spacer sequences promote in vitro transcription of Drosophila ribosomal DNA. Nucleic Acid Res. 10: 6879-6886.

KORN, E. (1978). Biochemistry of actomyosin-dependent cell motility (A Review). Proc.Natl.Acad.Sci. USA 75: No. 2, 588-599.

KORN, L.J. and BROWN, D.D. (1978). Nucleotide sequence of Xenopus borealis oocyte 5S DNA: Comparison of sequences that flank several related eucaryotic genes. Cell 15: 1145-1156.

KRUGER, K., GRABOWSKI, P.J., ZAUG, A.J., SANDS, J., GOTTSCHLING, D.E. and CECH, T.R. (1982). Self-splicing RNA: auto-excision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell 31: 147-157.

LAEMMLI, U.K. (1970). Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature 227: 680-685.

LAKE, J.A. (1983). Evolving ribosome structure: Domains in Archaebacteria, Eubacteria, and Eucaryotes. Cell 33: 318-319.

LAZARIDES, E. and BALZER, D.R. (1978). Specificity of desmin to avian and mammalian muscle cells. Cell 14: 429-438.

LAZARIDES, E. and LINDBERG, U. (1974). Actin is the naturally occurring inhibitor of deoxyribonuclease I. Proc.Natl.Acad. Sci. USA 71: 4742-4746.

LEON, W., FOUTS, D.L. and MANNING, J. (1978). Analysis of *Leishmania donovani* DNA and the arrangement of the 16S and 26S rDNA genes. Nucleic Acid Res. 5: 491-504.

LIS, J.T., PRESTIDGE, L. and HOGNESS, D.S. (1978). A novel arrangement of tandemly repeated genes at a major heat shock site in *Drosophila melanogaster*. Cell 14: 901-919.

LONG, E.O., COLLINS, M., KREFER, B.I. and DAWID, I.B. (1981). Expression of the rDNA insertions in bobbed mutants of *Drosophila* melanogaster. Molec. and Gen.Genet. 182: 377-384.

LONG, E.O. and DAWID, I.B. (1980). Repeated genes in eukaryotes. Ann.Rev.Biochem. 727-764.

LUDUENA, R.F. and WOODWARD, D.O. (1973). Isolation and partial characterization of α- and β-tubulin from outer doublets of sea urchin sperm and microtubules of chick embryo brain. Proc. Natl.Acad.Sci. USA 70: 3594-3598.

McBRIDE, J.S., WALLIKER, D. and MORGAN, G. (1982). Antigenic diversity in the human malaria parasite: *P. falciparum*. Science 217: 254-257.

MacDONALD, G. (1957). The Epidemiology and Control of Malaria. Oxford: Oxford Univ. Press.

McKEOWN, M., TAYLOR, W.C., KINDLE, K.L. and FIRTEL, R.A. (1978). Multiple, heterogeneous actin genes in Dictyostelium. Cell 15: 789-800.

McKEOWN, M. and FIRTEL, R.A. (1981). Differential expression and 5' end mapping of actin genes in Dictyostelium. Cell 24: 799-807.

McKEOWN, M. and FIRTEL, R.A. (1981a). Evidence for sub-families of actin genes in Dictyostelium as determined by comparison of 3' end sequences. J.Mol.Biol. 151: 593-606.

MAIZELS, N. (1976). Dictyostelium 17S, 25S and 5S rDNAs lie within a 38,000 base pair repeated unit. Cell 9: 431-438.

MANIATIS, T., FRITSCH, E.F. and SAMBROOK, J. (1982). Molecular cloning: a laboratory manual. C.S.H. Laboratory, N.Y.

MANNING, J.E., SCHMID, C.W. and DAVIDSON, N. (1975). Interspersion of repetitive and nonrepetitive DNA sequences in the Drosophila melanogaster genome. Cell 4: 141-155.

MARCO, Y. and ROCHAIX, J.-D. (1980). Organisation of the nuclear ribosomal DNA of Chlamydomonas reinhardii. Molec.Gen.Genet. 177: 715-723.

MARMUR, J., ROWND, R. and SCHILDRAUT, C.L. (1963). Denaturation and renaturation of DNA. Prog.Nucl.Acid.Res. 232- 300.

MERLINO, G.T., WATER, R.D., CAMBERLAIN, J.P., JACKSON, D.A., RAAFAT EL-GEWELY, M. and KLEINSMITH, L.J. (1980). Cloning of sea urchin actin gene sequences for use in studying regulation of actin gene transcription. Proc.Natl.Acad.Sci. USA 77: 765-769.

MILLER, F.W. and LLAN, J. (1978). The ribosomes of Plasmodium berghei: isolation and ribosomal ribonucleic acid analysis. Parasitology 77: 345-365.

MILLER, K.G. and SOLLNER-WEBB, B. (1981). Transcription of mouse rRNA genes by RNA polymerase 1: In vitro and in vivo initiation and processing sites. Cell 27: 165-174.

MORRISEY, J.H. (1981). Silver stain for protein in polyacrylamide gels. A modified procedure with enhanced uniform sensitivity. Analytical Biochemistry 117: 307-310.

MURRAY, N.E. (1983). Phage lambda and Molecular Cloning. in Lambda II (Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.S. eds.) Cold Spring Harbor, N.Y.

MURRAY, N.E., BRAMMER, W.J. and MURRAY, K. (1977). Lambdoid phages that simplify the recovery of in vitro recombination. Molec. and Gen.Genet. 150: 53-61.

NAKAZATO, H. and EDMONDS, M. (1974). Purification of mRNA and Heterologous Nuclear RNA Containing Poly(A) Sequences. in Methods in Enzymology 29 (Grossman, L. and Moldave, K. eds.) pp. 431-443, Academic Press, New York.

NATH, K. and BOLLON, A.P. (1977). Organisation of the yeast ribosomal RNA gene cluster via cloning and restriction analysis. J.Biol.Chem. 252: 6562-6571.

NG, R. and ABELSON, J. (1980). Isolation and sequence of the gene for actin in Saccharomyces cerevisiae. Proc.Natl.Acad.Sci. USA 77: 3912-3916.

NILES, E.G., CUNNINGHAM, K. and JAIN, R. (1981). Structure of the Tetrahymena pyriformis rRNA gene. J.Biol.Chem. 256: 12857-12860.

ORJIH, A.U., BANYAL, H.S., CHEVLI, R. and FITCH, C.D. (1981). Hemin lyses malaria parasites. Science 214: 667-669.

PAN, W.-C. and BLACKBURN, E.H. (1981). Single extrachromosomal ribosomal RNA gene copies are synthesized during amplification of the rDNA in Tetrahymena. Cell 23: 459-466.

PARDUE, M.L., BRAWN, D.D. and BIRNSTIEL, M.L. (1973). Location of the genes for 5S ribosomal RNA in Xenopus laevis. Chromosoma 42: 191-203.

PASVOL, G. and WILSON, R.J.M. (1982). The interaction of malaria parasites with red blood cells. British Medical Bulletin 38: 133-140.

PAVLAKIS, G.N., JORDAN, B.R., WURST, R.M. and VOURNAKIS, J.N. (1979). Sequence and secondary structure of Drosophila melanogaster 5.8S and 25 rRNAs and of the processing site between them. Nucleic Acid Res. 7: 2213-2238.

PERKOWSKA, E., MACGREGOR, H.C. and BIRNSTIEL, M.L. (1968). Gene amplification in the oocyte nucleus of mutant and wild-type Xenopus laevis. Nature 217: 649-650.

PERRIN, L.H., DAYAL, R. and REIDER, H. (1981). Characterization of antigens from erythrocytic stages of P. falciparum reacting with immune human sera. Trans.R.Soc.Trop.Med.Hyg. 75: 163-165.

PERRY, R.P. (1976). 3 major rRNAs derived from a single precursor. Ann.Rev.Biochem. 45: 605-629.

PETERS, W., CHANCE, M.L., LISSNAR, R., MOMEN, H. and WARHURST, D.C. (1978). The chemotherapy of rodent malaria. XXX. The enigmas of the 'NS lines' of P. berghei. Annals of Tropical Medicine and Parasitology 72: No. 1. 23-34.

PETERS, W. (1980). Chemotherapy of Malaria. in Malaria, Vol. I: Epidemiology, Chemotherapy, Morphology and Metabolism. (Krier, J.P. ed), pp. 145-283, Academic Press, London.

PETERS, W. (1982). Antimalarial drug resistance. British Medical Bulletin 38: 187-192.

PETES, T.D. and BOTSTEIN, D. (1977). Simple Mendelian inheritance of the reiterated ribosomal DNA of yeast. Proc.Natl.Acad.Sci. USA 74: 5091-5095.

POLLACK, Y., KATZEN, A., SPIRA, D.T. and GOLENSER, J. (1982). The genome of Plasmodium falciparum. 1. DNA base composition. Nucleic Acid Res. 10: 539-545.

PRESCOTT, D.M. and MURTI, K.G. (1973). Chromosome structure in ciliated protozoans. Cold Spring Harb.Symp.Quant.Biol. 38: 609-618.

RAMKARAN, A.E. and PETERS, W. (1969). Infectivity of chloroquine resistant Plasmodium berghei to Anopheles stephensi enhanced by chloroquine. Nature 223: 635-666.

RASCH, E.M., BARR, H.J. and RASCH, R.W. (1971). The DNA content of sperm of Drosophila melanogaster. Chromosoma 33: 1-18.

RIGBY, P.W., DIECKMANN, M., RHODES, G. and BERG, P. (1977). Nick translation of both circular and linear DNAs. J.Mol.Biol. 113: 237-251.

RITOSSA, F.M. (1968). Unstable redundancy of genes for ribosomal RNA. Proc.Natl.Acad.Sci. USA 60: 509-516.

RITOSSA, F.M. and SPIEGELMAN, S. (1965). Localization of DNA complementary to ribosomal RNA in the nucleolus organiser region of Drosophila melanogaster. Proc.Natl.Acad.Sci. USA 53: 737-745.

ROCHAIX, J.-D. and MALNOE, P. (1978). Anatomy of the chloroplast ribosomal DNA of Chlamydomonas reinhardii. Cell 15: 661-670.

ROSARIO, V. (1981). Cloning of naturally occurring mixed infections of malaria parasites. Science 212: 1037-1038.

RUBIN, G.M. (1983). Dispersed Repetitive DNAs in Drosophila. in Mobile Genetic Elements (Shapiro, J.A. ed.) Academic Press.

RUBSTOV, P.M., MUSAKHANOV, M.M., ZAKHARYEV, V.M., KRAYEV, A.S., SKRYABIN, K.G. and BAYEV, A.A. (1980). The structure of the yeast ribosomal RNA genes. 1. The complete nucleotide sequence of the 18S ribosomal RNA gene from Saccharomyces cerevisiae. Nucleic Acids Res. 8: 5779-5794.

SALIM, M. and MADEN, E.H. (1981). Nucleotide sequence of Xenopus laevis 18S ribosomal RNA inferred from gene sequence. Nature 291: 205-208.

SANCHEZ, F., TOBIN, S.L., RDEST, U.,ZULAUF, E. and McCARTHY, B.J. (1983). Two Drosophila actin genes in detail: Gene structure, protein structure, and transcription during development. J. Mol.Biol. 163: 533-551.

SCHACHAT, F., O'CONNOR, D.J. and EPSTEIN, H.F. (1978). The moderately repetitive DNA sequences of Caenorhabditis elegans do not show short-period interspersion. Biochmica et Biophysica Acta 520: 688-692.

SCHMID, C.W. and DEININGER, P.L. (1975). Sequence organization of the human genome. Cell 6: 345-357.

SCHULER, M.A. and KELLER, E.B. (1981). The chromosomal arrangement of two linked actin genes in the sea urchin Strongylocentrotus purpuratus. Nucleic Acid Res. 9: 591-604.

SHAPIRO, J.A. and CORDELL, B. (1982). Eukaryotic mobile and repeated genetic elements. Biol. Cell 43: 31-54.

SHERMAN, I.W. (1979). Biochemistry of Plasmodium (Malarial parasites). Microbiological Reviews 453-495.

SHINE, J. and DALGARNO, L. (1974). The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc.Natl.Acad.Sci. USA 71: 1342-1346.

SINDEN, R.E. (1983). The cell biology of sexual development in Plasmodium. Parasitology 86: 7-28.

SMITH, H.O. and BIRNSTIEL, M.L. (1976). A simple method for DNA restriction site mapping. Nucleic Acid Res. 3: 2387-2398.

SMITH, G.E. and SUMMERS, M.D. (1980). The bidirectional transfer of DNA and RNA to nitrocellulose or diazobensyloxymethyl-paper. Anal.Biochem. 109: 123-129.

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

SPRADLING, A.C. and RUBIN, G.M. (1981). Drosophila genome organisation: conserved and dynamic aspects. Ann.Rev.Gen. 15: 219-264.

SUTCLIFFE, J.G. (1978). Complete nucleotide sequence of the Escherichia coli plasmid pBR322. Cold Spring Harb.Symp.Quant. Biol. 43: 77-90.

TAIT, A. (1981). Analysis of protein variation in P. falciparum by two-dimensional gel electrophoresis. Molecular and Biochemical Parasitology 2: 205-217.

TARTOF, K.D. (1971). Increasing the multiplicity of ribosomal RNA genes in Drosophila melanogaster. Science 171: 294-297.

TAYLOR, D.W. and SIDDIQUI, W.A. (1982). Recent advances in malaria immunity. Ann.Rev.Med. 33: 69-96.

THAITHONG, S. and BEALE, G.H. (1981). Resistance of ten Thai isolates of P. falciparum to chloroquine and pyrimethamine by in vitro tests. Trans.R.Soc.Trop.Med.Hyg. 75: 271-273.

THAITHONG, S., BEALE, G.H., FENTON, B., McBRIDE, J., ROSARIO, V., WALKER, A. and WALLIKER, D. (1983). Clonal diversity in a single isolate of the malaria parasite Plasmodium falciparum. Trans.R.Soc.Trop.Med.Hyg. (in press).

TOBIN, S.L., ZULAUF, E., SANCHEZ, F., CRAIG, E.A. and McCARTHY, B.J. (1980). Multiple actin-related sequences in the Drosophila melanogaster genome. Cell 19: 121-131.

TOWBIN, H., STAEHELIN, T. and GORDON, J. (1979). Electrophoretic transfer of proteins from poly acrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc.Natl.Acad.Sci. USA 76: 4350-4354.

TRAGER, W. and JENSEN, J.B. (1976). Human malaria parasites in continuous culture. Science 193: 673-675.

TWIGG, A.J. and SHERRATT, D. (1980). Trans-complementable copy-number mutants of plasmid Col E1. Nature 283: 216-218.

VALENZUELA, P., BELL, G.I., VENEGAS, A., SEWEL, E.T., MASIARZ, F.R., DE GENNARO, L.J., WEIENBERG, F. and RUTTER, W.J. (1977). Ribosomal RNA genes of Saccharomyces cerevisiae. J.Biol.Chem. 252: 8126-8135.

VANDEKERCKHOVE, J. and WEBER, K. (1978). The amino acid sequence of Physarum actin. Nature 276: 720-721.

VANDEKERCKHOVE, J. and WEBER, K. (1980). Vegetative Dictyostelium cells containing 17 actin genes express a single major actin. Nature 284: 475-477.

VERBEET, M.Ph., KLOOTWIJK, J., v. HEERIKHUIZEN, H., FONTIJN, R., VREUGDENHIL, E. and PLANTA, R.J. (1983). Molecular cloning of the cDNA of Saccharomyces rosei, and comparison of its transcription initiation region with that of S. carlsbergensis. Gene 23: 53-63.

VICKERMAN, K. and COX, F.E.G. (1967). The Protozoa. in Introductory Studies in Biology (Murray, S.W. ed.) William Clowes & Sons Ltd., London and Beccles.

VILLALBA, E. and RAMIREZ, J.L. (1982). Ribosomal DNA of Leishmania brasiliensis: Number of ribosomal copies and gene isolation. J.Protozool. 29(3) 438-441.

WALKER, P.M.B. (1971). Origin of satellite DNA. Nature 229: 306-308.

WALKER, T.A. and PACE, N.R. (1983). 5.8S ribosomal RNA. Cell 33: 320-322.

WALLACE, H. and BIRNSTIEL, M.L. (1966). Ribosomal cistrons and the nucleolar organizer. Biochim.Biophys.Acta 114: 296-310.

WALLIKER, D. (1982). Genetic variation in malaria parasites. British Medical Bulletin 38: 123-128.

WALLIKER, D. (1983). The genetic basis of diversity in malaria parasites. Advances in Parasitology 22: 217-259.

WELLAUER, P.K. and DAWID, I.B. (1977). The structural organisation of ribosomal DNA in Drosophila melanogaster. Cell 10: 193-212.

WELLAUER, P.K. and DAWID, I.B. (1978). Ribosomal DNA in Drosophila melanogaster. J.Mol.Biol. 126: 769-782.

WELLAUER, P.K., DAWID, I.B., BROWN, D.D. and REEDER, R.H. (1976). The molecular basis for length heterogeneity in ribosomal DNA from Xenopus laevis. J.Mol.Biol. 105: 461-486.

WETMUR, J.G. and DAVIDSON, N. (1968). Kinetics of renaturation of DNA. J.Mol.Biol. 31: 349-370.

WHITE, G.B. (1982). Malaria vector ecology and genetics. British Medical Bulletin 38: 207-212.

WILD, M.A. and GALL, J.G. (1979). An intervening sequence in the gene coding for the 28S RNA of Tetrahymena pigmentosa. Cell 16:565-573.

WILSON, R.J.M. (1980). Serotyping of P. falciparum malaria with S-antigens. Nature 284: 451-452.

WIMBER, D.E. and STEFFENSEN, D.M. (1970). Localization of 5S RNA genes on Drosophila chromosomes by RNA-DNA hybridisation. Science 170: 639-641.

WOOL, I.G. (1979). The structure and function of eukaryotic ribosomes. Ann.Rev.Biochem. 48: 719-754.

WORLD HEALTH STASTICICS ANNUAL. World Health Organisation, Geneva, yearly updates.

YAO, M.-C., BLACKBURN, E. and GALL, J.G. (1979). Amplification of the rRNA genes in Tetrahymena. Cold Spring Harb.Symp.Quant.Biol. XLIII: 1293-19296.

YAO, M.-C. and GALL, J.G. (1977). A single integrated gene for ribosomal RNA in a eucaryote, _Tetrahymena pyriformis_. Cell _12_: 121-132.

YAO, M.-C. and YAO, C.-H. (1981). Repeated hexanucleotide C-C-C-C-A-A is present near free ends of macronuclear DNA of _Tetrahymena_. Proc.Natl.Acad.Sci. USA _78_: 7436-7439.

ZUKER, C. and LODISH, H.F. (1981). Repetitive DNA sequences co-transcribed with developmentally regulated _Dictyostelium discoideum_ mRNAs. Proc.Natl.Acad.Sci. USA _78_: 5386-5390.

ZULAF, E., SANCHEZ, F., TOBIN, S.L., RDEST, U. and McCARTHY, B. (1981). Developmental expression of a _Drosophila_ actin gene encoding actin I. Nature _292_: 556-558.