

***S.pombe* Artificial Chromosome (SPARC)  
Vectors for Cloning Large DNA Fragments**

**Dorothy J.D. Young**

**PhD**

**University of Edinburgh**

**1997**





## ABSTRACT

The fission yeast, *Schizosaccharomyces pombe*, has several potential advantages as a host for a cloning system for large DNA molecules when compared to the budding yeast, *Saccharomyces cerevisiae*. The possibility of constructing a large DNA cloning system in *S. pombe* has been investigated.

*S. pombe* has 3 large chromosomes (greater than 3.5 Mb) and may therefore have the ability to carry megabase sized artificial chromosomes. In addition artificial chromosomes of less than 3.5 Mb will migrate below the host chromosomes on pulsed field gels and thus will be easier to separate away from the host chromosomes. As in higher eukaryotes, the *S. pombe* centromeres have arrays of repetitive elements, therefore such sequences from heterologous sources which rearrange in *S. cerevisiae* may not do so in *S. pombe*.

It has been shown that small acentric linearised plasmids with cloned *S. pombe* telomeres at each end replicate intact in *S. pombe*. These plasmids were used to prepare vector arms each having a selectable marker, *S. pombe* telomere and *S. pombe* replication origin. The *S. pombe* artificial chromosome (SPARC) vectors do not contain any *S. pombe* centromere sequences as they are not necessary for the maintenance of the SPARCs under selection.

To test the potential of the SPARC vectors for cloning human sequences in *S. pombe* well characterised human fragments derived from cosmids were cloned in *S. pombe* using these SPARC vectors. Analyses of the DNA from recombinant transformants



carrying SPARCs of approximately 50 kb revealed that each of the cloned cosmid fragments cloned had remained intact.

Large fragments of *NotI* digested human DNA have now been cloned using this SPARC system. Several transformants were isolated which contain SPARCs of greater than 100 kb. These SPARCs were acentric and appeared to replicate at an average of approximately 5 copies per cell. The human inserts of a few of these SPARCs has shown that they map to single locations within the human genome using FISH analyses.

The ability to clone and propagate large fragments of human DNA in *S. pombe* has been demonstrated. It should be possible to construct libraries using these SPARC vectors in *S. pombe*. SPARC libraries may then be used in conjunction with libraries constructed in YACs, BACs and PACs.



## ABBREVIATIONS

A	adenine
<i>ade</i>	<i>S. pombe</i> adenine gene
<i>arg</i>	<i>S. pombe</i> arginine gene
ARS	autonomously replicating sequence in <i>S. cerevisiae</i>
<i>ars</i>	autonomously replicating sequence in <i>S. pombe</i>
BAC	bacterial artificial chromosome
bp	base pair of DNA
°C	degrees centigrade
C	cytosine
<i>C. elegans</i>	<i>Caenorhabditis. elegans</i>
CEN	<i>S. cerevisiae</i> centromere
<i>cen</i>	<i>S. pombe</i> centromere
CFTR	Cystic fibrosis transmembrane conductance regulator
cfu	colony forming units
Cot1	C <sub>Ot</sub> of 1, enriched for sequences repeated more than 10,000x/genome
dCTP	deoxycytidine triphosphate
dH <sub>2</sub> O	distilled water
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphates
<i>E. coli</i>	<i>Escherichia coli</i>
EBV	Epstein Barr virus
EDTA	ethylenediamineettra-acetic acid, disodium salt
EMMG	Edinburgh minimal media with glutamic acid
FISH	Fluorescence <i>in situ</i> hybridisation
G	guanine



g	grams
<i>HIS</i>	<i>S. cerevisiae</i> histidine gene
<i>his</i>	<i>S. pombe</i> histidine gene
Ig	immunoglobulin
kb	kilobase pairs of DNA
LCR	locus control region
<i>LEU</i>	<i>S. cerevisiae</i> leucine gene
<i>leu</i>	<i>S. pombe</i> leucine gene
LMP	low melting point
M	molar
ml	millilitre
$\mu$ l	microlitre
mg	milligram
$\mu$ g	microgram
mm	millimetre
mM	millimolar
MAC	mammalian artificial chromosome
Mb	megabase pairs of DNA
Mol. Wt.	molecular weight
Myr	million years
mRNA	messenger ribonucleic acid
N	not a conserved nucleotide
ng	nanogram
dNTP	deoxyribonucleotide triphosphates
p	short arm of chromosome
PAC	P1 derived artificial chromosome
PCR	polymerase chain reaction



Pu	purine
q	long arm of chromosome
RNase	ribonuclease
SPARC	<i>S. pombe</i> artificial chromosome
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	sodium dodecyl sulphate
<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i>
SSC	standard saline citrate (150 mM NaCl, 75 mM tri-sodium citrate)
T	thymine
TE	Tris:EDTA (10 mM Tris pH 8.0, 1 mM EDTA pH 8.0)
Tel	telomere
Tris	2-Amino-2-(hydroxymethyl) -1,3-propandiol (C <sub>4</sub> H <sub>11</sub> NO <sub>3</sub> )
<i>TRP</i>	<i>S. cerevisiae</i> tryptophan gene
tRNA	transfer ribonucleic acid
<i>URA</i>	<i>S. cerevisiae</i> uracil gene
<i>ura</i>	<i>S. pombe</i> uracil gene
UV	ultraviolet light
YAC	yeast artificial chromosome



## MANUFACTURERS

( $\alpha$ - <sup>32</sup> P)dCTP	Amersham
Agarose for PFGE	Sigma
$\beta$ -agarase	NE Biolabs
BCIP/NBT alkaline phosphatase substrate kit IV	Vector
Biotin-11-dUTP	Boehringer Mannheim
$\beta$ -mercaptoethanol	Sigma
Cot 1 DNA	Gibco BRL
dNTPs	Promega
Genescreen	Dupont
GeneClean kit	Bio 101
Glusulase	Biotechnology systems
Lipofectin	Gibco BRL
Novozym 234	Novo Biolabs
NICK columns	Pharmacia
Proteinase K	Boehringer Mannheim
Random prime kit	Boehringer Mannheim
Restriction enzymes	Boehringer Mannheim
Reagents for FISH	Vector
Sephadex G50	Pharmacia
Streptavidin alkaline phosphatase	Boehringer Mannheim
TAPs	Sigma
<i>Taq</i> polymerase	Boehringer Mannheim
Zymolyase	ICN Biochemical inc.

All other materials were obtained from Sigma or BDH.



## ACKNOWLEDGEMENTS

I would like to thank many people for all their help, support and encouragement throughout the duration of my PhD. Firstly I have to thank Robin Allshire, my supervisor. I also received assistance from Elaine Nimmo over the years. To those past and present from the "pombe" group - Gwen Cranston, Niki Redhead, Jean-Paul Javerat, and Karl Ekwall who have also contributed helping me with various problems throughout my project. For help and advice on high molecular weight DNA technology I have to thank David Kipling, Brenda Grimes and Nio McGill. I cannot neglect the other members of the Chromosome Biology section; Howard Cook, Shona Kerr, Dominique Broccoli, Mary Taggart, Helen Wilson, Rosey Bayne, Arthur Mitchell, Peter Warburton, Philip Lord and to the other PhD students, in particular Jeff Craig and Kathy Williamson. I must also extend my thanks to Peter Fantes for being my University supervisor.

I would like to express my gratitude to Wendy Bickmore for the provision of the human DNA sample and other FISH reagents, to Harris Morrison for the provision of prepared slides with human metaphases to allow me to carry out the FISH analysis and finally to Muriel Lee without whom I would not have been able to identify the human chromosomes.

Thanks also go to Norman Davidson, Sandy Bruce and Douglas Stuart in photography and to Alistair Brown in the computing section. Finally I would like to thank Donald and my dad for all their support.



# CONTENTS

<b>Declaration .....</b>	<b>ii</b>
<b>Abstract .....</b>	<b>iii</b>
<b>Abbreviations.....</b>	<b>v</b>
<b>Manufacturers .....</b>	<b>viii</b>
<b>Acknowledgements .....</b>	<b>ix</b>
<b>Contents .....</b>	<b>x</b>
<b>List of Figures.....</b>	<b>xiv</b>
<b>List of Tables.....</b>	<b>xvi</b>
<b>1 INTRODUCTION.....</b>	<b>1</b>
1.1 THE FISSION YEAST, <i>Schizosaccharomyces pombe</i> .....	2
1.2 COMPONENTS OF THE <i>S. pombe</i> CHROMOSOME .....	3
1.2(a) <i>S. pombe</i> CENTROMERES .....	5
1.2(b) REPLICATION ORIGINS .....	10
1.2(c) <i>S. pombe</i> TELOMERES .....	11
1.3 VECTOR SYSTEMS FOR CLONING LARGE DNA FRAGMENTS .....	12
1.4 DEVELOPMENT OF <i>S. cerevisiae</i> YEAST ARTIFICIAL CHROMOSOMES .....	12
1.5 THE YAC SYSTEM: ADVANTAGES AND DISADVANTAGES .....	16
1.6 OTHER VECTOR SYSTEMS FOR CLONING LARGE DNA FRAGMENTS .....	20
1.7 POSSIBLE ADVANTAGES OF AN <i>S. pombe</i> ARTIFICIAL CLONING SYSTEM .....	23
1.8 FEASIBILITY OF CONSTRUCTING SPARC VECTOR SYSTEM .....	25



<b>2 MATERIALS AND METHODS .....</b>	<b>27</b>
2.1 STRAINS .....	28
2.2 PLASMIDS, YACS & MINI CHROMOSOMES .....	28
2.3 MEDIA .....	29
2.4 PLASMID PREPARATION .....	31
2.5 PREPARATION OF <i>S. cerevisiae</i> DNA IN AGAROSE PLUGS .....	33
2.6 PREPARATION OF <i>S. pombe</i> DNA IN AGAROSE PLUGS .....	33
2.7 PREPARATION AND TRANSFORMATION OF <i>S. pombe</i> .....	34
2.8 20 ml <i>S. pombe</i> DNA PREPARATION .....	36
2.9 ANALYSES OF DNA USING RESTRICTION ENZYMES .....	37
2.10 GENETIC CROSS .....	37
2.11 GEL ELECTROPHORESIS .....	38
2.12 SOUTHERN BLOTTING .....	38
2.13 CATCH LINKER PCR .....	40
2.14 PREPARATION OF BIOTINYLATED PROBES .....	41
2.15 FLOURESCENCE <i>IN SITU</i> HYBRIDISATION .....	42
<b>3 TRANSFORMATION OF <i>S. pombe</i> WITH <i>S. cerevisiae</i> YACS CONTAINING HUMAN DNA.....</b>	<b>44</b>
3.1 INTRODUCTION .....	45
3.2 ISOLATION OF YAC DNA .....	46
3.3 TRANSFORMATION OF <i>S. pombe</i> WITH YACS.....	51
3.4 ANALYSIS OF TRANSFORMANTS .....	53
3.5 DISCUSSION .....	59



<b>4 CLONING COSMID FRAGMENTS IN SPARC VECTORS .....</b>	<b>61</b>
4.1 INTRODUCTION .....	62
4.2 <i>S. pombe</i> ARTIFICIAL CHROMOSOME (SPARC) VECTOR SYSTEM .....	64
4.3 CLONING HUMAN COSMID FRAGMENTS INTO SPARC VECTORS .....	66
4.4 CLONING OF COSMIDS I AND II .....	69
4.5 RECOMBINATION DEFICIENT MUTATION CROSSED INTO <i>S. pombe</i> .....	78
4.6 CLONING OF THE HUMAN FRAGMENTS FROM pWE15 COSMIDS .....	80
4.7 DISCUSSION .....	87
<b>5 GENERATION OF SPARCS WITH LARGE HUMAN INSERTS .....</b>	<b>88</b>
5.1 INTRODUCTION .....	89
5.2 ISOLATION OF HUMAN DNA AND LIGATION INTO SPARC VECTORS .....	90
5.3 ANALYSIS OF SPARC TRANSFORMANTS .....	93
5.4 ESTIMATION OF SPARC COPY NUMBER .....	97
5.5 TRANSFORMATION EFFICIENCY .....	101
5.6 DISCUSSION .....	105
<b>6 ANALYSIS OF SPARCS CONTAINING LARGE HUMAN INSERTS .....</b>	<b>107</b>
6.1 INTRODUCTION .....	108
6.2 ANALYSIS OF SPARCS THROUGH MITOSIS .....	109



6.3 DETERMINING LOSS OF SPARCS THROUGH MEIOSIS .....	116
6.4 ANALYSIS OF SPARCS FOR CHIMERIC INSERTS .....	118
6.5 ANALYSIS OF SPARCS TO INVESTIGATE POSSIBLE INSERT REARRANGEMENTS .....	121
6.6 DISCUSSION .....	124
<b>7 DISCUSSION .....</b>	<b>127</b>
7.1 ANALYSES OF THE SPARC VECTOR SYSTEM .....	128
7.2 FUTURE INVESTIGATIONS OF THE SPARC SYSTEM .....	132
7.3 MODIFICATIONS TO SPARC VECTORS .....	134
7.4 POTENTIAL OF SPARC VECTORS .....	137
7.5 CONCLUSIONS .....	140
<b>8 BIBLIOGRAPHY .....</b>	<b>141</b>



## LIST OF FIGURES

Figure 1.1 Structure of <i>S. pombe</i> centromeres: strain SP223 .....	6
Figure 1.2 An <i>S. pombe</i> minimal centromere (approximately 17 kb) .....	9
Figure 1.3 Consensus sequence of 125 bp functional <i>S. cerevisiae</i> centromere .....	14
Figure 1.4 Structure of Yeast Artificial Chromosome (YAC) .....	18
Figure 3.1 Structure of the pSp( <i>cen1</i> )7L YAC .....	47
Figure 3.2 Sizes of YACs .....	49
Figure 3.3 Treatment of high molecular weight DNA to prevent degradation .....	52
Figure 3.4 Analysis of DNA from <i>S. pombe</i> transformed with YACs .....	55
Figure 3.5 DNA from <i>S. pombe</i> transformants with centromere YAC and minichromosome .....	58
Figure 4.1 Method used to clone human DNA fragments into SPARC vectors .....	65
Figure 4.2 Schematic diagram showing the human <i>NotI</i> fragments produced from the cosmids .....	68
Figure 4.3 DNA from <i>S. pombe</i> transformants with cosmid I SPARCs .....	70
Figure 4.4 <i>NotI</i> restriction digest analysis on DNA from 4 cosmid I SPARC transformants .....	72
Figure 4.5 Further restriction analysis on DNA from 3 cosmid I transformants .....	74
Figure 4.6 <i>NotI</i> restriction digest analysis on DNA from 4 cosmid II SPARC transformants .....	76
Figure 4.7 Further restriction analysis on DNA form 4 cosmid II SPARC transformants .....	77
Figure 4.8 Construct used to determine recombination deficient <i>S. pombe</i> strains ...	79
Figure 4.9 DNA from <i>S. pombe</i> with SPARCs containing small human DNA	



inserts .....	82
Figure 4.10 <i>NotI</i> restriction digest analyses on DNA from transformants with SPARCs .....	84
Figure 4.11 Further restriction analyses on DNA from transformants with SPARCs .....	86
Figure 5.1 Ligation of SPARC vector arms to large DNA fragments .....	92
Figure 5.2 DNA from human SPARC transformants .....	94
Figure 5.3 Ten human SPARCs .....	96
Figure 5.4 Copy number of several SPARC strains .....	99
Figure 5.5 DNA from <i>S. pombe</i> transformants recovered after retransforming with SPARCs .....	104
Figure 6.1 Maintenance of SPARCs through many cell divisions .....	111
Figure 6.2 Maintenance of SPARC bands form strains with more then one initial band .....	113
Figure 6.3 Isolation of single SPARC bands derived from transformant 14 .....	115
Figure 6.4 Using FISH to determine chromosomal location of the SPARCs human insert .....	120
Figure 6.5 Comparing the human <i>NotI</i> fragment of SPARC 16 with total human DNA digest .....	122



## LIST OF TABLES

Table 5.1 Size Of SPARC Bands In Each Transformant .....	97
Table 5.2 Copy Number Of SPARCs Shown In Figure 5.4 .....	100
Table 6.1 Numbers Of Spores Through Meiosis Containing SPARC .....	117



# 1 INTRODUCTION



## 1.1 THE FISSION YEAST, *Schizosaccharomyces pombe*

The unicellular eukaryote *Schizosaccharomyces pombe* (fission yeast) is well established as a model organism in many areas of research such as in the study of the eukaryote cell cycle and chromosome structure and function (Allshire, 1995; Moreno *et al.*, 1991; Russell, 1989). These studies have demonstrated that *S. pombe* can be manipulated using techniques similar to those used when working with the budding yeast *Saccharomyces cerevisiae*. *S. pombe* and *S. cerevisiae* are not closely related, it has been estimated that they diverged from a common ancestor approximately 1200 Myr ago (Huysmans *et al.*, 1983). The *S. pombe* genome is approximately 13.7 Mb, similar to the genome size of *S. cerevisiae*. The *S. cerevisiae* genome is organised into 16 chromosomes ranging from 225 kb to 1600 kb (Mortimer *et al.*, 1989), whereas the *S. pombe* genome is organised into three chromosomes; I, II and III of 5.5, 4.7 and 3.5 Mb respectively (Fan *et al.*, 1988; Kohli *et al.*, 1977; Robinow, 1977). Several aspects of *S. pombe* biology make it appear more similar to higher eukaryotes.

The cell cycle of *S. pombe* serves as a useful model system for higher eukaryotes (reviewed Moreno *et al.*, 1991; Russell, 1989). Such studies have identified many of the proteins involved in the *S. pombe* cell cycle and comparisons can be made with homologues from other eukaryotes. The three *S. pombe* chromosomes condense through mitosis and when stained are visible by light microscopy (Robinow, 1977). Fluorescence *in situ* hybridisation (FISH) has been developed for *S. pombe*, therefore enabling chromosomal structures to be visualised throughout the *S. pombe* cell cycle (Uzawa and Yanagida, 1992; Funabiki *et al.*, 1993). This technique has proved useful for the analysis of chromosomal structures such as centromeres and telomeres



and other features of fission yeast chromosome biology (Funabiki et al., 1993; Allshire et al., 1994, 1995; Ekwall et al., 1996).

The splicing machinery of *S. pombe* is similar to that of higher eukaryotes (Kaufer, et al., 1985). The mRNA of SV40 small t antigen is correctly spliced in *S. pombe* but not in *S. cerevisiae*. This is probably because within *S. cerevisiae* the recognition sequences for splicing are highly conserved whereas in other eukaryotes they are more degenerate. *S. pombe* also has a higher percentage of introns per gene than *S. cerevisiae*.

*S. pombe* has complex centromeres, like those of higher eukaryotes, with arrays of repetitive elements arranged as an inverted repeat around a central core (reviewed Clarke, 1990; Carbon and Clarke, 1990). The centromeres are 35, 65 and 110 kb for chromosomes I, II and III respectively. In comparison, the centromere of *S. cerevisiae* is much simpler in that complete centromere function can be provided by only 125 bp of DNA.

## **1.2 COMPONENTS OF THE *S. pombe* CHROMOSOME**

Maintenance of a linear eukaryotic chromosome through mitotic and meiotic divisions requires three essential elements: a centromere (Blackburn and Szostak, 1984; Schulman and Bloom, 1991; Pluta et al., 1995); telomeres (Blackburn and Szostak, 1984; Blackburn, 1991; Kipling, 1995); and replication origins (reviewed Bell, 1995). Centromeres have long been recognised as the primary constriction on highly condensed metaphase chromosomes. The centromere is the chromosomal region where the kinetochore is assembled allowing microtubules to attach and



segregate chromosomes on mitotic and meiotic spindles (Willard, 1990; Schulman and Bloom, 1991). Centromeres tend to be associated with a class of chromatin which remains condensed throughout the cell cycle. This is known as heterochromatin and in general such regions are transcriptionally inert, late replicating and suppressed for recombination (Lohe & Hilliker., 1995; Pluta *et al.*, 1995; Vogt, 1992; Rattner, 1991; Singer, 1982). Telomeres are present at the termini of the chromosomes and protect them from damage and are structured to allow end replication of linear chromosomes by telomerase (Blackburn and Szostak, 1984; Blackburn, 1991). DNA synthesis is initiated at discrete sites known as replication origins. Firing of replication origins is tightly regulated with respect to cell cycle events. Replication of chromosomes is initiated and completed during S phase providing the substrates for segregation of in mitosis (for review Bell, 1995). In *S. pombe* several replication origins have been identified and their DNA sequences determined (Maundrell *et al.*, 1988; Dubey *et al.*, 1996).

A strain of *S. pombe* disomic for chromosome III has been utilised to create minichromosome derivatives (Niwa *et al.*, 1989; Matsumoto *et al.*, 1990). The disomic strain was  $\gamma$ -irradiated to break chromosome III and isolates were propagated which contained stable minichromosomes of various sizes. The fragmented chromosomes were found to have been healed with 300 bp of telomere sequence at the breakpoint (Matsumoto *et al.*, 1987). The minichromosomes were used to study the minimal centromeric region of chromosome III. These studies have shown that acentric minichromosomes can be maintained as linear molecules which behave like a linear plasmid and can exist under selection without a centromere (Niwa *et al.*, 1989; Matsumoto *et al.*, 1990). All three fission yeast centromeres have been cloned in *S. cerevisiae* as Yeast Artificial Chromosomes (YACs) which have facilitated the analysis of centromere structure (Hahnenberger *et al.*, 1989). One of these *S.*



*cerevisiae* YACs (75 kb) containing *S. pombe* centromere I was transformed into *S. pombe* and while in most of the transformants the YAC had integrated or rearranged, in approximately 10% of transformants 300 bp of *S. pombe* telomere repeats were added onto the ends of the YAC creating a linear minichromosome of 75 kb (Hahnenberger *et al.*, 1991). The ability to construct such minichromosomes suggested that construction of a YAC-like cloning system in *S. pombe* would be feasible.

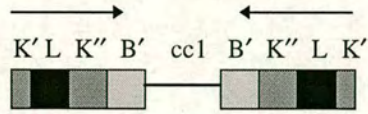
### 1.2(a) *S. pombe* CENTROMERES

It is clear that *S. cerevisiae* centromeres can not function in *S. pombe* (Clarke, 1991; Carbon & Clarke, 1991; Chikashige *et al.*, 1989). The analysis of the *S. pombe* centromeres (figure 1.1) revealed a centromeric structure vastly different from the simpler centromere structure of *S. cerevisiae*. The structure of each fission yeast centromere differs. The largest fission yeast chromosome, chromosome I, has the smallest centromere of 35 kb, the centromere of chromosome II is 65 kb and the largest centromere of 110 kb is on chromosome III, the shortest *S. pombe* chromosome. Due to the size and complexity of the *S. pombe* centromeres only smaller fragments could be recovered in *Escherichia coli*. Entire centromeres were cloned into *S. cerevisiae* YACs, on *SalI* fragments of 50 kb, 90 kb and 120 kb for chromosomes I, II and III respectively (Fishel *et al.*, 1988). Initial observations using chromosome walking from markers adjacent to the centromeres revealed that *S. pombe* centromeres contain repetitive elements in a symmetrical arrangement (Nakaseko *et al.*, 1986; Clarke *et al.*, 1986). The structure of all three centromeres was elucidated by various methods including cloning into YACs, excision and repair methods and by creating minichromosomes through radiation induced chromosome

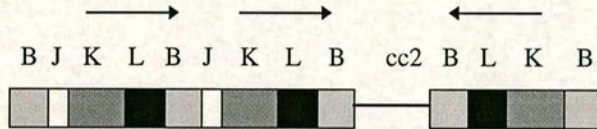


**Figure 1.1 Structure of *S. pombe* Centromeres**

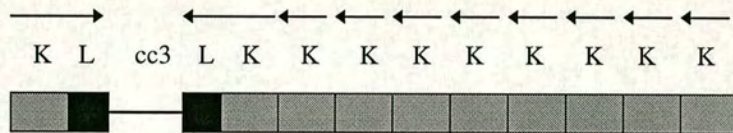
Centromere 1



Centromere 2



Centromere 3



Structure of *S. pombe* centromeres from strain SP223 (adapted from Steiner *et al.*, 1993) showing the repetitive nature of the *S. pombe* centromeres. Each centromere is arranged with different numbers and orders of the repeat units. However each centromere has an overall inverted repeat structure.



breakage (Fan *et al.*, 1988; Fishel *et al.*, 1988; Chikashige *et al.*, 1989; Niwa *et al.*, 1986, 1989; Hahnenberger *et al.*, 1989, 1991; Takahashi *et al.*, 1992).

Although the centromeres are different in size, the basic underlying structure of each is the same, an inverted repeat arranged around a non repetitive central core (figure 1.1). The central core is unique to each centromere except for the element, *tm*, present in centromeres 1 and 3 (Matsumoto *et al.*, 1990; Hahnenberger *et al.*, 1991). The inverted repetitive structure is composed of a different number of repeat units around the central core. Two of these repeat units are found in centromere 1, while the number of repeat units can vary from 2 to 4 in centromere 2 and from 11 to 13 repeat units in centromere 3. The smallest natural centromere appears to require at least 2 repeat units, one either side of the central core and the number of repeat units varies between different strains of *S. pombe* (Steiner *et al.*, 1993). Each of these repeats have been defined further into repetitive elements which can be shown to be present in different configurations within each centromere.

The B repeat (or *imr*) elements of the centromere have a unique sequence apart from the presence of tRNA genes (Takahashi *et al.*, 1991; Kuhn *et al.*, 1991). It is not known whether these genes are transcriptionally active or if they play a role in the centromere function. The outer repeats (*otr*) of the three centromeres mainly consist of the elements known as K and L (Clarke and Baum, 1990; Hahnenberger *et al.*, 1989 1991)(also known as *dg* and *dh* in different configurations (Murakani *et al.*, 1991; Chikashige *et al.*, 1989; Fan *et al.*, 1988)). These repeat elements of around 5-6 kb in length are homologous. However in centromere I the K repeat is split into K', adjacent to the inner repeat, and K'' next to the L repeat (Hahnenberger *et al.*, 1991). The K'' repeat is also in the opposite orientation to the same sequences in



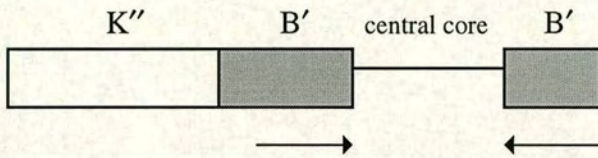
centromeres 2 and 3. In addition there is another repeat found on centromere 2 which has been termed the J repeat (Clarke and Baum, 1990).

Several deletion derivatives of centromere 1 have been constructed to determine which repeat units are important in centromere function (Hahnenberger *et al.*, 1991). The minimal deletion construct with mitotic centromere function is a 17 kb fragment containing K' through the central core into the B' repeat on the other side (figure 1.2). The mitotic segregation frequency is reduced although still partially maintained while meiotic function is abolished. It is thought that the outer repeats may play a role in sister chromatid separation during meiosis (Hahnenberger *et al.*, 1991). A centromere construct which consists only of the central core from centromere 2 with a K' repeat element from centromere 1 has been shown to have some mitotic function (Baum *et al.*, 1994). Tandem arrays of K' included in this construct result in an increase in the centromeric activity, therefore the K repeats and central core appear to be important for mitotic function. The roles of the other repeats are not known.

*S. pombe* centromeres appear to have properties similar to domains of heterochromatin in higher eukaryotes. Genes placed within fission yeast centromeres are transcriptionally repressed (Allshire *et al.*, 1994) and recombination across *cen1* has been shown to be suppressed (Niwa *et al.*, 1989). The main regions recognised as heterochromatic in higher eukaryotes are centromeres and telomeres (Lohe & Hilliker., 1995; Pluta *et al.*, 1995; Rattner, 1991; Singer, 1982). When inserted at any one of several sites tested within *S. pombe* centromeres the *ura4* or *ade6* genes are repressed for transcription (Allshire *et al.*, 1994, 1995). However this repressed state is reversible and results in a variegated phenotype similar to position effect variegation of *Drosophila* (Muller, 1931; Tartoff and Bremner, 1990). This variable gene expression is indicative of a heterochromatin-like structure being



**Figure 1.2 An *S. pombe* minimal centromere (approximately 17 kb)**



This 17 kb central region of *S. pombe* centromere1 was cloned intact in *E. coli* and shown to retain some centromere function for mitotic segregation. The K'' is essential for centromere function but only with additional centromere elements. The construct also contains an inverted repeat of part of B' around the central core of *cen1*.



formed at *S. pombe* centromeres. Mutations in a number of genes such as *rik1*, *swi6* and *clr 4* affect expression of genes inserted within *S. pombe* centromeres (Allshire *et al.*, 1995). These mutations were originally identified as being required to maintain the silent mating type loci (*mat2* and *mat3*) in a repressed state in *S. pombe* (Ekwall and Ruusala, 1994; Egal *et al.*, 1989; Thon *et al.*, 1994).

## 1.2(b) REPLICATION ORIGINS

Autonomously replication sequences (*ars*) of *S. pombe* were isolated and shown to allow plasmids to replicate as circular episomes in *S. pombe* and be maintained without rearrangement (Maundrell *et al.*, 1988). Estimates suggest that *ars* elements occur on average once every 20 kb in the *S. pombe* genome. It is not known whether all *ars* elements represent chromosomal origins of replications. It was difficult to identify an *S. pombe ars* because plasmids without an *ars* can transform *S. pombe* and be maintained at high frequency although such plasmids are frequently rearranged. Frequently these rearranged plasmids were found to have incorporated fragments of host chromosomal DNA which contain an *ars* element (Maundrell *et al.*, 1988).

Fragments with *ars* activity were cloned from *EcoRI* digested *S. pombe* DNA in a plasmid as part of an investigation studying transcription of the *URA3* gene (Losson and Lacroute, 1983). The smallest fragment identified was *ars1* of 2.1 kb. This *ars1* element is widely used on plasmids in *S. pombe*. When several more *ars* elements were cloned from *Sau3A* digested *S. pombe* DNA fragments with *ars* activity ranged from 0.8-1.8 kb (Maundrell *et al.*, 1988). Even with the average restriction digest fragments used to generate the clones being 200 to 300 bp, the smallest fragment containing a functional *ars* was 800 bp. These elements have a higher than usual AT



content (69-75%) and contain a consensus sequence of 5' (A/T)PuTTTATTTA(A/T) 3'. This consensus sequence however turns out to be unnecessary since it can be deleted from the *ars* on a plasmid without affecting *ars* activity. Assuming that these *ars* elements represent chromosomal origins of replication, the existence of this consensus sequence may be significant and of importance in a chromosomal context. Three additional *S. pombe ars* sequences have been identified which have large modular structure similar to, but more complex than, the *S. cerevisiae* *ARS* (Dubey *et al.*, 1996).

### 1.2(c) *S. pombe* TELOMERES

A linear plasmid with telomeres from the ciliate, *Tetrahymena thermophila*, a selectable marker and *ars1* was shown to transform *S. pombe* and the termini were healed by the addition of *S. pombe* telomere repeats (Szostak and Sugawara, 1986). *Tetrahymena* telomeres seeded new *S. pombe* telomere sequences, averaging around 300 bp in 1 of 11 transformants. Treatment of DNA from this transformant with Bal31 exonuclease, followed by ligation allowed a circularised form to be recovered in *E. coli*. The broken ends of *S. pombe* chromosomes were also shown to be healed with approximately 300 bp of the same telomeric repeat (Matsumoto *et al.*, 1987). Libraries were constructed from the ends of chromosomes and clones for all *S. pombe* telomeres recovered. Sequencing of these clones demonstrated that the terminal repeat arrays are formed from units with the consensus sequence T<sub>1-3</sub>AC(A)(C)G<sub>1-8</sub>. Telomere associated sequences of approximately 20 kb are adjacent to telomere repeats on chromosomes I and II, however, chromosome III has no telomere associated sequences but each telomere repeat is adjacent to a stretch of rDNA repeats (Sugawara, 1989; Mizakami *et al.*, 1993). Linear plasmids were constructed with a 258 bp tract of *S. pombe* telomere repeats on both ends and



transformed *S. pombe* (Nimmo *et al.*, 1994) therefore demonstrating that 258 bp of telomere repeats is sufficient to seed new telomere growth in *S. pombe*.

### **1.3 VECTOR SYSTEMS FOR CLONING LARGE DNA FRAGMENTS**

The study of complex higher eukaryotic genomes is difficult because some functional genetic units span several hundred kilobases and frequently contain a large number of different repetitive elements. Plasmid and cosmid vectors are capable of containing only small fragments of up to 40 kb. One problem encountered was that certain repetitive elements were found to be unstable in *E. coli* (Yokobata *et al.*, 1991; Kim *et al.*, 1992; Wyman *et al.*, 1985; Leach & Stahl, 1982). The development of YACs in *S. cerevisiae* provided a major research tool for the analysis of higher eukaryotic genomes (Burke *et al.*, 1987). YACs allow large fragments of DNA to be cloned and manipulated which is essential in modern mammalian molecular genetics. YAC technology assisted the physical isolation of genes involved in human genetic diseases, the study of whole genes in their genomic context and has facilitated the physical mapping of whole genomes (Monaco and Lorin, 1994; Cohen *et al.*, 1993). In addition YACs have been transformed or microinjected into mammalian cell lines and embryonic stem cells allowing functional analysis of genes within mammalian cell lines or in transgenic mice (Paven *et al.*, 1990; Huxley *et al.*, 1991; Schedl *et al.*, 1993; Schedl *et al.*, 1996).



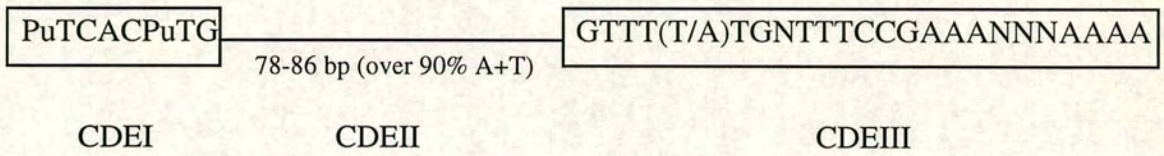
## 1.4 DEVELOPMENT OF *S. cerevisiae* YEAST ARTIFICIAL CHROMOSOMES

YACs were created in *S. cerevisiae* by incorporating replication origins, telomeres and centromere into a single linear DNA molecule. *S. cerevisiae* autonomous replication sequences (ARS) were isolated as fragments of DNA which had the ability to allow plasmids to replicate extrachromosomally (Stinchcomb *et al.*, 1979). These plasmids are maintained in multiple copies per cell and are mitotically unstable. *S. cerevisiae* ARS elements contain an 11 bp consensus sequence, (A/T)TTTATPuTTT(A/T) (Brewer *et al.*, 1987).

An *S. cerevisiae* centromere was cloned by its ability to confer mitotic stability on ARS containing plasmids (Murray and Szostak, 1983). These *CEN* plasmids were maintained as circular episomes but because of the mitotic stability provided by the centromere their copy number was reduced to approximately one molecule per cell. *S. cerevisiae* centromeres are small, only 125 bp of DNA is required to provide complete centromere function (Clarke, 1990; Carbon & Clarke, 1990). Unlike *S. pombe* centromeres, *S. cerevisiae* centromeres are devoid of repetitive sequences and all centromeres are interchangeable. Each *S. cerevisiae* centromere is composed of three centromere DNA elements, CDEs, as shown in figure 1.3 (Clarke, 1990; Carbon and Clarke, 1990). Genes are found within a few hundred bases on either side of the centromere therefore in total the *S. cerevisiae* kinetochore covers no more than 250 bp. These centromeres can function in either orientation and are not chromosome specific. *S. cerevisiae* centromeres are atypical since centromeres in other eukaryotes, including *S. pombe*, have complex repetitive array structures that span up to several megabases and unlike other eukaryotes, *S. cerevisiae* centromeres do not appear to be heterochromatic since transcription and recombination occur



**Figure 1.3 Consensus sequence of 125 bp functional *S. cerevisiae* centromere**



Consensus sequence of the functional *S. cerevisiae* centromere. This centromere contains three centromere DNA elements (CDE) and is structurally relatively simple with no repeated DNA sequences.



normally (Clarke, 1990; Carbon and Clarke, 1990). Heterochromatic like structures appear to be confined to the telomeres and silent mating type locus in *S. cerevisiae* (Louis, 1995; Laurenson and Rine, 1992; Aparicio *et al.*, 1991).

Linearised *ARS* plasmids with arrays of the (TTGGGG)<sub>n</sub> *Tetrahymena* telomeric repeats ligated on both ends were found to form linear episomes upon transformation of *S. cerevisiae*. Further analyses revealed that the (TTGGGG)<sub>n</sub> repeats acted as a seed allowing the addition of *S. cerevisiae* repeats. The *S. cerevisiae* telomere repeats were cloned from the ends of these plasmids and the telomere repeat sequence was found to be TG<sub>(1-3)</sub> and present at the ends of all natural yeast chromosomes (Szostak and Blackburn, 1982). These linear plasmid replicate at a high copy per cell and are mitotically unstable since they lack a centromere.

The first complete yeast artificial minichromosome constructed was a small linearised plasmid carrying *CEN3*, *Tetrahymena* telomeres and a replication origin, *ARS1* (Murray and Szostak, 1983). This was still mitotically unstable. When the size of the minichromosome was increased, by cloning 50 kb of lambda DNA between the arms, the resulting YACs behaved more like the endogenous chromosomes indicating that the size of YAC influences their mitotic stability.

Burke *et al.* (1987) constructed the first YAC vectors with plasmids which could replicate in *E. coli*. The vector included a replication origin (*ARS1*), a centromere (*CEN4*), selectable markers (*URA3* and *TRP1*) and telomere repeats, separated by a *HIS3* gene. The vector also included a cloning site within a *SUP4* gene, as the *SUP4* gene encodes an ochre suppressing tRNA which allows wildtype expression in an ochre *Ade2* host background. The cloning of an insert disrupts the *SUP4* gene



resulting in the transformed yeast colony being red instead of white (i.e. mutant phenotype). The first YACs developed contained inserts of around 200 kb.

## **1.5 THE YAC SYSTEM: ADVANTAGES AND DISADVANTAGES**

With the development of pulsed field gel electrophoresis, the full potential of the YAC libraries was realised. Pulsed field gel electrophoresis is a sophisticated system which uses a combination of voltage and switching of current direction to separate large fragments of DNA (Schartz and Canter, 1984). Fragments of DNA up to several megabases can be separated including the *S. pombe* chromosomes of 3.5, 4.7 and 5.7 Mb (Arshad *et al.*, 1993). The size of the inserts cloned into the YAC system has increased to Megabase proportions with the production of the human MegaYAC library (Evans *et al.*, 1993).

Improved genome coverage was evident when a contig was constructed to cover the 100 Mb genome of *C. elegans*. Cosmid contigs assembled for the nematode contained many gaps which were spanned using YACs which created larger contigs covering regions not previously cloned. The YAC library of *C. elegans* was estimated to be only 5% chimeric and was put in ordered arrays on filters and is now in use for assisting the DNA sequencing this organism's genome (Coulson *et al.*, 1991).

Human YAC libraries have been extremely useful in the study of large regions of chromosomes associated with disease loci (Monaco and Lorin, 1994). YAC contigs were constructed across these disease loci as defined by chromosomal abnormalities such as deletions, translocations and fragile sites. This positional cloning method has



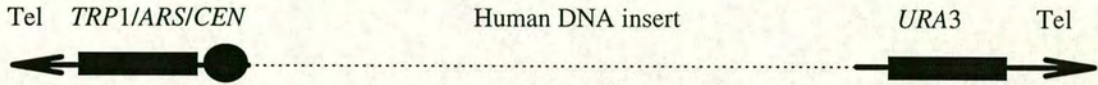
been used to identify the genes implicated in genetic diseases e.g. the CFTR gene in cystic fibrosis (Keren *et al.*, 1989; Riordan *et al.*, 1989; Rommens *et al.*, 1989).

YAC contigs have been assembled to map large chromosomal regions. Contigs of overlapping clones were first constructed for small chromosomal regions such as that of chromosome 21q (Chumakov *et al.*, 1992) and the euchromatic region of the Y chromosome (Foote *et al.*, 1992). This has led to the construction of the first generation map of the human genome. This combined the Centre d'Etude du Polymorphisme Humain (CEPH) MegaYAC library with approximately 2000 Généthon Markers (Cohen *et al.*, 1993). The YACs were ordered according to the Généthon 1993 genetic map which gave 30 % coverage of the genome. Additional MegaYACs were integrated into this map by pairwise clone linking derived from their fingerprints or using Alu PCR YAC products to screen the MegaYAC library. The MegaYACs were hybridised to Alu PCR products from monosomic somatic cell mapping, thus each MegaYAC was chromosomally positioned on a cytogenetic map. This integrated genetic and physical map reported by Cohen *et al.* (1993) was estimated to cover 90% of some chromosomes. A fully integrated genetic and physical map of the human genome made analysis of the human genome easier and allowed progress in the identification of disease genes.

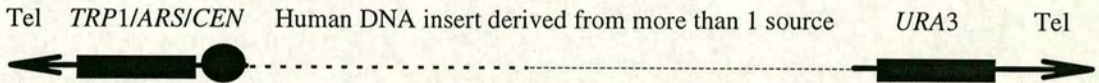
YAC libraries have undoubtedly been very useful in the study of higher eukaryotic genomes. However some problems have been encountered, i.e. chimerism and the instability of some inserts (figure 1.4; Monaco and Larin, 1994; Anderson, 1993; Green *et al.*, 1991; Albertsen *et al.*, 1990). Chimeric YACs have a non contiguous DNA insert derived from more than one location. This problem is particularly prevalent with mammalian libraries in which 40-60% of YACs have been reported to be chimeric. Originally co-ligation events were thought to be the cause of chimerism



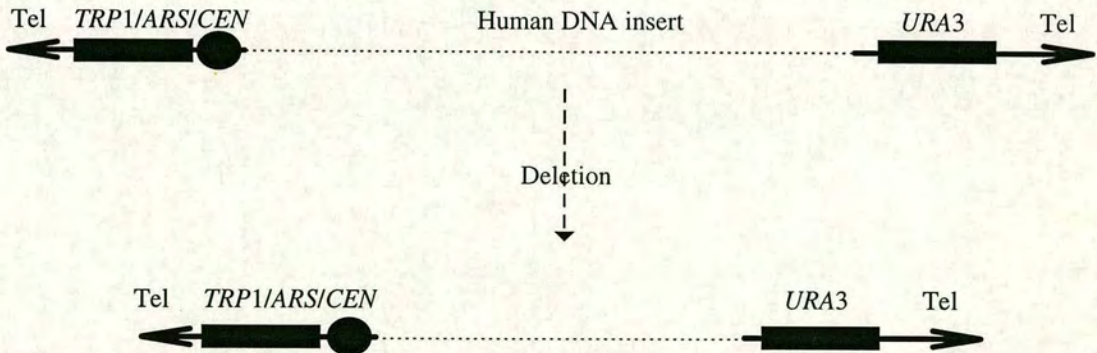
**Figure 1.4 (A) Structure of Yeast Artificial chromosome (YAC)**



**Figure 1.4 (B) Chimeric YAC**



**Figure 1.4 (C) Rearranging YAC**



(A) Structure of YAC clone with telomeres (Tel), yeast selectable markers (*TRP1* and *URA3*), an origin of replication (*ARS*) and centromere (*CEN*) and large insert of human DNA. (B) is the chimeric YAC with a non contiguous human insert. (C) An unstable YAC clone which has rearranged with the loss of human insert sequences. Adapted from Larin and Monoco, 1994.



but a study has shown that suppressing co-ligation events does not significantly decrease chimerism (Wada *et al.*, 1994). Many investigations now point to recombination between repeats as the source of the problem (Larinov *et al.*, 1994; Ling *et al.*, 1993; Green *et al.*, 1991). Most non contiguous YAC inserts can be screened out using Fluorescent *In situ* Hybridisation (FISH) onto mammalian cell metaphase spreads as chimeric YAC probes display more than one signal. Mapping the end clones from YACs can be used to determine any chimeric inserts outside the resolution of FISH. The higher resolution of this technique gives a higher estimate for chimerism than that determined by FISH analysis. In *S. cerevisiae* there are mutants which are deficient in mitotic recombination, such as *rad1* and *rad52* (Klein, 1988). *RAD1* and *RAD52* are involved in different types of recombination in *S. cerevisiae*. In a *rad1* mutant background the rate of recombination in the YACs is similar to wildtype. However, a *rad52* mutant background is shown to reduce the percentage of chimeric YACs in a library to around 10 % (Larinov *et al.*, 1994).

There is a problem of instability of some regions of higher eukaryote genomes when propagated in YACs (Monaco & Larin, 1994; Kouprina *et al.*, 1994; Anderson, 1993). Such instability generally involves internal deletions within YACs and causes poor representation of some regions of the genome in YAC libraries. This could be due to a large number of repetitive elements in particular regions of mammalian genomes that can recombine and cause deletions. Such deletions within YACs are more difficult to detect since part of the region of interest may be deleted although known adjacent markers may still be present. YAC instability appears to be related to recombination between Alu repeats (Green *et al.*, 1991), therefore the highly repetitive nature of the mammalian centromeres mean that such sequences are poorly propagated in YACs (Chartier *et al.*, 1992; Neil *et al.*, 1990). YACs with inserts from the centromere of human chromosome Y were highly rearranged with



substantial deletions. However three of the four classes of repeats in this centromere have been shown to be stabilised in a *RAD52* mutant while the fourth class remained unstable (Neil *et al.*, 1990). Therefore the study of mammalian centromeres has been hampered by the inability to clone these large repetitive domains intact on YACs. The centromere of the Y chromosome has been further analysed making use of 33 rearranged Y chromosomes which included rearrangements of centromeric repeats (Tyler-Smith *et al.*, 1993). Using these the minimal centromere was concluded to consist of 200 kb array of alphoid satellite repeats and 300 kb of other repeat sequences as defined by two deletion derivatives. Recently a YAC containing 70 kb of alphoid sequences has been shown to form centromeric structures (Grimes, personal communication). The importance of the higher order structure of these satellite repeat arrays in the formation of the functional centromere has yet to be fully understood.

The 17 natural chromosomes of *S. cerevisiae* range in size from 200 kb to 1600 kb (Mortimer *et al.*, 1989). YACs with sizes from 100 kb up to 2000 kb in *S. cerevisiae* frequently migrate with an endogenous chromosome making their detection and purification more difficult. The contamination of purified YACs with endogenous yeast chromosomes may interfere with further manipulations of YACs, for example, in transgenic experiments it is crucial to have as pure a sample as possible.

## **1.6 OTHER VECTOR SYSTEMS FOR CLONING LARGE DNA FRAGMENTS**

Cosmids have been used extensively and can contain fragments up to 40 kb. A cosmid is a modified plasmid carrying the bacteriophage lambda *cos* site. Cosmid



vectors are ligated to target DNA and molecules of 37-52 kb are packaged *in vitro* into phage heads by cleavage at the *cos* sites. These cosmid molecules can then be introduced efficiently into *E. coli* as phage where they are propagated as large circular plasmids. Many cosmid libraries are available for several different organisms. The size of the inserts limited the construction of contigs across entire genomes of higher eukaryotes as a large number of cosmids would be required to cover mammalian genomes (YACs reduced this number by around 10 fold). Some regions of higher eukaryote genomes appear to be unstable in cosmid libraries (Kim *et al.*, 1992; Yokobata *et al.*, 1991) however cosmids remain useful in allowing the isolation of smaller fragments from a region of interest.

Recently other *E. coli* systems such as the P1 cloning system (Sternberg, 1990), Bacterial artificial chromosomes, BACs (Shizuya, *et al.*, 1992) and P1-derived artificial chromosomes, PACs (Ioannou *et al.*, 1994) have been developed.

The P1 bacteriophage vector system was developed to carry up to 90 kb (Sternberg, 1990). This vector has been based on the same principles as the cosmids. The P1 vectors contain a *lox* site allowing the packaging of a particular amount of DNA equivalent to that normally found in a P1 phage. This vector bridges the gap between YACs and cosmids, however the size average of YAC libraries is usually around 500 kb and now there are MegaYAC libraries whose average size exceeds 1 Mb. There are several libraries for the P1 vector system commercially available for human (Sternberg, 1992) mouse (Pierce *et al.*, 1992) and *Drosophila* (Smoller *et al.*, 1991).

As a result of the problems experienced with some YACs, several groups have been investigating alternative methods for cloning large fragments of heterologous DNA. The bacteriophage systems were further exploited to produce two additional vector



systems which are capable of containing fragments of several hundred kilobases.

The BAC (Bacterial Artificial Chromosome) vector system makes use of the F origin of replication from *E. coli* that permits the replication of large DNA fragments which are propagated as a large circular plasmid maintained at a single copy per cell. These large circular DNA molecules can transform *E. coli* using electroporation. The first study proved that fragments up to 350 kb can be cloned in the BAC system (Shizuya *et al.*, 1992). Libraries of genomic DNA from a wide range of organisms have been constructed such as rice (Wang *et al.*, 1995), and human (Wang *et al.*, 1996, 1994; Kim *et al.*, 1995; Schmitt *et al.*, 1996).

The most recent system to be developed is the PAC (P1 derived Artificial Chromosome) vector system (Ioannou *et al.*, 1994). This vector has some features from both BACs and P1 vectors producing a vector with the ability to propagate fragments in the 100-300 kb size range. Cloning DNA into this vector disrupts a gene *SacB* whose gene product is toxic to cells in sucrose media, therefore the PAC system allows the positive selection for the presence of an insert. Like BACs, these PACs are also large circular DNA molecules which can transform *E. coli*, also by electroporation. A human PAC library has been constructed and twenty of these clones were analysed for insert rearrangements and the presence of chimeric inserts (Ioannou *et al.*, 1994). Analysis of these PACs after retransformation into bacteria has shown that there were few rearrangements of the human insert when ten strains were analysed. FISH analyses on these 20 PACs confirmed that their human inserts were contiguous (not chimeric).

These new vector systems are adaptations of systems already in use. Increased use of these systems will reveal any problems, it may be expected that the problems encountered when cloning arrays of repetitive sequences in a bacterial background



will remain. Human libraries of both vector systems have been used by Ashworth *et al.* (1995) in conjunction with human YACs in order to produce a complete contig of the human chromosomal region of 19q13.2.

The development of MACs should contribute to our understanding of many aspects of chromosome structure and function in mammalian cells (Farr *et al.*, 1991, 1995; Brown *et al.*, 1994; Heller *et al.*, 1996; Grimes, personal communication). In addition, MACs might eventually provide a useful vector for somatic gene therapy (Huxley, 1994). One approach towards constructing a MAC includes telomere associated chromosome fragmentation (TACF) (Farr *et al.*, 1991, 1995; Heller *et al.*, 1996). This method has produced minichromosomes of 4-10 Mb. However, recently it has been found that YACs of approximately 70 kb containing human chromosome 21 alphoid satellite DNA can form minichromosomes in some transformants and it is likely that the YAC has formed or incorporated into structures of 1-5 Mb (Grimes, personal communication).

## **1.7 POSSIBLE ADVANTAGES OF AN *S.pombe* ARTIFICIAL CLONING SYSTEM**

The construction of a cloning system in the fission yeast, *S. pombe* may provide a useful alternative and complement the existing vector systems. An *S. pombe* artificial chromosome system will require vectors with appropriate telomeres, selectable markers and replication origins to produce stable linear DNA molecules. The *S. cerevisiae* YAC system provides a useful model to guide the development of the *S. pombe* artificial chromosome system.



Burke *et al.* (1987) developed the YAC technology which has allowed the cloning of large DNA fragments as linear extrachromosomes in *S. cerevisiae*. The Human Genome Mapping Project has made advances, including the first integrated genetic and physical map, due to the development of human YAC libraries (Cohen *et al.*, 1993). As problems with rearrangements in YACs have emerged other vector systems, such as BACs (Shizuya *et al.*, 1992) and PACs (Ioannou *et al.*, 1994) have been developed. While these additional vector systems were being investigated, the *S. pombe* system was also being developed. The potential advantages of a *S. pombe* system are described below.

It is possible that sequences which are difficult to clone in *S. cerevisiae* may be propagated in the *S. pombe* vector system since repetitive sequences are more common in *S. pombe*. Arrays of repeats are present at the mammalian centromeres of the order of several hundred kilobases, however in YACs these arrays are highly unstable. To study repetitive regions e.g. the mammalian centromere, it would be advantageous to have a vector capable of cloning such arrays in a stable form. The reason for this instability is unknown but may be due to the repetitive sequences recombining and causing deletions.

The ability of the *S. pombe* vector system to maintain repetitive sequences is currently unknown. The *rec55-36* mutation with a 10 fold reduction in mitotic recombination was included in the strain used in the study presented in the hope that it might reduce the frequency of rearrangements (Gysler-Junker *et al.*, 1991). The construction of a mammalian library should give an indication of how well this vector system copes with mammalian DNA.



As *S. pombe* has three large chromosomes, any artificial chromosomes smaller than 3.5 Mb can be easily identified and isolated using pulsed field gel electrophoresis. One of the problems with YACs is that many of the artificial chromosomes comigrate with host chromosomes. Using *S. pombe* artificial chromosomes should give purer isolations of DNA for various manipulations, including transformation into mammalian cell lines.

## 1.8 FEASIBILITY OF CONSTRUCTING SPARC VECTOR SYSTEM

The essential components of *S. pombe* chromosomes have all been cloned and are available. However a problem arises when considering the use of an intact centromere because it is too large and complex to be propagated in *E. coli*. Intact *S. pombe* centromeres have been cloned in YACs (Hahnenberger *et al.*, 1989) or are carried on minichromosome derivatives in *S. pombe* (Niwa *et al.*, 1989).

In the absence of a useful, small functional centromere an S. pombe artificial chromosome (SPARC) vector will at least require telomeres, replication origins and selectable markers to carry large fragments of heterologous DNA. The fission yeast has to be transformed at a high enough transformation efficiency to enable a library to be constructed. Transformation of *S. pombe* using lipofectin reagent improved the ability of protoplasts to take up large molecules of DNA (Allshire, 1990). Lipofectin is a cationic liposome which has been developed to facilitate the uptake of DNA into mammalian cells (Felgner *et al.*, 1987, 1989). Liposomes are thought to function by binding to the DNA and the cell membranes thus aiding DNA transfer into cells. The original protoplast transformation efficiency was about  $1 \times 10^4$  colony forming units (cfu) per  $\mu\text{g}$  of plasmid DNA. Using lipofectin, this transformation efficiency was



increased to approximately  $1 \times 10^6$  cfu per  $\mu\text{g}$  of plasmid (Allshire, 1990). This efficiency of transformation should be sufficient for use with an appropriate cloning system.

The potential of the SPARC system for cloning large fragments of heterologous DNA will become apparent once developed and tested. The *S. pombe* system should complement the other vector systems currently used to clone large DNA molecules. The primary objective of this project was to determine whether the SPARC vectors could be utilised for cloning large fragments of heterologous DNA in *S. pombe*.



## **2 MATERIALS AND METHODS**



## 2.1 STRAINS

*Schizosaccharomyces pombe*:

SP813	h <sup>+</sup> N <i>leu1-32 ade6-210 ura4-D18</i>
BG47	h <sup>-</sup> <i>ura4-D18 ade6-M375 int::pUC8/ura4/ade6-L469 rec55-36</i> and
BG48	same as above except h <sup>+</sup> (Gysler-Junker <i>et al.</i> , 1991)
FY562	h <sup>-</sup> <i>ura4-D18 leu1-32 rec55-36</i>
FY561	h <sup>+</sup> <i>ura4-D18 leu1-32 rec55-36</i>

## 2.2 PLASMIDS, YACS & MINI CHROMOSOMES

pUR19	pUC19 with the <i>S. pombe ars1</i> and <i>ura4</i> gene cloned into the <i>NdeI</i> site (Barbet <i>et al.</i> , 1992)
pEN51	pUR19 and telomere cassette (Nimmo <i>et al.</i> , 1994)
pEN53	pRS305 (with <i>S. cerevisiae LEU2</i> gene and <i>S. pombe ars1</i> ) and telomere cassette (Nimmo <i>et al.</i> , 1994)

YACs with human inserts were selected by size from the ICI library (Anand *et al.*, 1991)

pSp(*cen1*)7L YAC with *S. pombe* insert containing *S. pombe* centromere 1 and

pSp(*cen1*)7L - Tel is the minichromosome derived from above YAC through added *S. pombe* telomeres (Hahnenberger *et al.*, 1991)

Cosmids (Hermanston *et al.*, 1992):

I - H11148 and II - J44 (from lawrist4 library);

III - U22, IV - U143, V - U447 and VI - U833 (from pWE15 library).



*Sau3A* Linkers and Primers:

D921 (23mer) 5'-GTCAAGAATTCTGTACCGTGGAC-3'

D922 (21mer) 3'- TCTTAAGCGATGGCACGTGCTAG-5'

### 2.3 MEDIA

YMM minimal media was used for *S. cerevisiae* YAC strains (McCormick *et al.*, 1990):

Per litre: 900 mls (autoclaved) 6g ammonium sulphate,

1.7g yeast nitrogen base without amino acids or  
ammonium sulphate,

22g casamino acids,

50mg each amino acid

100mls of filter sterile 20% glucose was added.

*S. pombe* media used as directed by Moreno *et al.* (1991):

YE Per litre: 5 g yeast extract

30 g glucose

YEA YE + Adenine (approximately 100 mg/litre)

ME Per litre: 30 g malt extract

250 mg amino acids supplements



EMMG	per litre:	Pthallic Acid	3.0g
		Di-sodium hydrogen orthophosphate	1.8g
		Glutamic acid	3.76g
		2% Glucose	20.0g
		50× salts	20 mls
		1000×vitamins:	1 ml
		10000×minerals:	0.1 ml

Adjust to pH 5.6 and autoclave.

50x salts	per litre:	MgCl <sub>2</sub> .6H <sub>2</sub> O	53.5 g
		Ca Cl <sub>2</sub> .6H <sub>2</sub> O	1.0 g
		KCl <sub>2</sub>	50.0 g
		Na <sub>2</sub> SO <sub>4</sub>	2.0 g
1000x vitamins per 100 mls:		Inositol	1.0 g
		Nicotinic acid	1.0 g
		Calcium pantothenate	0.1 g
		Biotin	1.0 mg
10000xminerals per 100 mls:		H <sub>3</sub> BO <sub>3</sub>	0.5 g
		MnSO <sub>4</sub> .4 H <sub>2</sub> O	0.52 g
		ZnSO <sub>4</sub> .7 H <sub>2</sub> O	0.4 g
		FeCl <sub>3</sub> .6 H <sub>2</sub> O	0.2 g
		H <sub>2</sub> MoO <sub>4</sub>	0.144 g
		CuSO <sub>4</sub> .5 H <sub>2</sub> O	0.04 g
		Citric acid	1.0 g
		KI	0.01 g

EMMG with 0.5% glucose (used to grow cells for protoplasting)



Amino acids added as required to EMMG to a final concentration of 75-100 mg/litre.

Preparation of media for agar plates with the addition of 20g/litre of Oxoid agar.

Regeneration of protoplasts: 1.2 M sorbitol added to plates.

Bacterial media (Sambrook *et al.*, 1989):

Terrific Broth per litre:	900 mls	Bacto-tryptone	12 g
		Bacto-yeast extract	24 g
		glycerol	4 ml
	100 mls	KH <sub>2</sub> PO <sub>4</sub>	2.31 g
		K <sub>2</sub> HPO <sub>4</sub>	12.54 g

Solutions are prepared and autoclaved separately then mixed together when required.

LB broth	per litre:	Bacto-tryptone	10 g
		yeast extract	5 g
		NaCl	5 g

Adjust pH to 7.2 and autoclave. LB agar was prepared with the addition of 20g/litre Oxoid agar.

## 2.4 PLASMID PREPARATION

The method of plasmid preparation is adapted from Sambrook *et al.* (1989).

Inoculate 100 mls of terrific broth (+ antibiotic selection) with single colony from a fresh plate or a fresh overnight culture. Incubate at 37°C overnight. Spin down in a 50 ml Falcon tube at 1000 x g for 5 minutes in a Sorvall centrifuge. The pellet was resuspended in 4.5 mls of solution I. Leave at room temperature for 10 minutes.



Add 10 mls of solution II and leave on ice for 5 minutes. Add 5 mls of ice cold solution III and leave on ice for a further 15 minutes. Spin at 1000 x g for 10 minutes in the Sorvall. Filter supernatant through muslin and precipitate DNA by adding 0.6 volume with propan-2-ol. Spin at 1000 x g for 10 minutes. Wash pellet in 70% ethanol and leave to air dry.

The pellet was dissolved in 400 ml of dH<sub>2</sub>O, 0.7g of caesium chloride added and allowed to dissolve, before adding 67 µl of 10µg/ml ethidium bromide. Aliquot 1.3mls of 65% CsCl solution into a 1.8 ml Beckman ultracentrifuge tube and layer the DNA mix underneath. Balance tube must be within 0.01g and tubes sealed. Spin in TL-100 Ultracentrifuge at 100,000 rpm, 2.30 hours at 20°C with 2 minutes acceleration and 5 minutes deceleration.

The band is carefully removed using a needle and 1 ml syringe and put into a 5 ml Falcon tube. The ethidium bromide is extracted with dH<sub>2</sub>O saturated butan-1-ol, and the DNA solution made up to 1 ml. Precipitate DNA with 3 volumes of 70% ethanol. Spin at 1000 x g for 10 minutes in sorvall centrifuge. Wash pellet in 70% ethanol and air dry. Resuspend in 200 ml of TE.

To prepare cosmid DNA the volume of solutions were doubled to extract larger quantities of DNA.

Solution I: 50 mM glucose, 25 mM Tris, 10 mM EDTA

Solution II: 0.2 M sodium hydroxide, 1% SDS

Solution III: 3 M potassium acetate pH 4.8



## 2.5 PREPARATION OF *S.cerevisiae* DNA IN AGAROSE PLUGS

DNA was prepared in agarose plugs as adapted from McCormick *et al.* (1990). An *S. cerevisiae* colony was grown to late log phase (approximately  $1 \times 10^8$  cell/ml) at 30°C in yeast minimal media (100 mls). The cells were pelleted and washed twice in 50 mM EDTA pH 7.5 at 0°C and resuspended in 3.25 ml of 50 mM EDTA pH 7.5 and warmed to 37°C. A 1% low melting point agarose was prepared in 0.125 M EDTA pH 7.5 at 42°C. Three mls of cell suspension was mixed with 5 mls of low melting agarose and 1 ml of solution I (SCE) and aliquoted into plug moulds. Once the plugs set, they were suspended into 5 mls of solution II and incubated overnight at 37°C. Solution II was then replaced with 5 ml of NDS solution with a final concentration 1 mg/ml Proteinase K and incubated at 50°C overnight.

Solution I (SCE): 1 M sorbitol, 0.1 M sodium citrate, 0.06 M EDTA pH 8.

Solution II: 0.45 M EDTA pH 9, 10 mM Tris-chloride pH 8, 7.5 %  $\beta$ -mercapthoethanol.

NDS solution (III): 0.45 M EDTA pH 9, 10 mM Tris-chloride pH 8, 1 % N lauryl sarcosine

## 2.6 PREPARATION OF *S. pombe* DNA IN AGAROSE PLUGS

Method for the isolation of DNA in agarose plugs as given by Moreno *et al.* (1991).

An *S. pombe* colony is grown to late log phase (approximately  $1 \times 10^7$  cell/ml) at 32°C in EMMG (100 mls). The cells are pelleted and washed twice in 50 mM EDTA pH 7.5 at 0°C. The cells are resuspended in 0.5 mls solution SP1 to approximately  $1 \times 10^9$  cells/ml and Zymolyase T100 is added to a final concentration



of 0.5 mg/ml. This cell suspension was incubated for 1-2 hours at 37°C then an equal volume of 1% LMP agarose (in 0.125 M EDTA pH 9) at 42°C was added. This mixture is then aliquoted into 100 µl plug moulds. Once the plugs have set, the plugs are suspended NDS solution (as above) with Proteinase K added to a final concentration of 0.5 mg/ml and incubated overnight at 50°C. The NDS solution is replaced with fresh NDS with Proteinase K and left for a further 24 hours incubated at 50°C.

SP1: 1.2 M sorbitol, 50 mM sodium citrate, 50 mM sodium phosphate, 40 mM EDTA pH 8

## **2.7 PREPARATION AND TRANSFORMATION OF *S. pombe***

This method was modified from Allshire (1992). An important component of this method is the lipofectin reagent which is a positively charged liposome-forming molecule (*N*-[1-(2,3-dioleoyloxy)propyl]-*N,N,N*-trimethylammonium chloride) known to facilitate the uptake of DNA during the transformation of *S. pombe* cells (Allshire, 1990) and of mammalian cells (Felgner *et al.*, 1987, 1989).

Inoculate 5 ml of YEA or EMMG (+ required amino acid supplements) with healthy fresh growing colony. Incubate at 32°C (or permissive temperature) overnight. Add 0.5 ml and 1 ml of this overnight culture into 2 separate flasks containing 100 mls of EMMG + 0.5% glucose + required amino acid supplements. Incubate cultures at the permissive temperature for 12-20 hours until density is between  $1.0$  and  $2.0 \times 10^7$  cells/ml. Harvest cells by centrifugation for 5 minutes at  $1000 \times g$  in two 50 ml falcon tubes. Resuspend cell pellets in 5 ml of SP1 and combine to a single tube.



Add 20 ml of  $\beta$ -mercaptoethanol, mix and incubate at room temperature for 10 minutes, pellet cells and resuspend in 2 mls of SP2, add 5-10 mgs of Novozym 234 (use a face mask and/or fume hood when weighing out Novozym because repeated exposure can cause an allergic reaction) and incubate at 37°C until 70-90% of the cells are spherical. This can take between 20 and 30 minutes, protoplast transformation is monitored by observing aliquots under the microscope every 10 minutes. Add 10 ml of SP3 to the protoplasts, mix by inversion, pellet protoplasts by centrifugation at 600 x g for 5 minutes. Resuspend pellet in 10 ml of SP3 and spin again as above and repeat this wash once more. During these washes the remaining cells form protoplasts. To resuspend the protoplast pellet it is best to use a pipette to resuspend in a small amount of solution first and then add remainder (the cells are fragile therefore should not be vortexed). After washes with SP3 resuspend the final pellet in 1 ml of SP4.

Aliquot 100  $\mu$ l of protoplasts into separate 4 ml sterile Falcon (2070) tubes. Add transforming DNA and mix gently by hand and incubate at room temperature for 15 minutes. Add an equal volume of SP4 containing 66 mg/ml of lipofectin (final concentration 33 mg/ml) and incubate at room temperature for 15 minutes. Add 1 ml of SP5, mix by inversion and incubate at room temperature for 15 minutes, by this stage cells are phase dark and clump.

Pellet cells at 350 x g for 5 minutes. Carefully remove supernatant with a pipette. Resuspend pellet in 100-1000  $\mu$ l of SP4 (cells will now be wrinkled in appearance). Pipette aliquots of each transformation onto EMMG + 1.2 M sorbitol (+ required amino acid supplements) agar plates. When plating less than 50  $\mu$ l, pipette into 100  $\mu$ l of SP4 already on the plate. Cells should be spread over the plate very gently



using a glass spreader (bent pasteur pipettes are adequate). Do not lean too heavily on the plate with the spreader since protoplasts are very fragile.

Transformants can be streaked out on fresh EMMG plates without added uracil or leucine.

Protoplasting solutions:

SP1: 1.2 M sorbitol, 50 mM sodium citrate, 50 mM sodium phosphate, 40 mM EDTA pH 8;

SP2: 1.2 M sorbitol, 50 mM sodium citrate, 50 mM sodium phosphate;

SP3: 1.2 M sorbitol, 10 mM Tris-HCl pH 8;

SP4: 1.2 M sorbitol, 10 mM Tris-HCl pH 8, 10 mM CaCl<sub>2</sub>:2H<sub>2</sub>O;

SP5: 20% Polyethyleneglycol Mol. Wt. 3500-4000, 10 mM Tris-HCl pH 8, 10 mM CaCl<sub>2</sub>:2H<sub>2</sub>O.

## 2.8 20 ml *S.pombe* DNA PREPARATION

This method has been adapted from Holm *et al.* (1986). A colony of cells was grown in 20 mls of selective media (EMMG) overnight at 32°C in shaking incubator.

Pellet cells, resuspend in 1.5 mls SP1 and add 15 µl β-mercaptoethanol and incubate at room temperature for 20 minutes. Pellet cells and resuspend in 1ml SP1 and 20 µl of 10 mg/ml Zymolyase 100T and incubate for 1-2 hours for 37°C. Pellet cells and resuspend in 0.1 ml SP1 (as above) plus 20 µl 20 mg/ml Proteinase K, 0.9 mls 1% sarcosyl, 0.1 M EDTA, 0.1 M Tris-HCl pH 8, mix by inversion (do not vortex) and incubate at 60-65°C for at least 1 hour. Add 500 µl phenol/chloroform, mix and spin for 5 minutes, remove approximately 0.8-0.9 ml of the supernatant and repeat, then



follow with 2 chloroform treatments. Precipitate DNA by addition of  $1/10$  volume 3 M sodium acetate pH 5.2 and 0.6 volume isopropanol. Spin at high speed in microcentrifuge for 10 minutes. Wash pellet in 70% ethanol, dry and resuspend in 100  $\mu$ l of TE with 0.1 mg/ml RNase.

## **2.9 RESTRICTION ENZYMES DIGESTS**

Restriction analyses were performed using restriction enzymes from Boehringer Mannheim according to manufacturers specifications. For digestion of DNA in agarose plugs, the plugs are washed at least four times in TE solution (including an overnight wash), then equilibrated in buffer equivalent to the recommended buffer for that restriction enzyme with two washes in that buffer. The enzyme reaction mixture is prepared on ice in a volume sufficient to cover the plugs. The reaction is incubated at the restriction enzymes optimum temperature overnight approximately 30 units/ $\mu$ g DNA.

## **2.10 GENETIC CROSS**

The crosses are carried out as described in Moreno *et al.* (1991) using malt extract media to initiate nutrient starvation conditions required for conjugation and sporulation of *S. pombe*. The cross is prepared with an average sized colony from each of the  $h^+$  and  $h^-$  strains mixed in sterile distilled water. Once dried, the plate is incubated at 25°C (as higher temperatures affect conjugation) and after 2-3 days spores can be seen. The spores are released when mixed into 1 ml sterile distilled



water with 5 µls of glucylase. The spores are plated out and the resulting colonies tested on selective media for their phenotype.

## **2.11 GEL ELECTROPHORESIS**

Pulse field gel electrophoresis (PFGE) was carried out using a BIORAD CHEF-DR II. Routine gels were 1% agarose in 0.5xTBE whereas preparative gels were 1% low melting point (LMP) agarose in 0.5xTAE. The running conditions for each gel are given in the figure legends. Preparative gels were used to isolate YAC or SPARC DNA in a LMP agarose gel slice which was cut from a wide track on the gel, containing around 6 to 10 plugs, of which only the edges of each track was stained with ethidium bromide. The high molecular weight DNA bands were detected and marked using a clean coverslip to nick their position which was used to direct the isolation of the band. This gel slice was stored in 50 mM EDTA pH 8.

Conventional gel electrophoresis used 1% agarose in 0.5 x TAE buffer at 100 volts for a few hours or 30 volts overnight.

## **2.12 SOUTHERN BLOTTING**

The gel was processed with the Nic program (60 mJoules) in the BIORAD Gene Linker which nicks large pieces of DNA as directed in manufacturers guidelines. The Southern transfer and hybridisation were carried out as in Sambrook *et al.*, 1989. The gel was denatured for approximately 40 minutes, followed by approximately 40 minutes in neutralise solution.



In chapter 3 nitrocellulose was used as the blotting membrane. After blotting, the nitrocellulose was baked at 80°C under vacuum for 1 hour. In chapters 4 to 6 the membrane used was Genescreen which was used as directed by the manufacturers. The DNA was crosslinked to the membrane using 150 mJoules in the BIORAD Genelinker.

Filters were pre-hybridised for 30 minutes in approximately 25 mls of Church and Gilbert's hybridisation mix (Church & Gilbert, 1984) at 65°C in a rotating oven or an oven with a shaking platform. Probes were made by random priming to linear, denatured (i.e. single stranded) DNA using a Bcl random prime kit as directed. The probe was labelled by the incorporation of ( $\alpha$ -<sup>32</sup>P)dCTP (approximately 3000 Ci/mM) for 30 minutes to 1 hour at 37°C. Once incorporation was checked ( to be greater than 50%) the probe was separated on a Pharmacia NICK column then denatured using a boiling water bath and added to the pre-hybridisation bag which was resealed and hybridised at 65°C overnight. The membranes were then washed 2x30 minute washes in 0.2xSSC, 0.1% SDS at 65°C. Autoradiograph film was stored at -70°C overnight and the film developed in an X-ray developing machine or a phosphor-image cassette was used and the image scanned and processed using the Imagequant program.

Fragments of DNA for use as probes were prepared using the GeneClean kit or  $\beta$ -agarase. The probe used in figure 6.5 was prepared from an inter Alu PCR probe derived from the SPARC human insert (Arveiler and Porteous, 1992). The probe was prepared as above with approximately 1000 Ci/mM ( $\alpha$ -<sup>32</sup>P)dCTP. The probe was preannealed with cot1 DNA following manufacturer's instructions.



Denature solution: 1.5M NaCl, 0.5M NaOH

Church & Gilbert: 0.25M Na<sub>2</sub>HPO<sub>4</sub>, 7% SDS with added 0.5% marvel and  
100mg/ml Denatured Salmon Sperm DNA

### 2.13 CATCH LINKER PCR

The probe is prepared from a catch linker PCR of SPARC inserts (Shibasaki *et al.*, 1995). The transformant DNA is subject to pulse field gel electrophoresis in a low melting point (LMP) gel and the SPARC band excised in LMP gel slice. The gel slices are stored in high salt buffer. To digest the DNA a 50-100 µl portion of the gel slice (approximately 10-100 ng DNA) was equilibrated in restriction buffer for one hour then change the buffer for 2 further washes then set up on overnight restriction digest with an excess of enzyme (approximately 50 units of *Sau3A*). Heat to 65-70°C for 15 minutes to inactivate enzyme, cool to 37°C, add ligase buffer, 25 units of ligase and linkers (100 ng each of D921 and D922) and ligate at 15°C overnight. Melt the gel slice at 65°C and add 10 µl aliquot to 50 µl PCR reaction (5 µls 10 x TAPs buffer, 2 mM NTPs, 250 ng/ul D921, 0.3 µl 5 units/µl *Taq* polymerase).

The PCR program: 1 cycle: 94°C - 3 minutes,  
60°C - 1 minute,  
72°C - 1 minute;  
30 cycles: 94°C - 45 seconds,  
60°C - 1 minute,  
72°C - 1 minute;  
1 cycle: 94°C - 45 seconds,  
60°C - 1 minute,  
72°C - 10 minutes.



10x TAPS buffer (Hultman *et al.*, 1991): 250 mM TAPs HCl, pH 9.3, 500 mM KCl, 20 mM MgCl<sub>2</sub>, 1% Tween 20.

## 2.14 PREPARATION OF BIOTINYLATED PROBES

The probe is prepared by labelling the catch linker PCR products with biotin-11-dUTP by nick translation as described in Fantes *et al.* (1992). The probe was separated from the free nucleotides through a Sephadex G50 spin column prepared in a 1 ml falcon syringe blocked with glass wool then filled with Sephadex G50. Once prepared the excess liquid can be spun out at 200 x g in Sorval centrifuge. The probe was spun through the column and collected in an eppendorf.

Stratagene nitrocellulose gridded circles were washed in dH<sub>2</sub>O briefly, then soaked in 20x SSC for 10 minutes and air dried. Spot onto filter 10<sup>-2</sup> and 10<sup>-3</sup> dilutions of each reaction, in addition 1, 2 10 and 20 pg standards of appropriately labelled lambda DNA. The DNA on this filter was crosslinked using the program C4 in the BIORAD Stratalinker (250 mJoules). The filter is then washed in buffer 1 (0.1 M Tris, 0.15 M NaCl) for 5 minutes, incubated for 30 minutes at 37°C in buffer 1 + 3% bovine serum albumin (BSA) fraction 5, then incubated in 10 ml of buffer 1 containing 10 µl of strepavidin-alkaline phosphatase for 30 minutes at room temperature. The filters are then washed 2 x 15 minutes in 200 mls of buffer 1, then washed for 5 minutes in 0.1 M Tris pH 9.5. The filter was placed in a hybridisation bag to which 5 mls of buffer 3 containing 2 drops each from bottles 1,2 and 3 of Vector BCIP/NBT kit. The filter was incubated in the dark for 2-4 hours. The



concentration of each probe was estimated from the standards with the optimum concentration for FISH of 10 ng/ $\mu$ l.

## **2.15 FLOURESCENCE *IN SITU* HYBRIDISATION**

The method for hybridisation and detection used was from Fantes *et al.* (1992). A volume of probe was dispensed (allowing 30 -50 ng of DNA per slide) with 1  $\mu$ g of Cot 1 DNA and 5  $\mu$ g salmon sperm DNA. Two volumes of ethanol were added and the probe was spun down under vacuum until they had precipitated and resuspended in 10  $\mu$ l of hybridisation mix. This was mixed well then spun to the bottom of the tube and the DNA left to dissolve into the hybridisation mix. The prepared slides were provided prepared by H. Morrison. The position of the cells was marked using a pencil, the slides placed into a metal rack which was incubated in 2 x SSC with 100  $\mu$ g/ml RNase at 37°C for 1 hour. Formamide denaturant was warmed to 45°C, then transferred to 70°C. The slides were washed quickly in 2 x SSC then dehydrated through 70%, 90% and 100% ethanol for 2 minutes each, then desiccated for 10 minutes using a vacuum pump. The slides were then warmed to 70°C in an oven for 5 minutes and placed into 70% formamide, 2 x SSC for 3 minutes at 70°C. They were transferred to ice cold 70% ethanol for 2 minutes, then through 90% and 100% ethanol, then desiccated using a vacuum pump. The probe was denatured at 70°C for 5 minutes, then transferred to 37°C waterbath to preanneal for 15 minutes before spotting on a pre-cleaned coverslip and the slide placed carefully on top. The slide was turned over and the edges sealed with rubber cement solution (Tip-Top). The slides were incubated in a covered tray floating in a 37°C waterbath overnight. Care has to be ensured that the slides are not allowed to dry out.



Blocking buffer (4 x SSC, 5% Marvel) was prepared and used to prepare the correct dilutions of antibodies and conjugates. The antibodies/conjugates used were avidin-texas red (vector) in a 1:500 dilution and biotinylated anti-avidin (BAA) in a 1:100 dilution. The dilutions of these antibody/conjugates were spun in an eppendorf for 15 minutes in a centrifuge in the cold room to remove any precipitate. The rubber solution is pulled off the slide but the coverslip is left to fall off in the first few washes. The slides were put through four 3 minute washes in 50% formamide, 2 x SSC at 45°C with periodic agitation. By this stage the coverslips should have fallen off. Wash four times 3 minutes in 2 x SSC at 45°C, followed by four 3 minute washes in 0.1 x SSC at 60°C. The slides are then transferred to a solution of 4 x SSC, 0.1% Tween 20. The slide was incubated with 40 ml of blocking buffer under a 22 x 40 mm coverslip for 5 minutes at room temperature ensuring no air has been trapped. Working one slide at a time the coverslip is removed and excess fluid is drained off. A 40 µl aliquot of the avidin-texas red is spotted onto the coverslip and the slide is placed carefully on top to pick up the coverslip. The slides were incubated in a moistened chamber at 37°C for 30-60 minutes, the coverslips removed and the slides given three 2 minute washes in 4 x SSC, 0.1% Tween 20 at 37°C. Then biotinylated anti-avidin was added to each slide as above followed by the washes. Finally, this was repeated for avidin-texas red. Once the final washes were complete. Dapi staining was added as 20 µl into 500 ml mountant (vectrashield) this and approximately 40 µl of this solution was added per slide.

The coverslips were sealed with rubber solution (Pang). The slides were viewed using a Ziess microscope. The images were captured and processed using a program written by digital scientific (Cambridge) for the Iplab Spectrum software.



**3 TRANSFORMATION OF *S. pombe* WITH *S. cerevisiae*  
YACS CONTAINING HUMAN DNA**



### 3.1 INTRODUCTION

Construction of an *S. pombe* artificial chromosome vector requires a centromere, telomere repeats and autonomous replication origins. Previously *S. pombe* minichromosomes had been derived from fragmented *S. pombe* chromosomes (Matsumoto *et al.*, 1987) or *S. cerevisiae* YACs (Hahnenberger *et al.*, 1991) which were healed by the addition of approximately 300 bp of telomere repeats. Prior to the development of suitable *S. pombe* vectors to construct artificial chromosomes, the initial investigations focused on the ability of *S. pombe* to propagate human sequences with minimal rearrangements. In the absence of an *S. pombe* system, *S. cerevisiae* YACs carrying human sequences were transformed into *S. pombe*. *S. cerevisiae* YACs can be propagated in *S. pombe* and selected for, although they may be mitotically unstable without *S. pombe* telomeres, replication origins and centromere. *S. cerevisiae* YACs can be maintained as linear extrachromosomal elements if the YAC ends are healed with 300 bp of *S. pombe* telomere repeats (Hahnenberger *et al.*, 1991). This healing process was found to occur at a frequency of 1 in 11 transformants (Szostak and Sugawara, 1986). The transformation of *S. pombe* by YACs with human inserts will serve as a useful indicator as to how intact human sequences are maintained in *S. pombe*. In addition to the possible rearrangements caused by recombination events within the human repeats, it has to be assumed that the YACs may be altered due to mitotic instability of the YAC as an extrachromosomal element. Any conclusions reached may assist in the construction of the SPARC vectors.



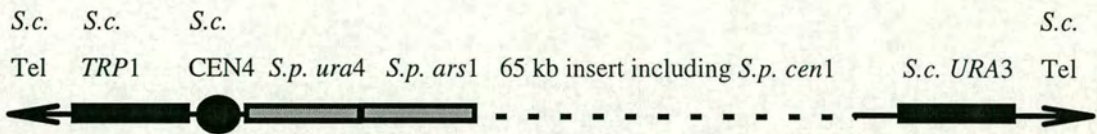
### 3.2 ISOLATION OF YAC DNA

Six YACs from *S. cerevisiae* and an *S. pombe* minichromosome were initially chosen for these investigations. Five of these YACs were selected by their size from an ICI human YAC library (Anand *et al.*, 1989). This library was constructed using pYAC4 vectors (Burke *et al.*, 1987) and *EcoRI* partially digested human DNA that was size fractionated to isolate fragments larger than 200 kb. The library transformed the *S. cerevisiae* strain AB1380 with an estimated efficiency of  $1 \times 10^4$  cfu/ $\mu$ g DNA. The sixth test YAC, pSp(*cen1*)7L (figure 3.1) was previously constructed by cloning *SalI* fragments of *S. pombe* DNA into pYAC4 to isolate *S. pombe* centromeres (Hahnenberger *et al.*, 1989). In this case the pYAC4 vector had been modified to include the *S. pombe ura4* gene as a selectable marker and *S. pombe ars1* as the replication origin to allow this YAC to be propagated in both *S. pombe* and *S. cerevisiae*.

Hahnenberger *et al.* (1989) demonstrated that when the pSp(*cen1*)7L YAC transformed *S. pombe*, one out of 20 transformants analysed had formed a stable linear minichromosome of 75 kb while in the rest of the transformants the YAC had been rearranged. Szostak and Sugawara (1986) had also estimated that the frequency of telomere healing to be 1 in 11 transformants when linear plasmids with *Tetrahymena* telomere repeats are transformed into *S. pombe*. It was therefore assumed that *S. pombe* telomere repeats had been added onto the ends of the pSp(*cen1*)7L YAC in *S. pombe* to produce a stable linear minichromosome (Hahnenberger *et al.*, 1989). This stable linear minichromosome propagated in *S. pombe* has been designated pSp(*cen1*)7L-Tel. During the course of my analyses, it has been shown by hybridisation with an *S. pombe* telomeric repeat that



**Figure 3.1 Structure of the pSp(cen1)7L YAC**



The pSp(*cen1*)7L YAC was based on the pYAC4 vector (Burke *et al.*, 1987) with *S. cerevisiae* telomeres (*S.c. Tel*), *S. cerevisiae* selectable markers (*S.c. TRP1* and *S.c. URA3*) and *S. cerevisiae* centromere (*S.c. CEN4*). The vectors were modified to include an *S. pombe* selectable marker (*S.p. ura4*) and *S. pombe* replication origin (*S.p. ars1*). The insert was a *SalI* fragment from *S. pombe* which included the 35 kb centromere 1.

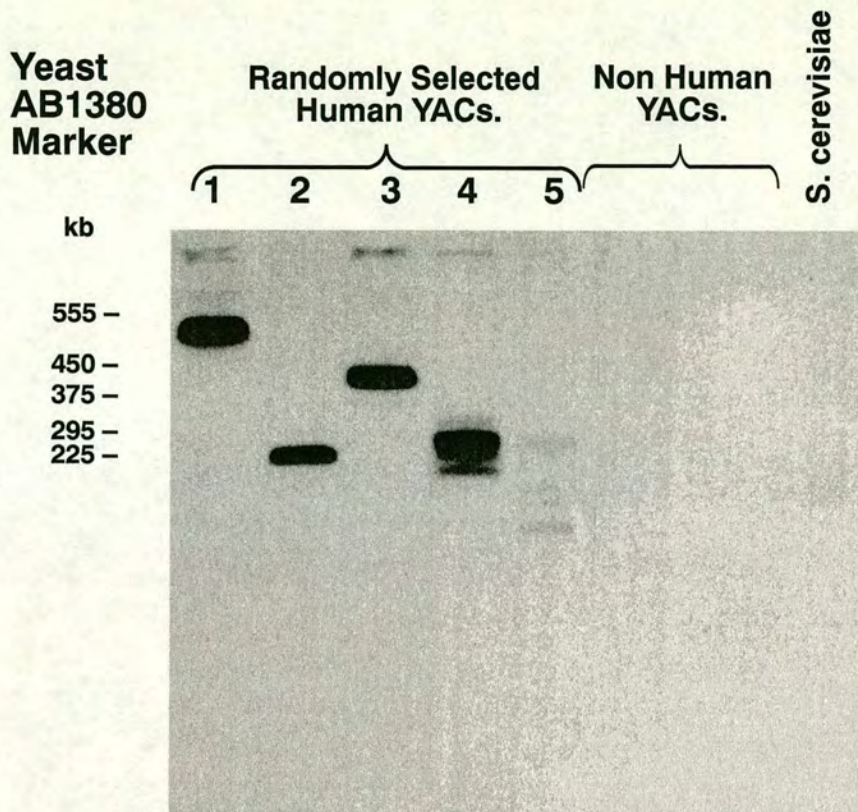


pSp(*cen1*)7L-Tel had gained *S. pombe* telomere repeats (data not shown). The pSp(*cen1*)7L YAC propagated in *S. cerevisiae* and its minichromosome version, pSp(*cen1*)7L-Tel which is propagated in *S. pombe*, were used as controls in the following experiments.

DNA was prepared in agarose plugs from each of the *S. cerevisiae* YAC strains and from the *S. pombe* strain containing the pSp(*cen1*)7L-Tel. These DNAs were subjected to pulsed field gel electrophoresis (PFGE) to separate the host chromosomes from the YAC and subsequently blotted to a nitrocellulose filter. Figure 3.2 shows an autoradiograph of the resulting blot with the five *S. cerevisiae* YACs containing human DNA with pSp(*cen1*)7L and pSp(*cen1*)7L-Tel as controls. The filter was first hybridised with digested total human DNA labelled with ( $\alpha$ - $^{32}\text{P}$ )dCTP. This probe detects human sequences by hybridising to repetitive elements present in abundance throughout the human genome. The first three lanes contain YACs 1,2 and 3 of 530, 225 and 360 kb respectively which were also visible on the pulsed field gel stained with ethidium bromide and viewed using an ultraviolet transilluminator (data not shown).

The other two YACs, in lanes 4 and 5 could not be seen by staining with ethidium bromide, however several bands in each lane are visible on the autoradiograph. This suggests that YACs 4 and 5 are highly unstable and undergoing rearrangement, therefore they were excluded from the following experiments. This serves to demonstrate one of the problems encountered using *S. cerevisiae* YACs. Lanes 6 and 7 contained pSp(*cen1*)7L YAC and pSp(*cen1*)7L-Tel minichromosome that are negative when probed with human DNA but were shown to give a single band





### Probe: Total Human DNA

#### Figure 3.2. Sizes of YACs

Autoradiograph of DNA from a pulsed field gel transferred to nitrocellulose membrane and probed with ( $\alpha$ - $^{32}$ P)dCTP labelled total human DNA. The approximate sizes of the *S. cerevisiae* chromosomes from strain AB1380 are shown. The first 5 lanes contained the DNA from randomly selected YACs in *S. cerevisiae* strain AB1380. The non-human YACs corresponds to the pSp (*cen1*) 7L YAC in *S. cerevisiae* and the pSp (*cen1*) 7L-Tel in *S. pombe* used as controls in this experiment. The pulsed field gel electrophoresis conditions were 150 volts, pulse switch time of 20 seconds to 70 seconds, 24 hours, in 0.5xTBE at 12°C.



of 75 kb when the blot was stripped and hybridised with ( $\alpha$ - $^{32}$ P)dCTP labelled plasmid DNA (data not shown). In the following experiments the *S. cerevisiae* human YACs 1, 2 and 3 were used, pSp(*cen1*)7L and pSp(*cen1*)7L-Tel served as controls.

To isolate the quantity of YAC DNA required for transformation of *S. pombe*, approximately 10 plugs of each YAC were loaded into wells (3-4 cm in length) on a low melting agarose pulsed field gels. After PFGE, the YAC DNA material was isolated by staining the outside edges of the wide lane with ethidium bromide and visualising under ultraviolet light the position of the YAC. With the position identified, the YAC DNA was isolated as a gel slice from the unstained portion. Pieces of the gel slice were washed in TE then melted at 65°C. Aliquots of the melted gel slice were used to transform the *S. pombe* strain, SP813, by the modified protoplast protocol. This method includes lipofectin (a cationic liposome) that had previously been shown to give improved transformation efficiency (Allshire, 1990). The transformation efficiency of *S. pombe* protoplasts was originally between 2- $3 \times 10^4$  per  $\mu$ g of plasmid DNA transformed into  $3 \times 10^7$  *S. pombe* protoplasts (Beach *et al.*, 1981; Beach *et al.*, 1982). The inclusion of lipofectin increases this to approximately  $7 \times 10^5$  transformants per  $\mu$ g of plasmid DNA. Transformation efficiencies close to  $1 \times 10^6$  cfu per  $\mu$ g of plasmid DNA were required to facilitate the transfer of YACs with large inserts into *S. pombe*. Therefore the inclusion of lipofectin in transformations of *S. pombe* protoplasts assisted the uptake of large DNA molecules. One problem encountered when preparing *S. pombe* protoplasts was the time taken for cells in log phase to be converted to protoplasts. By simply reducing the volume of the Novozym reaction from 10 ml to 2 ml it was consistently found that protoplast formation approached 90 % completion by 30 minutes.



### 3.3 TRANSFORMATION OF *S. pombe* WITH YACS

The transformation efficiency of *S. pombe* was consistently around  $1 \times 10^6$  cfu per  $\mu\text{g}$  of plasmid DNA. However, YAC uptake into *S. pombe* protoplasts was poor, with 1 to 5 colonies recovered per 10  $\mu\text{l}$  aliquot of melted gel slice (around 0.1-1 ng DNA). The inefficiency of YAC transformation maybe a result of the low concentration of YAC DNA in the gel slices. YAC DNA is most susceptible to degradation when at low concentrations during the melting of gel slices. Impurities in the water have been suggested as the cause of high molecular weight DNA being sheared during melting. Such degradation can be prevented if the gel slice is equilibrated in a melting buffer with a high salt concentration or by addition of low concentrations of polyamines (Larin *et al.* 1991). To test this, samples of each gel slice with DNA from YACs 1,2 and 3 were treated with TE, TE with polyamines (0.3 mM spermine, 0.75 mM spermidine) or TE with 100 mM NaCl. The treated gel slices were melted at 65°C and the samples loaded onto a pulsed field gel with the appropriate *S. cerevisiae* YAC run alongside for comparison. Figure 3.3(a) shows that treatment of all three YAC DNAs with TE alone resulted in their degradation, producing a smear which is lower than the 225 kb yeast marker. Treatment with 100 mM salt appears to give some protection from degradation although partial degradation is apparent for YACs 2 and 3. Protection with polyamines appeared to be particularly effective since YACs 1 and 3 remained intact. For YAC 2, there is clear evidence of a smear below the 225 kb yeast marker and no sign of a band at 225 kb under any conditions used.

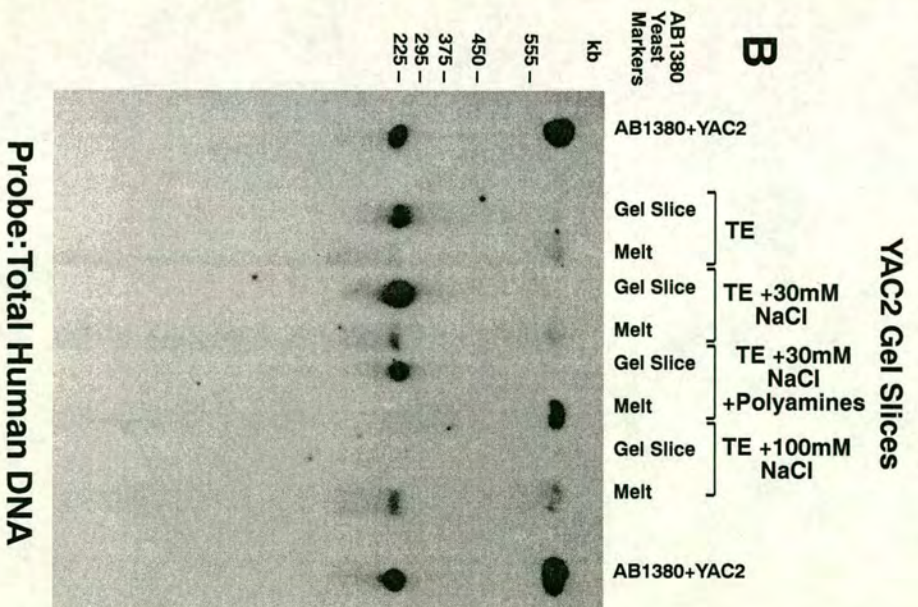
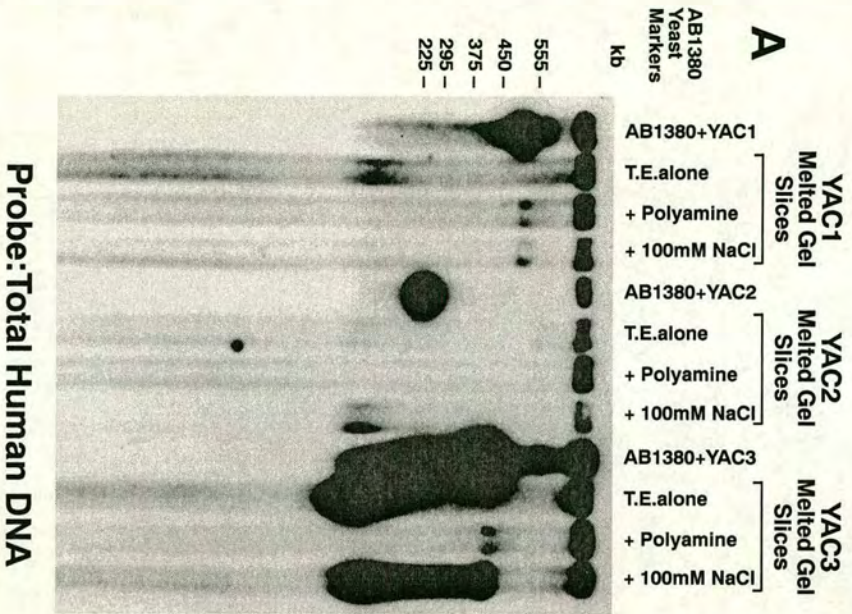




**Figure 3.3. Treatment of high molecular weight DNA to prevent degradation.**

Autoradiographs of pulsed field gels of DNA from melted agarose slices transferred to nitrocellulose membrane and probed with ( $\alpha$ - $^{32}$ P)dCTP labelled total human DNA. The sizes of the smaller *S. cerevisiae* chromosomes from the AB1380 strain are given on both autoradiographs. (A) YAC DNA after melting at 65°C after the pre-treatment of the agarose gel slice with TE alone or TE with polyamines (0.3 mM spermine, 0.75 mM spermidine) or TE with addition of 100 mM NaCl. Results for YACs 1, 2 and 3 are shown. The extent of DNA degradation after melting for each treatment is measured by comparison to the agarose embedded *S. cerevisiae* AB1380 strain containing the corresponding YAC. (B) YAC2 DNA from treated gel slices before and after melting at 65°C. Comparison of treating DNA with TE alone to TE plus 30 mM NaCl, or TE plus 30 mM NaCl and polyamines, or TE plus 100 mM NaCl. The pulsed field gel electrophoresis conditions used were 150 volts, 20 - 70 seconds, 24 hours, in 0.5xTBE at 12°C.







Refinements of these treatments were tested on a fresh sample of YAC 2 DNA to investigate the effect of treating the YAC DNA with 30 mM NaCl, 100 mM NaCl or a combination of 30 mM NaCl with polyamines (0.3 mM spermine, 0.75 mM spermidine). Figure 3.3(b) shows the DNA in the treated gel slices before and after melting. Compared to treatment with TE alone where the YAC DNA was degraded, more of the DNA remained intact when treated with either 30 mM or 100 mM NaCl. However the treatment with a combination of 30 mM NaCl and polyamines resulted in all the YAC DNA being trapped in the well. In this situation the DNA has become bound together due to the combination of salt and polyamines in the equilibration buffer. In some of the other lanes a signal can also be seen from residual DNA trapped in the wells which is probably the large DNA fragments in the sample that have become entangled.

It can be concluded that both the polyamines and addition of salt can help to prevent the degradation of large fragments of DNA. Treatment with polyamines is particularly useful for protecting YAC DNA isolated in low concentration and when the samples do not contain added sodium chloride. The addition of polyamines was used to protect the DNA for the YAC transformation experiments. In all subsequent experiments the extent of DNA degradation was assessed for each transformation by running a sample of the melted gel slice on a pulsed field gel.

### **3.4 ANALYSIS OF TRANSFORMANTS**

The *S. cerevisiae* *URA3* gene can complement mutations in the *S. pombe* *ura4* gene, allowing growth of *ura4*<sup>-</sup> strains on plates lacking uracil (Russell, 1989), therefore the *S. cerevisiae* YACs bearing the *URA3* marker can be transformed into and



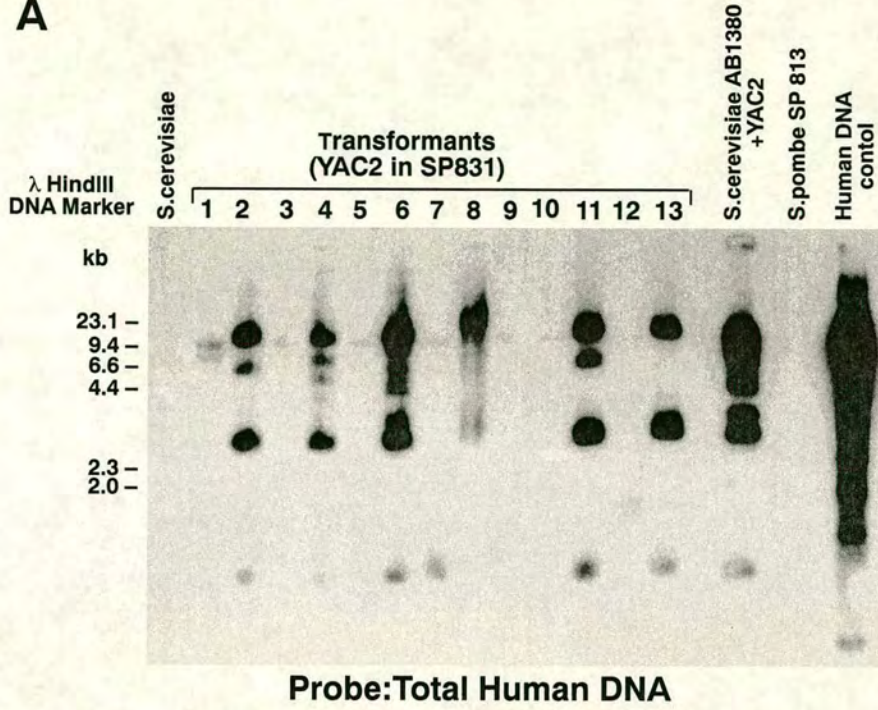
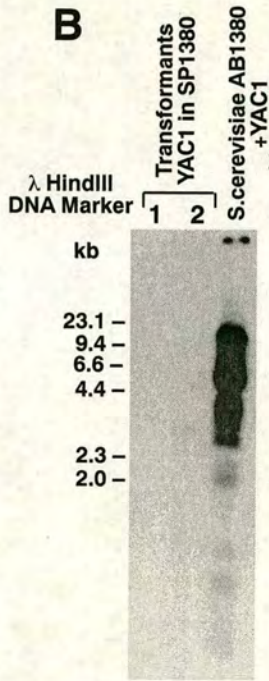
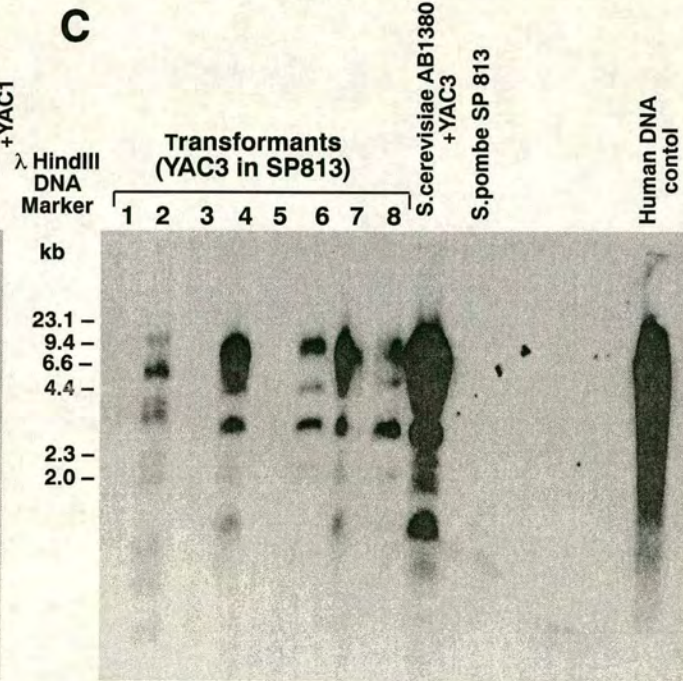
selected for in an *S. pombe* strain with a *ura4* deletion. The gel purified YAC DNA was demonstrated to be intact prior to transformation into *S. pombe* as an aliquot of this agarose was subject to pulsed field gel electrophoresis. Additional aliquots of the same molten agarose slices were used to transform *S. pombe* SP813 protoplasts. Transformants containing YAC DNA were selected on plates lacking uracil and several of these were selected for further analyses. DNA was prepared as agarose plugs from several transformants of each YAC and analysed by Southern blotting. The filter was probed with ( $\alpha$ - $^{32}$ P)dCTP labelled total human DNA. Figure 3.4 shows a Southern blot of *Eco*RI digested DNA of 13 transformants with YAC 2, 8 transformants with YAC 3 and 2 transformants with YAC 1. Figure 3.4(a) shows the pattern of human DNA *Eco*RI fragments in the 13 YAC 2 transformants compared to that of the DNA digest of the original YAC 2 DNA. Two of the transformants, 4 and 6, have the same pattern of bands as the original YAC. Transformants 2, 8, 11 and 13 have retained a few of the bands present in the YAC 2 profile. In approximately half of the transformants (lanes 1, 3, 5, 7, 9 and 10) many of the human bands have been lost with only one or two retained. Transformant 12 contains very little human DNA and one of the bands (of approximately 1 kb) does not correspond to a band within the YAC 2 profile. Therefore 2 of the 13 transformants analysed appeared to have retained YAC 2 with the human DNA insert unaltered.



**Figure 3.4. Analysis of DNA from *S. pombe* transformed with human YACs**

Autoradiographs of DNA digested with *Eco*RI subject to conventional gel electrophoresis, transferred to nitrocellulose membrane and probed with ( $\alpha$ - $^{32}$ P)dCTP labelled total human DNA. (A) Comparison of *Eco*RI digests from 13 *S. pombe* transformants containing YAC2 to *Eco*RI digest of *S. cerevisiae* strain AB1380 with YAC2. (B) Comparison of *Eco*RI digests from 2 *S. pombe* transformants containing YAC1 to *Eco*RI digest of *S. cerevisiae* strain AB1380 with YAC1. (C) Comparison of *Eco*RI digests from 8 *S. pombe* transformants containing YAC3 to *Eco*RI digest of *S. cerevisiae* strain AB1380 with YAC3. The sizes of bands are shown for the  $\lambda$  *Hind*III marker. In (A) and (C), positive and negative control lanes are shown as the human DNA and SP813 *S. pombe* strain DNA respectively.



**A****B****C****Probe: Total Human DNA**



Transformation of *S. pombe* with YAC 1 produced only two transformants, the digests of which were compared to that of YAC 1 in figure 3.4(b). Transformant 1 appeared not to have retained any human DNA while transformant 2 has a few of the human bands present in the YAC 1 profile. Therefore both of these transformants appear to have undergone rearrangement with very little of the human DNA remaining.

Digested DNA of YAC 3 was compared to the digests of the YAC 3 *S. pombe* transformants in figure 3.4(c). Here transformant 4 had a banding pattern of human fragments closely matching the YAC 3 pattern. Transformants 2, 6, 7, and 8 had many of the same bands seen in the YAC 3 profile. Transformants 3 and 5 had only a faint band of around 3 kb while transformant 1 had apparently no human DNA remaining. Therefore one of the eight transformants had the same repetitive human DNA fingerprint as the YAC.

Pulsed field gel electrophoresis was also performed on the undigested DNA from these transformants. Unfortunately attempts to detect human sequences in these transformants failed even though the Southern blots from conventional gel electrophoresis has shown human sequences to be present. It is possible that these YACs have become integrated into the endogenous *S. pombe* chromosomes which all migrate at 3.5 Mb or greater. Such large molecules may have failed to transfer under the conditions used. Regardless of the inability to detect human sequences by pulsed field gel electrophoresis, the conventional gel analyses clearly demonstrated that some or all the human sequences of the *S. cerevisiae* YACs transformed into *S. pombe* are frequently lost. These *S. cerevisiae* YACs did not contain a *S. pombe* centromere, replication origins or telomeres and therefore would not be expected to be maintained as linear episomes. It was important to determine whether this loss

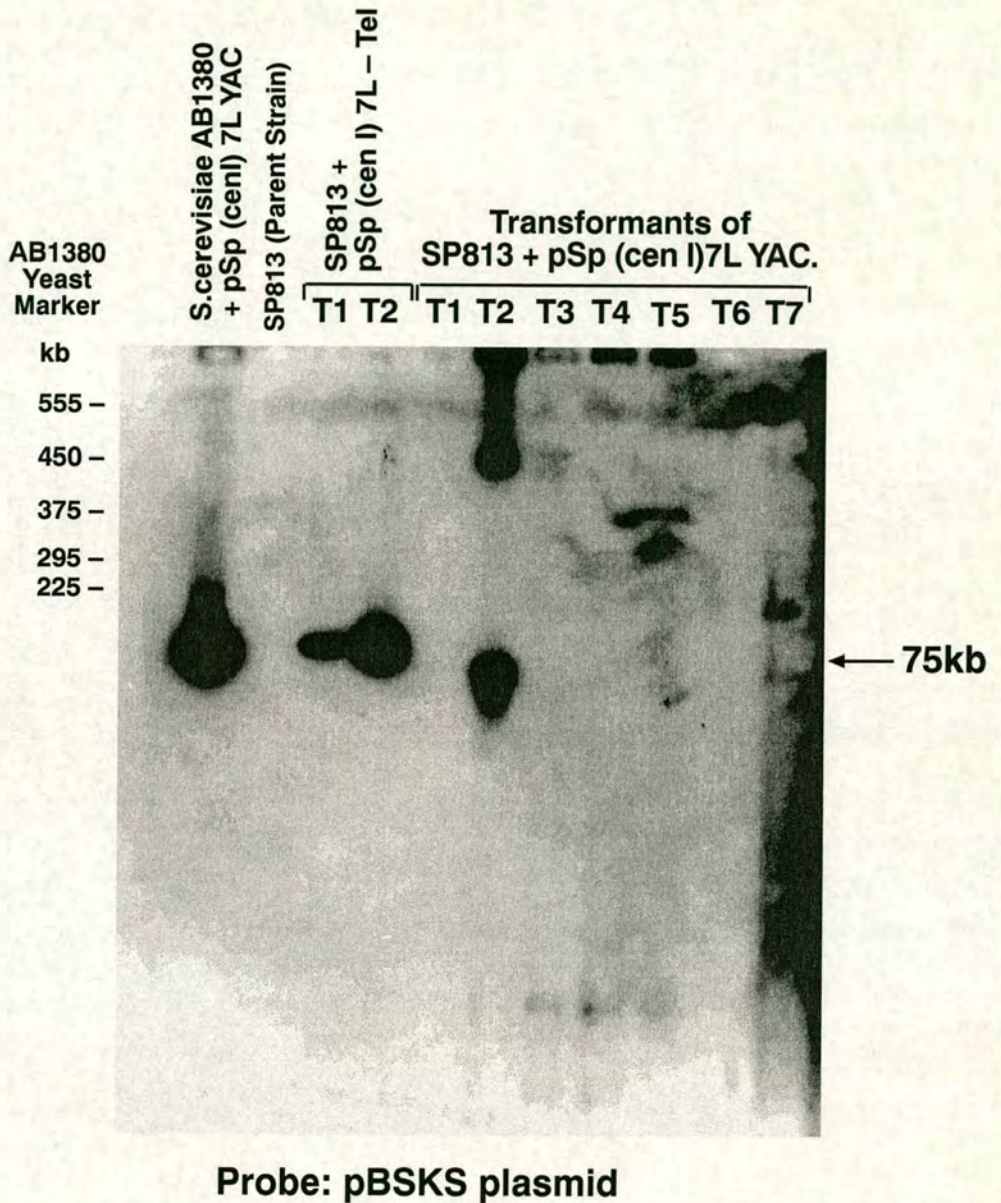


of DNA was due to an inability of *S. pombe* to cope with human sequences or results from the lack of *S. pombe* centromeres or telomeres on the vector. Control experiments were carried out with the pSp(*cen1*)7L YAC (which was prepared from *S. cerevisiae*) and the derivative pSp(*cen1*)7L-Tel (prepared from *S. pombe*) which were transformed into *S. pombe*.

Transformation of the SP813 strain by the control pSp(*cen1*)7L YAC and the pSp(*cen1*)7L-Tel occurred at a transformation efficiency of approximately 5-10 colonies per 10  $\mu$ l of molten agarose (around 1-10 pg of DNA). DNA from the resulting transformants was prepared and analysed on pulsed field gels. Figure 3.5 shows DNA from these transformants subject to pulsed field gel electrophoresis, then transferred to nitrocellulose and probed with ( $\alpha$ - $^{32}$ P)dCTP labelled plasmid DNA. In *S. cerevisiae* pSp(*cen1*)7L replicates as a linear episome of 75 kb. If *S. pombe* telomeric repeats were added to this YAC upon transformation into *S. pombe* then it would also be expected to migrate at 75 kb. Clearly this is not the case with the seven pSp(*cen1*)7L YAC transformants. The pSp(*cen1*)7L YAC in these transformants appear to have undergone a variety of rearrangements including multiple bands (as seen in 2 and 7) and smaller molecules (in 3, 4 and 5). Two of the transformants, 1 and 6, do not appear to have retained the plasmid sequences of the YAC vector arms. To test if the addition of *S. pombe* telomeres can prevent such rearrangements, DNA was prepared from the *S. pombe* strain containing pSp(*cen1*)7L-Tel which has *S. pombe* telomeres added to the YAC termini. The addition of *S. pombe* telomeric repeats is the only difference between this *S. pombe* version, pSp(*cen1*)7L-Tel, and the pSp(*cen1*)7L YAC. Shown in figure 3.5 are two of the resulting transformants containing pSp(*cen1*)7L-Tel retransformed into *S. pombe*, both of which contain a stable linear episome of the expected size of 75 kb.



Figure 3.5. DNA from *S. pombe* transformants with centromere YAC and minichromosome.



Autoradiograph of DNA from pulsed field gel transferred to nitrocellulose membrane and probed with ( $\alpha$ - $^{32}$ P)dCTP. DNA from 7 transformants from a transformation of the 75 kb pSp(*cen1*)7L *S. cerevisiae* YAC into *S. pombe* strain SP813 are shown. For comparison the pSp(*cen1*)7L YAC in *S. cerevisiae* AB1380 and 2 transformants of the 75 kb pSp(*cen1*)7L-Tel minichromosome from *S. pombe* strain SP223 transformed into *S. pombe* strain SP813 are also shown. DNA from the SP813 *S. pombe* strain serves as a negative control. The sizes of the yeast markers are shown for *S. cerevisiae* strain AB1380. The pulsed field gel electrophoresis conditions used were 150 volts, 20 - 70 seconds, 24 hours, in 0.5xTBE at 12°C.



Therefore, even in the presence of a functional *S. pombe* centromere, *S. cerevisiae* YACs which lack *S. pombe* telomere repeats become rearranged upon transformation into *S. pombe*.

Without *S. pombe* telomere repeats at their termini, YACs are unstable when transformed into *S. pombe* causing the rearrangements which resulted in the loss of DNA. Similar rearrangements have also been described when *S. cerevisiae* YACs are introduced into mammalian cell lines (Paven *et al.*, 1990; Huxley *et al.*, 1991). Again *S. cerevisiae* telomere repeats could not seed mammalian telomere repeats. It has been demonstrated that mammalian cells have stringent sequence requirements for seeding of new telomeres (Hanish *et al.*, 1994).

### 3.5 DISCUSSION

In this chapter, experiments have been described, which tested whether *S. cerevisiae* YACs containing human DNA inserts could transform *S. pombe* and be maintained in a stable form. Modifications to the *S. pombe* transformation protocol using lipofectin were made to improve the overall efficiency of the transformation procedure. Large fragments of DNA tend to become fragmented if heated to 65°C. Equilibrating agarose slices in a buffer containing polyamines or salt prior to melting protected the DNA from degradation. An improvement to the protoplasting technique, i.e. reduction of the volume of the Novozym reaction reduced the time required for the production of *S. pombe* protoplasts. Overall these modifications improved the transformation efficiency as the YAC DNA was intact and protoplasts were produced more reliably.



The human DNA containing YACs appeared to rearrange with loss of material when transformed into *S. pombe*. The pSp(*cen1*)7L YAC, which carries a 60 kb *SaII* fragment from *S. pombe* chromosome I, was also highly unstable producing several different rearranged forms. However pSp(*cen1*)7L-Tel, which has the same structure except for *S. pombe* telomere repeats added at each end, was stably maintained and could be detected as a linear band of 75 kb.

These results indicate that the absence of terminal telomeric sequences recognised by *S. pombe* largely contributed to the YAC instability. Since some of the transformants retained much of the human sequences without apparent rearrangement, these preliminary experiments indicated that it is feasible to propagate human sequences in *S. pombe*. It was clear that cloning large DNA fragments into *S. pombe* would require the development of a specific *S. pombe* vector system which includes *S. pombe* selectable markers, replication origins and telomeres. It is possible that a centromere would also be required but given their large size (a minimal centromere is at least 12 kb) this would be awkward and would make vector manipulation difficult.



## **4 CLONING COSMID FRAGMENTS IN SPARC VECTORS**



## 4.1 INTRODUCTION

An *S. pombe* vector system, like the *S. cerevisiae* YAC vectors, for cloning large fragments of DNA will require telomere repeats, replication origins and selectable markers. A centromere may not be essential for the stable maintenance of an artificial chromosome vector in *S. pombe*.

It was shown that transformation of *S. pombe* with *S. cerevisiae* YACs with human inserts, resulted in the substantial loss and or rearrangement of human sequences (chapter 3) and the *S. cerevisiae* pSp(*cen1*)7L YAC became rearranged when used to transform *S. pombe*. However the same YAC with *S. pombe* telomere repeats added to the YAC ends (pSp(*cen1*)7L-Tel, Hahnenberger *et al.*, 1989) was maintained as a stable, linear minichromosome. Therefore the instability of the *S. cerevisiae* YACs in *S. pombe* is apparently due to the lack of the appropriate *S. pombe* terminal repeats. *S. pombe* telomeric repeats are important in the formation of minichromosomes *in vivo*. Irradiation of *S. pombe* chromosome III resulted in derivatives where the damaged ends of the broken chromosomes were healed by the addition of 300 bp of telomere repeats (Matsumoto *et al.* 1987). Linear plasmids with 258 bp of *S. pombe* telomere repeats on both ends transformed *S. pombe* and maintained as linear episomes showing that these repeats are sufficient to seed and provide telomere function in *S. pombe* (Nimmo *et al.*, 1994). Therefore 258 bp of *S. pombe* telomere repeats should be sufficient to allow telomere formation *in vivo*.

The instability of the YACs in *S. pombe* may also be due to the differences in the replication origins between *S. cerevisiae* and *S. pombe*. To ensure their extrachromosomal maintenance any *S. pombe* vector requires the inclusion of *S. pombe* autonomous replication sequences. The most extensively used autonomous



replication sequence in *S. pombe* vectors is *ars1* isolated by Losson and Lacroute (1983).

Selectable markers such as the *S. cerevisiae* *URA3* and *LEU2* genes can be used to complement the corresponding mutations in *S. pombe* in *ura4* and *leu1* genes respectively. As shown in the last chapter the *S. cerevisiae* YACs with *URA3* can complement an *S. pombe* strain with a deletion of the *ura4* gene. Alternatively there are several *S. pombe* selectable markers like the *ura4*, *ade6*, *leu1*, *arg3* and *his3* genes which could be incorporated into the *S. pombe* vectors (Russell, 1989; Waddell and Jenkin, 1995; Burke & Gould, 1994).

*S. cerevisiae* YACs, in addition to *ars* elements, telomeres and selectable markers, carry a centromere which ensures that each cell receives a single copy of the YAC during mitosis. To include a centromere in a *S. pombe* YAC-like vector is not feasible as *S. pombe* centromeres have a complex inverted repeat structure.

Centromere 1 has the simplest structure and is the smallest at 35 kb. Efforts to map the minimal functional centromere sequences have resulted in the generation of several centromere deletion derivatives cloned in plasmids which retain reasonable mitotic function. The smallest of these constructs contains 10 kb of centromeric DNA including the central core of centromere 2 and the K' repeat of centromere 1 (Baum *et al.*, 1994). The addition of centromeric sequences to a *S. pombe* cloning vector would increase its size substantially, therefore making it difficult to manipulate. Experiments carried out by Matsumoto *et al* (1990) indicated that an acentric vector system might be tolerated in *S. pombe*. Two acentric minichromosomes were created which were shown to be maintained as stable, linear episomes of 390 kb and 130 kb at a high copy number. An increased copy number of *S. pombe* artificial chromosomes would be advantageous since this would allow more



of the cloned DNA to be produced per cell. A system has been developed which allows *S. cerevisiae* YACs to be modified to include a conditional centromere which can be inactivated thereby increasing the copy number of the YAC.

Thus an acentric *S. pombe* artificial chromosome cloning system will require telomeres, replication origins and selectable markers to create the *S. pombe* artificial chromosome vector system. There are two pre-existing plasmids which have all of these desired components. In this chapter, these vectors were tested to determine if they can be used to propagate human sequences in *S. pombe* on acentric linear episomes.

#### **4.2 *S. pombe* ARTIFICIAL CHROMOSOME (SPARC) VECTOR SYSTEM**

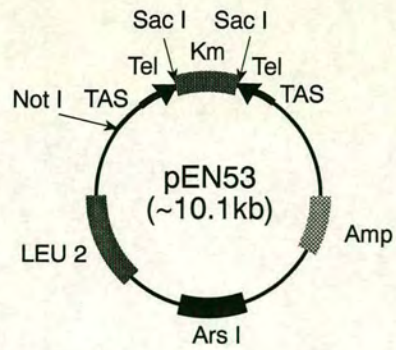
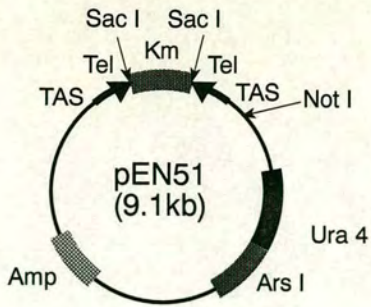
The linear versions of plasmids pEN51 and pEN53 have been used to show that 258 bp of telomere repeats can seed new telomere growth in *S. pombe* (Nimmo *et al.*, 1994). These plasmids contain the elements necessary for providing acentric vector arms to test the *S. pombe* cloning system. Both plasmids contain a telomere cassette consisting of 2 copies of a 1.1 kb fragment, which contains 258 bp of *S. pombe* telomere repeats plus telomere associated sequences, arranged in an inverted orientation around the kanamycin resistance gene (figure 4.1). Each plasmid has a copy of *S. pombe ars1* and a selectable marker (*S. pombe ura4* in pEN51 and *S. cerevisiae LEU2* in pEN53).



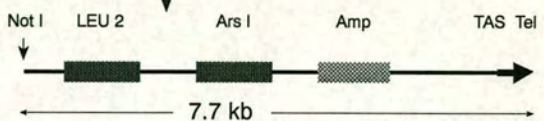
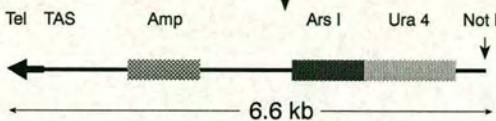
**Figure 4.1. Method used to clone human DNA fragments into SPARC vectors**

Human DNA *NotI* fragments were cloned into *S. pombe* using vector arms formed from pEN51 and pEN53. The products produced were acentric *S. pombe* artificial chromosomes (SPARCs).

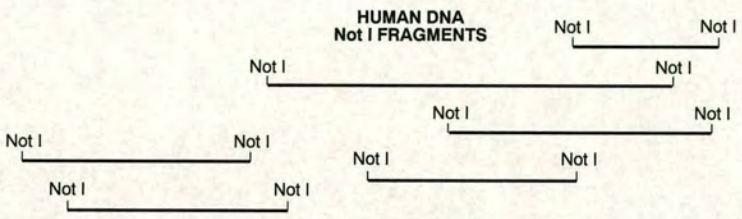




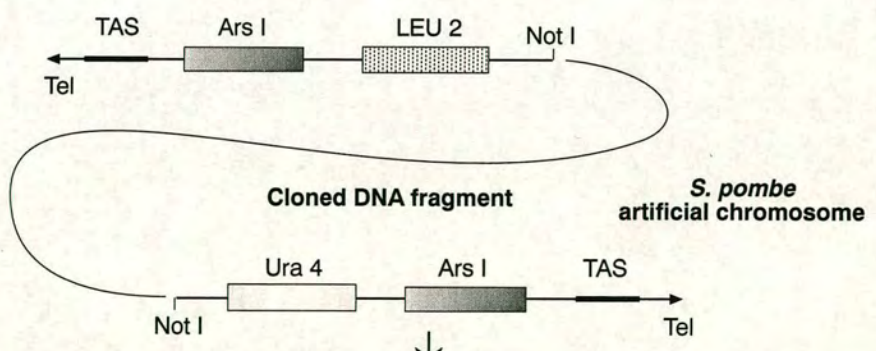
DIGEST WITH Sac I,  
PHOSPHATASE, DIGEST WITH Not I,  
GEL PURIFY THE VECTOR ARMS.



+



LIGATION



ISOLATE CORRECTLY SIZED LIGATION PRODUCTS ON LMP PULSE FIELD GEL

TRANSFORM INTO *S. pombe* PROTOPLASTS  
(*S. pombe* strains: SP813 - *ura 4-D18, leu 1-32, ade 6-210*  
FY562 - *ura 4-D18, leu 1-32, rec 55-36*)

SELECT FOR URA<sup>+</sup> LEU<sup>+</sup> TRANSFORMANTS

ANALYSES OF DNA FROM TRANSFORMANTS

KEY:

KM	- kanamycin resistance gene
AMP	- ampicillin resistance gene
TEL	- telomere (300bp)
TAS	- telomere associated sequences (800bp)
ARS I	- autonomously replicating sequence
ura4/LEU2	- selectable markers



It was envisaged that these vectors could be used as vector arms between which target DNA could be cloned and propagated in *S. pombe*, therefore creating *S. pombe* artificial chromosomes (SPARCs). The strategy for constructing these SPARCs is shown in figure 4.1. The *S. pombe* vector arms are created by first releasing the telomeres by digesting the plasmids with *SacI* and treating with phosphatase to prevent the telomeres religating. One telomere can subsequently be removed by digestion with *NotI* so that *NotI* fragments can be cloned between the vector arms. With an 8 bp recognition sequence, *NotI* is an infrequent cutting enzyme in the human genome therefore large human fragments can be isolated from a total digest of human DNA to clone into the SPARC system. This should allow the cloned fragments to be analysed for rearrangements by comparison with a *NotI* digest of the original total genomic DNA. The ligation products can be isolated away from vector arms on low melting point agarose pulsed field gels. Aliquots of this ligated DNA from the molten gel slice were used to transform *S. pombe* protoplasts. Selection for both *ura4* and *LEU2* can be used to isolate the transformants which are expected to carry recombinant SPARC products. Double selection should prevent the isolation of products with only one of the vector arms or two identical vector arms. DNA from each transformant recovered can be easily checked for the presence of a SPARC molecule by pulsed field gel electrophoresis.

### **4.3 CLONING HUMAN COSMID FRAGMENTS INTO SPARC VECTORS**

As an initial test of the SPARC system, human cosmid fragments were cloned to determine if these could be stably maintained in *S. pombe*. Six well-characterised cosmids derived from human chromosome 11 were selected. A DNA fragment of 30-50 kb in size could be derived from each of these cosmids (Hermanston *et al.*,

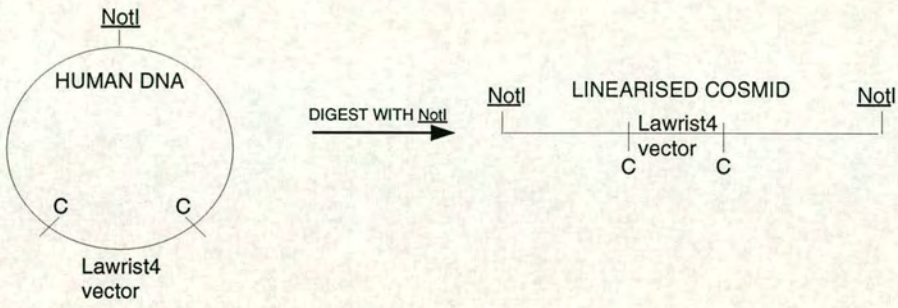


1992). Cosmid I (HIII48) of 45 kb and cosmid II (J44) of 50 kb were isolated from a lawrist4 vector library. Within the human insert of cosmids I and II there is a single *NotI* site which was utilised to linearise and clone the entire cosmid into the SPARC system (figure 4.2(a)). The other four cosmids III, IV, V and VI (U22, U143, U447 and U833 respectively) have human inserts cloned into the *BamHI* cloning site of the pWE15 cosmid vector. The design of the pWE15 polylinker allows the cosmid insert to be removed as a 30-50 kb fragment using *NotI* or *EcoRI* (figure 4.2(b)). All six cosmids were digested with *NotI* and the appropriate fragments were isolated from pulsed field gels and ligated between the SPARC vector arms. The large ligation products of interest were isolated away from the vector arms on pulsed field gels, transformed into *S. pombe* protoplasts and the resultant transformants analysed.

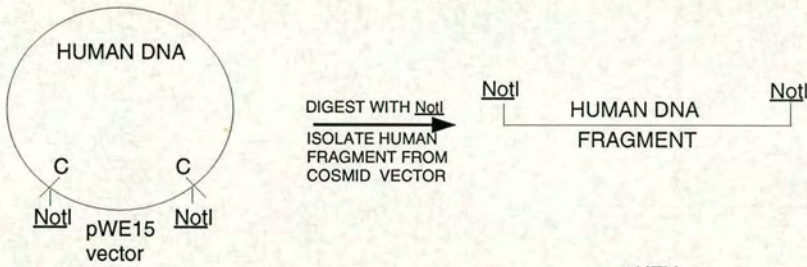


**Figure 4.2. Schematic diagram showing the human fragments produced from cosmids**

(A) *NotI* fragment produced from cosmids I and II



(B) *NotI* fragment produced from cosmids III and VI



KEY:

*NotI* Cloning site for *S. pombe* vector  
 C Cosmid cloning site



#### 4.4 CLONING OF COSMIDS I AND II

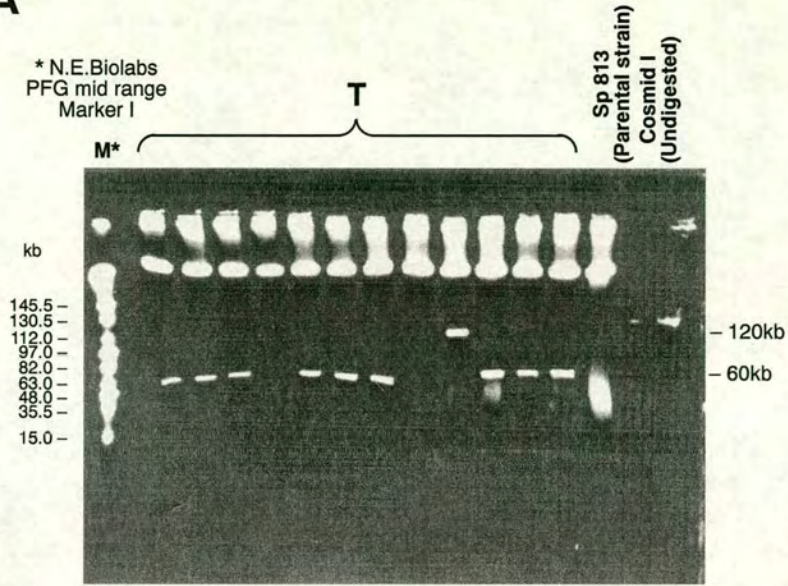
The first human cosmid *NotI* fragment cloned was approximately 45 kb. It was derived from the cosmid I and included the lawrist 4 cosmid vector sequences. This fragment was ligated to the SPARC vectors and transformed into SP813 protoplasts. Twenty-five transformants were recovered, DNA was prepared in agarose plugs from all. Analyses by pulsed field gel electrophoresis revealed that twenty-three of these transformants have an intense band of approximately 60 kb. Figure 4.3(a) shows the ethidium stained pulsed field gel with DNA from twelve of the transformants (the other eleven transformants gave similar results but are not shown). Approximately 60 kb is the expected size of recombinant SPARC products from the ligation of a pEN51 arm of 6.6 kb and a pEN53 arm of 7.7 kb to the 45 kb cosmid I fragment. However transformant 9 had an intense band of 120 kb shown in the tenth lane. Another transformant (lane 9) had no extrachromosomal band. This pulse field gel was Southern blotted and then probed with ( $\alpha$ -<sup>32</sup>P)dCTP labelled total human DNA (figure 4.3(b)). Of the 25 transformants, two did not contain the expected 60 kb SPARC band. Apart from the 60 kb band, each transformant also contains larger bands which migrate at approximately 120 kb and 180 kb. It is likely that these represent multimers of the 60 kb SPARC. These bands decrease with intensity as their size increases except in the tenth lane where the 120 kb dimer band is the most intense. These multimers may be produced by the SPARCs becoming associated through their telomere repeats. Alternatively these multimers could result from a recombination event between two SPARC molecules.



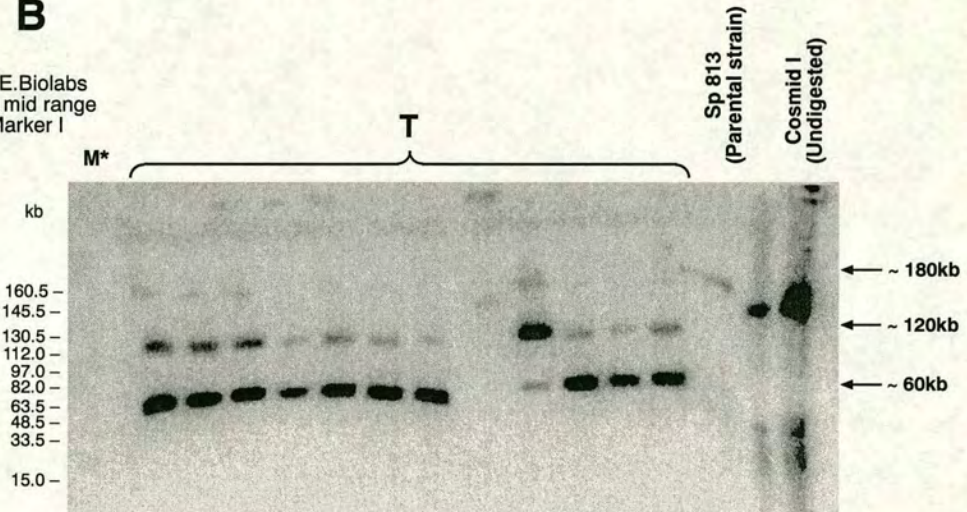
**Figure 4.3. DNA from *S. pombe* transformants with cosmid I SPARCs**

DNA from 12 of the 25 transformants recovered when cloning cosmid I into the SPARC system subjected to pulsed field gel electrophoresis. The sizes of the New England Biolabs pulsed field gel mid range marker I are shown for the lower bands. (A) Shows the photograph of the ethidium bromide stained gel. (B) Phosphor-image of DNA from this pulsed field gel transferred to Genescreen membrane and probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled total human DNA. The pulsed field gel electrophoresis conditions were 150 volts, 2 - 20 seconds, 20 hours, 0.5xTBE, 12°C.



**A****B**

\* N.E. Biolabs  
PFG mid range  
Marker I



Probe: Total Human DNA



Further analyses were carried out on four of these cosmid I SPARC transformants. DNA from the four transformants was digested with *NotI* which should release the 45 kb cosmid fragment and separated on a pulsed field gel alongside undigested DNA of the same transformant. The resulting gel was blotted and probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled human DNA (figure 4.4(a)). All four transformants contain a human insert of the expected size of 45 kb which comigrates with the *NotI* digested cosmid I DNA in the end lane. Therefore it appears that the *NotI* digested cosmid I has been cloned in the SPARC system without major rearrangement. To allow detection of both vector arms this Southern blot was stripped and reprobed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled plasmid DNA. The resulting phosphor-image is shown in figure 4.4(b). The vector arms should migrate at 6.6 kb and 7.7 kb at the bottom of the gel. In three of the transformants the two vector arms are clearly visible, however in the third transformant there appears to be no pEN51 arm at 6.6kb. Other analyses have shown that in this transformant the pEN51 vector arm is smaller than the expected size (data not shown) and therefore this fragment was run off the bottom of the gel. The reduction in vector arm size may be due to a rearrangement but it is clear the human insert was unaffected.

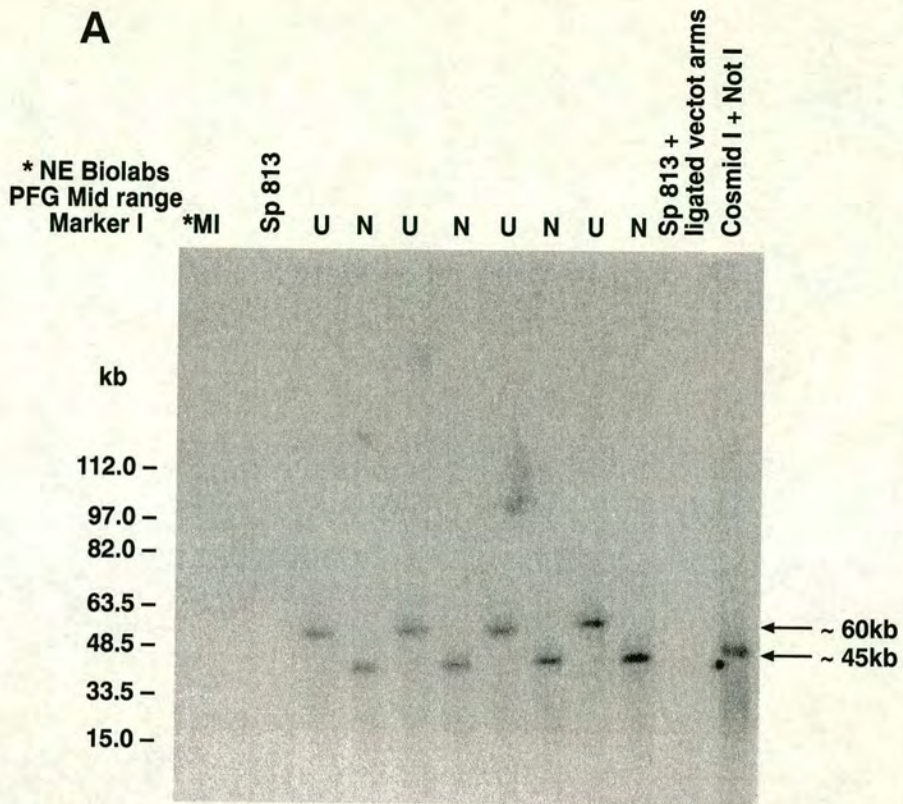


**Figure 4.4. *NotI* restriction digest analysis on DNA from 4 cosmid I SPARC transformants.**

Phosphor-images of DNA from a pulsed field gel transferred to Genescreen membrane. DNA from four of the transformants are shown before and after digestion with *NotI*. The sizes are shown for the lower bands of the New England Biolabs pulsed field gel mid range marker I. (A) The membrane was probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled total human DNA. The undigested tracks (U) show the extrachromosomal band of around 60 kb while the *NotI* digested DNA (N) shows the human cosmid fragment corresponding to the 45 kb cosmid fragment cloned. (B) The membrane shown in (A) was stripped and reprobbed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled plasmid. The pulsed field gel electrophoresis conditions were 150 volts, 2 - 20 seconds, 20 hours, 0.5xTBE, 12°C.

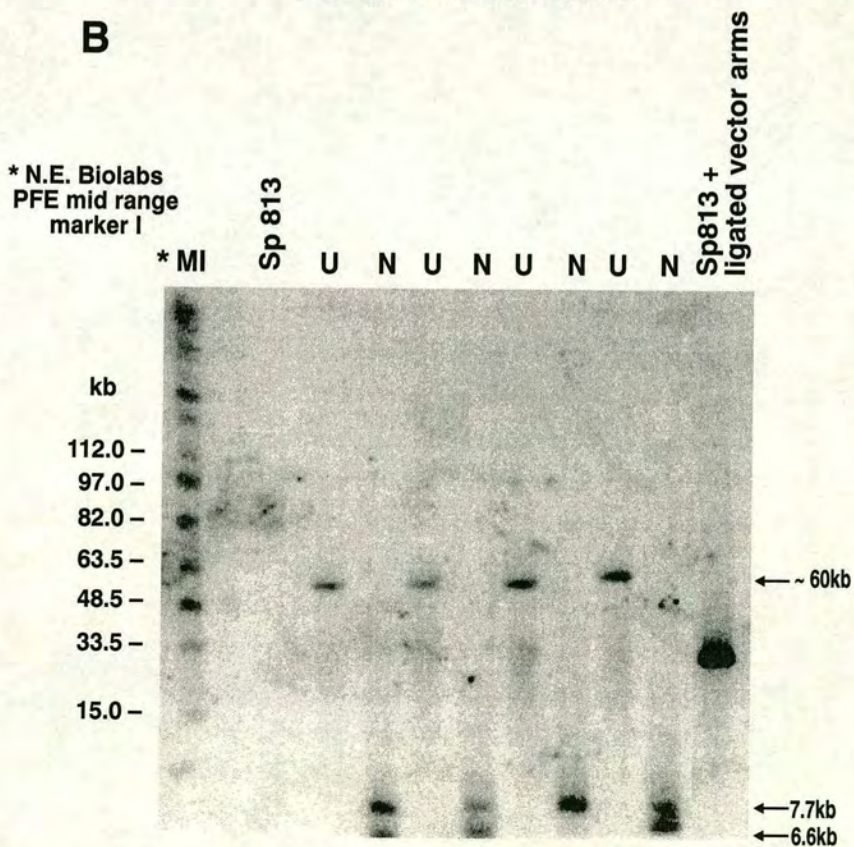


**A**



Probe: Total Human DNA

**B**



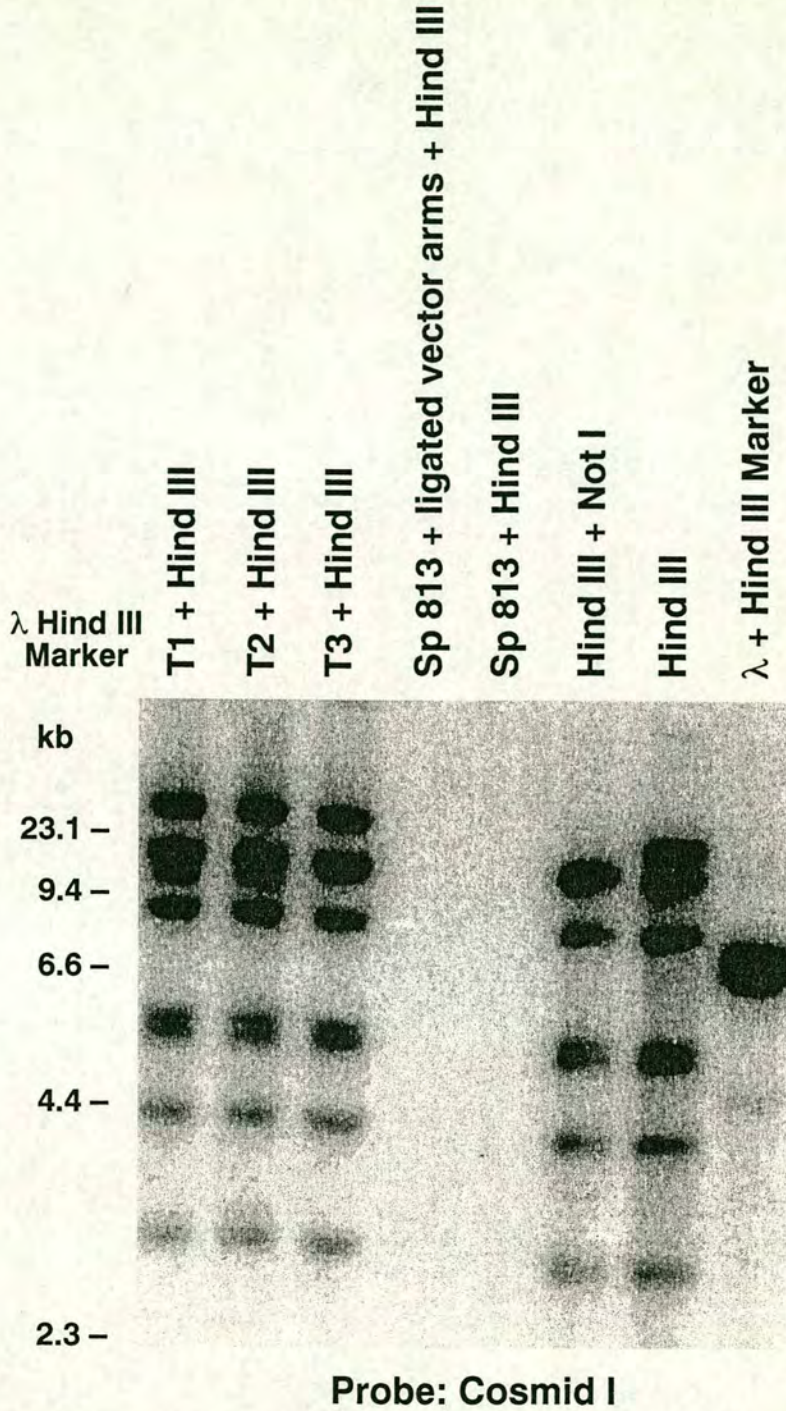
Probe: pBSKS plasmid



To determine whether there had been any smaller deletions or rearrangements within the human insert these cosmid I SPARCs were further analysed by digestion with other restriction enzymes cutting within the insert. Total DNA from three transformants was digested with *HindIII* and separated by conventional gel electrophoresis alongside various controls. Figure 4.5 is the phosphor-image of a Southern blot, using ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled cosmid I as the probe. The cosmid itself was used as the probe since it should detect all bands derived from the cosmid whereas total human DNA may not. The restriction pattern of digested cosmid I DNA is shown in lanes 6 and 7 (the DNA in lane 6 was digested with *HindIII* and *NotI*, lane 7 with *HindIII*) and can be compared to the digest pattern of the transformant DNA samples, in lanes 1,2 and 3. The restriction patterns of the cosmid I fragment and the human cosmid SPARC are identical apart from the larger bands in the transformants' DNA digest. These fragments contain the *NotI* junction between the cloned cosmid I fragment and SPARC vector. This was confirmed when the blot was reprobbed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled plasmid and these bands were detected (data not shown). Therefore this data in combination with the analyses shown in figures 4.3 and 4.4 indicate that these human cosmid SPARCs had no detectable deletions or rearrangements in any of these randomly selected transformants.



Figure 4.5. Further restriction analysis on DNA from 3 cosmid I SPARC transformants.

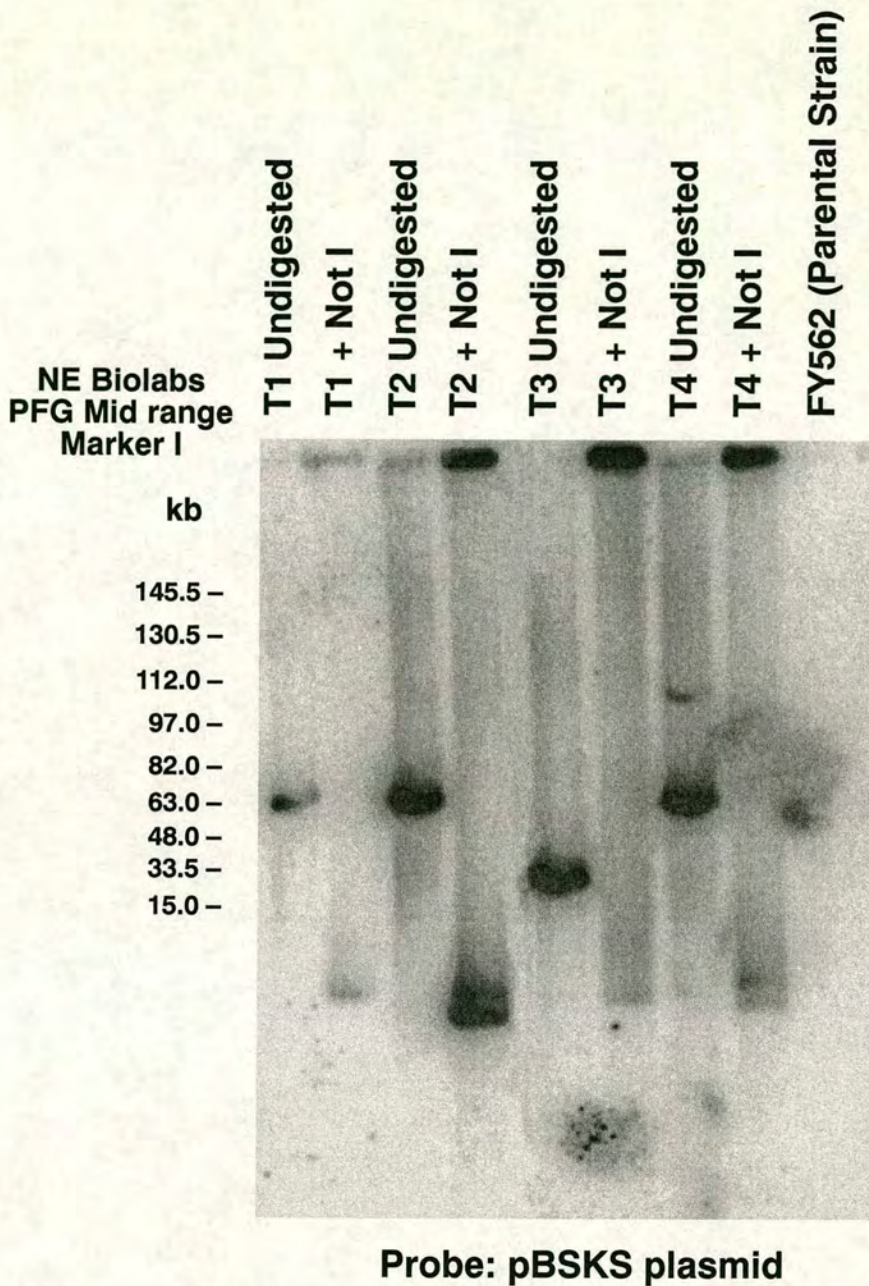


Phosphor-image of DNA digested with *Hind*III subject to conventional gel electrophoresis, transferred to Genescreen and probed with ( $\alpha$ - $^{32}$ P)dCTP labelled cosmid I fragment. The digest pattern of three transformants (T1, T2 and T3) have been compared to the digest pattern of the cosmid digest shown in lanes 6 and 7 where the cosmid has been digested with *Hind*III plus *Not*I and *Hind*III alone respectively. The sizes of the larger bands of the lambda *Hind*III digest marker are shown. Lanes 4 and 5 contain the DNA from SP813 with vector arms and SP813 alone were used as negative controls.



Only four transformants were recovered when cosmid II was ligated to the SPARC arms. This was probably due to a poor recovery of ligation products from the isolation gel. DNA from the four recovered transformants was digested with *NotI*, this along with undigested DNA was subject to PFGE. A phosphor-image of the resulting Southern blot probed with ( $\alpha$ - $^{32}$ P)dCTP labelled plasmid DNA is shown in figure 4.6. Three of the four transformants produced a band of the expected size. There are 2 extrachromosomal bands in transformant 4 one of the expected size 65 kb in addition to a larger band of approximately 100 kb. Further analyses with *NotI* plus *EcoRI* restriction digests are shown in figure 4.7. *NotI* and *EcoRI* digests of the four transformants were compared with the same digest of cosmid II. The digest profile of the transformant T1, T2 and T4 matches the profile of the cosmid II. It is clear that transformant T3 contains only a few of the *EcoRI* fragments present in human cosmid II. Two of the bands of approximately 1 kb are seen in the cosmid II restriction pattern however a band of approximately 4.4 kb can be seen which is different from the cosmid II digest. Out of the four transformants, all *EcoRI* fragments are apparently maintained in three of these SPARCs with no rearrangements detected. In one transformant, the human insert has clearly been rearranged wither due to events before, during or after transformation. Regardless of this, both cosmid I and II can clearly be propagated intact in *S. pombe*.





**Figure 4.6.** *NotI* restriction digest analysis on DNA from 4 cosmid II SPARC transformants.

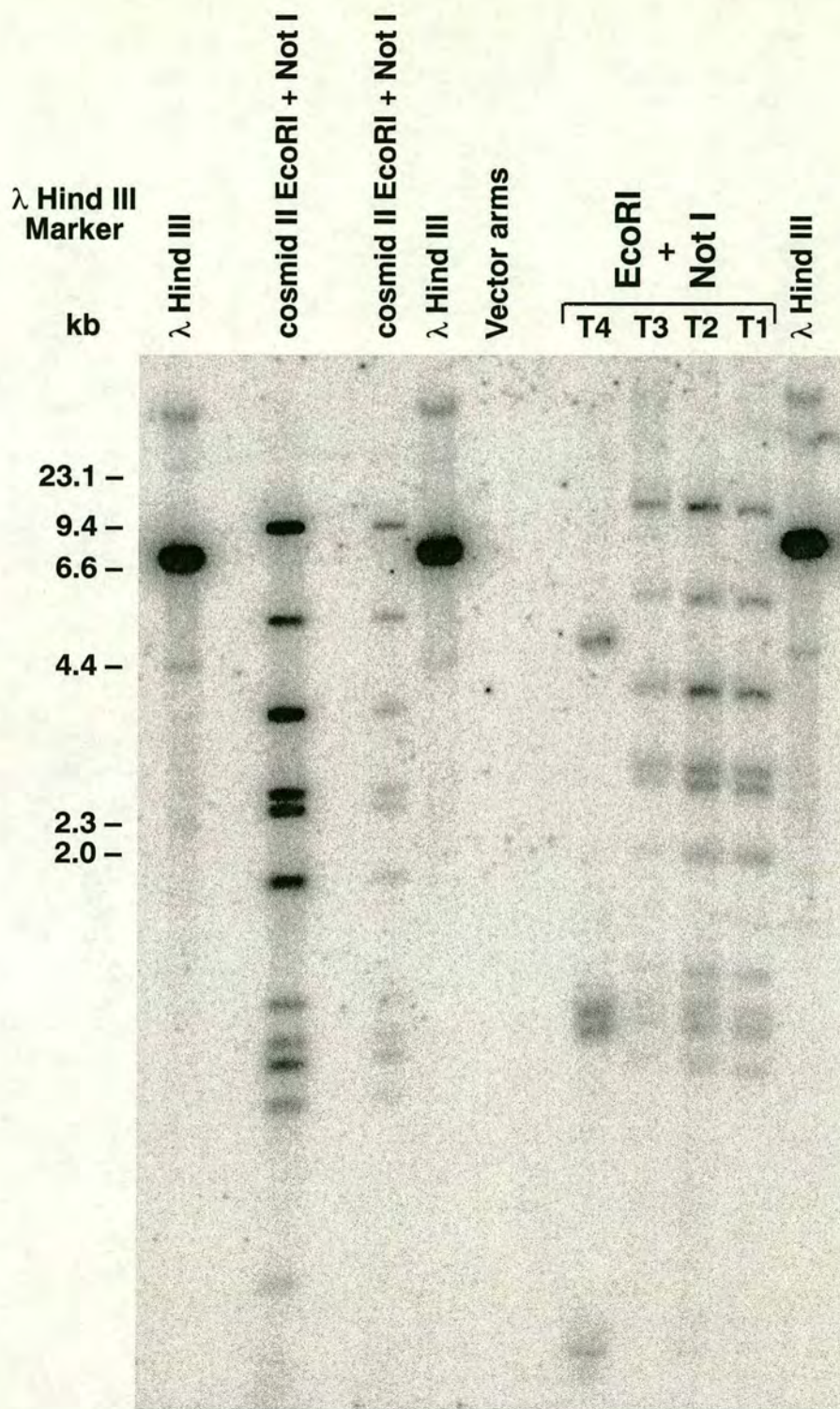
Phosphor-image of DNA from transformants with SPARCs with a cosmid II insert subject to pulsed field gel electrophoresis, transferred to Genescreen then probed with ( $\alpha$ - $^{32}$ P)dCTP labelled plasmid. The sizes are shown for the lower bands of the New England Biolabs pulsed field gel mid range marker I. DNA from four transformants (T1 to T4) with undigested DNA and DNA digested with *NotI* are shown. The pulsed field gel electrophoresis conditions were 150 volts, 2 - 20 seconds, 20 hours, 0.5xTBE, 12°C.



**Figure 4.7. Further restriction analysis on DNA from 4 cosmid II SPARC transformants.**

Phosphor-image of DNA from transformants with SPARCs with a cosmid II insert digested with *EcoRI* plus *NotI*, subjected to conventional gel electrophoresis, transferred to Genescreen membrane then probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled cosmid II fragment. The larger bands for the lambda *HindIII* digest markers have been shown. The digests patterns of the four transformants (T1 - T4) have been compared to the digest profile of the cosmid II.





**Probe: Cosmid II**



#### 4.5 RECOMBINATION DEFICIENT MUTATION CROSSED INTO *S. pombe*

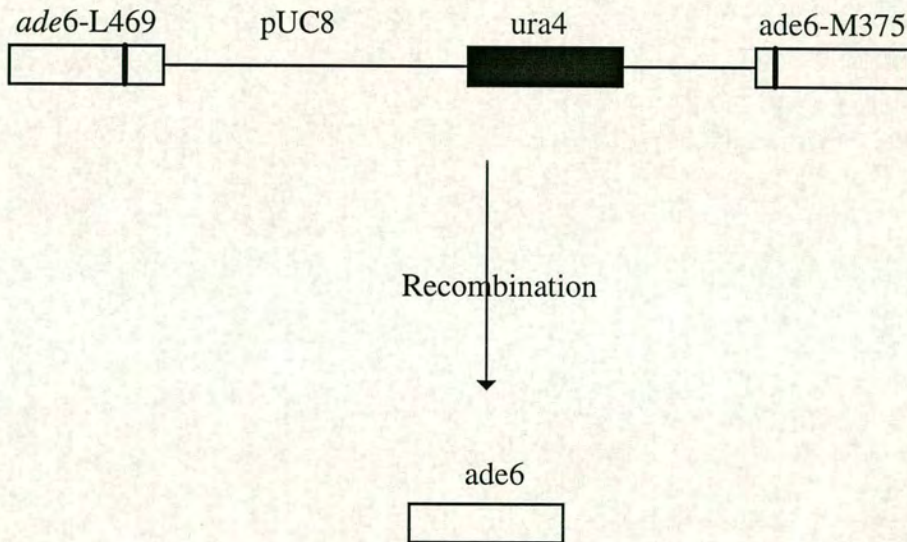
In the analysis of the cosmid I SPARCs there were multimers forming a ladder increasing in size by 60 kb but decreasing in intensity. The insert itself was intact suggesting that the vector arms played a role in this phenomenon, either by telomere associations or due to a recombination event between vector sequences.

Recombination is known to cause rearrangements within *S. cerevisiae* YACs when transformed into *S. cerevisiae* spheroplasts (Larinov *et al*, 1994; Ling *et al*, 1993). *S. cerevisiae* recombination deficient strains mutated in RAD52 have been shown to reduce the occurrence of chimeric YACs from around 50% to 10% (Larinov *et al*, 1994), suggesting that recombination deficient mutants help to prevent rearrangement of transformed DNA.

In *S. pombe* a recombination deficient mutant, with a 10 fold reduction in mitotic recombination, was isolated in a screen for meiotic recombination mutants (Gysler-Junker *et al.*, 1991). The recombination frequency of the mutant strains compared to the parental strain was determined using a marker system constructed to detect recombination events. The tester construct had the *ura4* gene and plasmid sequences integrated between two copies of the *ade6* gene with point mutations in different regions of the gene. Normally this strain would be *ura*<sup>+</sup> and *ade*<sup>-</sup>, producing a red pigment on plates with reduced adenine. The occurrence of a recombination event between the two mutant *ade6* genes can result in the deletion of the intervening *ura4* gene, plasmid sequences and the 3' end of one *ade6* gene and the 5' end of the other *ade6* gene creating a wildtype copy of the *ade6* (figure 4.8). This results in cells which are *ade*<sup>+</sup>*ura*<sup>-</sup> are white on limiting adenine plates and unable to grow without uracil.



**Figure 4.8 Construct used to determine recombination deficient *S. pombe* strains**  
(Gysler-Junker *et al.*, 1991)



This construct shows the *ade6* point mutation duplication with *ura4* present in the intervening sequences which gives a *ura4*<sup>+</sup>, *ade6*<sup>-</sup> strain. Recombination between the *ade6* mutant genes can occur to produce an *ade6* wildtype gene with the loss of *ura4* which results in a *ura4*<sup>-</sup>, *ade6*<sup>+</sup> strain.



The *rec55-36* mutant has a 10 fold reduction in the frequency of recombination between tandem copies of the *ade6* gene as measured by the appearance of *ade*<sup>+</sup> or white colonies. The frequency of mitotic recombination was reduced from  $3.31 \times 10^{-4}$  to  $3.45 \times 10^{-5}$  in this mutant (Gysler-Junker *et al.*, 1991).

In order to cross the *rec55-36* mutation into a background appropriate for SPARC propagation, SP813 and BG47 were crossed and the *rec55-36* mutation followed by different recombination frequency as monitored by colony colour. The *rec55-36* mutant produced a predominantly red patch with only a few white colonies as a result of 10 fold reduction in recombination while the wildtype strain produced pink patches. Once confirmed as having the correct genotype the white segregants with the *rec55-36* background were screened to isolate a strain lacking the intervening *ura4* gene. To check whether the loss of uracil was complete the colonies were plated on media containing 5-fluoro-orotic acid (FOA) on which *ura4*<sup>+</sup> cells cannot grow. The resulting strain FY562 (*h*<sup>-</sup> *ura4*-DSE, *leu1-32*, *rec55-36*) carrying this mitotic recombination mutation was used in all the following experiments.

#### **4.6 CLONING OF THE HUMAN FRAGMENTS FROM pWE15 COSMIDS**

The human inserts of the pWE15 cosmids (cosmids III, IV, V and VI) were isolated by first digesting with *NotI* then subjecting the DNA to PFGE to separate the 30-42 kb human inserts from the cosmid vector. These human fragments were ligated to the SPARC vector arms and run on a PFGE to isolate the ligation products away from the vector arms. For each human cosmid insert, approximately 50 colonies were recovered per 10 µl of the ligation products (50-100 ng DNA). DNA was prepared in agarose plugs from five transformants derived from each of the different



cosmid SPARCs. This DNA was subject to pulse field gel electrophoresis and the resulting Southern blot was hybridised with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled total human DNA (figure 4.9). Most of the transformants analysed produced a band of the expected size for each of the different cosmid SPARCs. Again higher molecular weight bands representing presumably dimer bands were detected in transformants for each cosmid SPARC. Therefore using this *rec55-36* recombination deficient strain did not decrease the frequency of multimerisation of these cosmid SPARCs.

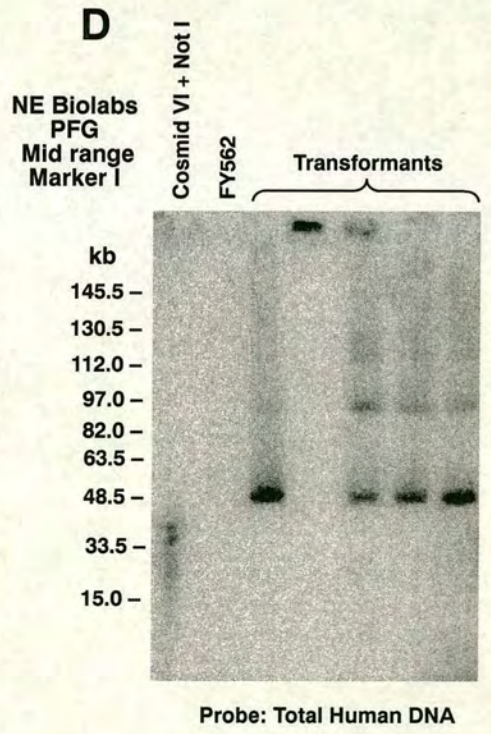
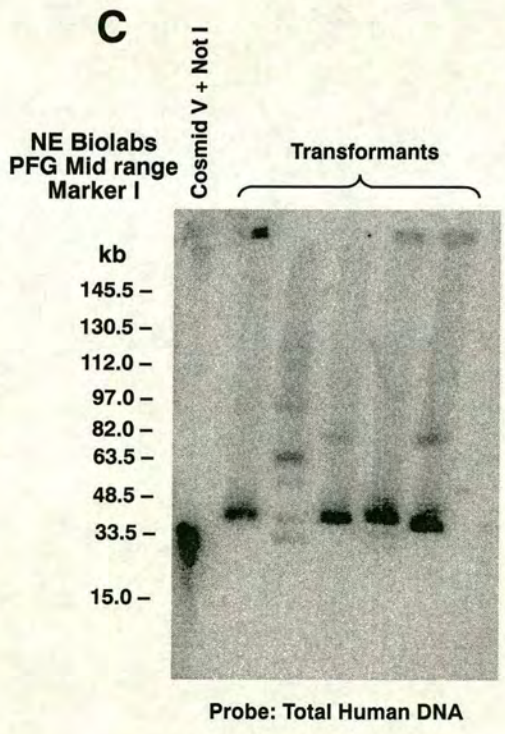
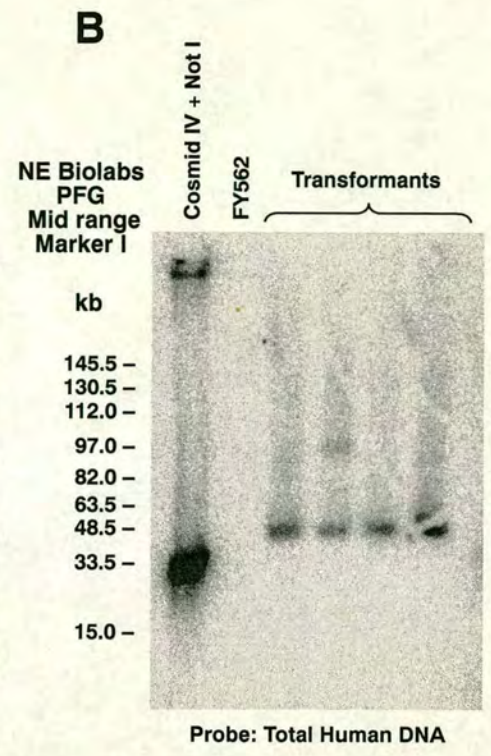
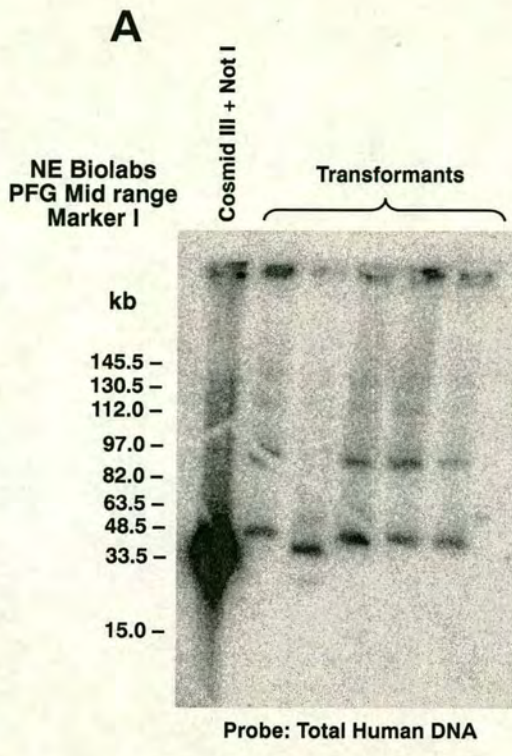
Cosmid III carries a *NotI* insert of approximately 30 kb and produced SPARCs of 45 kb in 4 of the five transformants analysed, shown in figure 4.9(a). However one of these (second transformant) had clearly rearranged to produce a smaller band of approximately 35 kb. Figure 4.9(b) shows 4 of the SPARCs derived from the 30 kb human insert of cosmid IV had a band of the expected size, 45 kb. A fifth transformant analysed did not have an extrachromosomal band. Figure 4.9(c) shows for cosmid V, 4 of the 5 resulting SPARC transformants had the expected extrachromosomal band of 45 kb. In the second transformant multiple bands are visible, one band of the correct size (45 kb) is present, with 4 others of approximately 30, 60, 65 and 90 kb. The cosmid VI results are shown in figure 4.9(d), where 4 of the five SPARC transformants had the expected extrachromosomal band of approximately 50 kb as a result of cloning 35 kb human insert from the cosmid. In the second lane there was no extrachromosomal cosmid SPARC band but the human DNA appears to have remained in the well.



**Figure 4.9. DNA from *S. pombe* transformants with SPARCs containing small human DNA inserts.**

Phosphor-images probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled total human DNA. The smaller bands from the New England Biolabs pulsed field gel marker I have been indicated. (A) SPARC bands derived from cosmid III. Transformants from cloning of human (B) cosmid IV, (C) Cosmid V and (D) Cosmid VI. In each case the cosmid fragment has been used as a positive control and the parental strain, FY562, used as a negative control (shown in figures B and D). The pulsed field gel electrophoresis conditions were 150 volts, 2 - 20 seconds, 20 hours, 0.5xTBE, 12°C.







Approximately 80% of the SPARC transformants with cosmids III, IV, V and VI with human inserts of 30-40 kb were found to be present as acentric linear episomes. Two transformants did not contain any extrachromosomal band, although the transformant derived from the human insert from cosmid VI had some human DNA retained within the well. This may suggest that the SPARC is perhaps circular in nature as such molecules tend not to migrate normally into pulsed field gels, alternatively the SPARC may have integrated into a host chromosome.

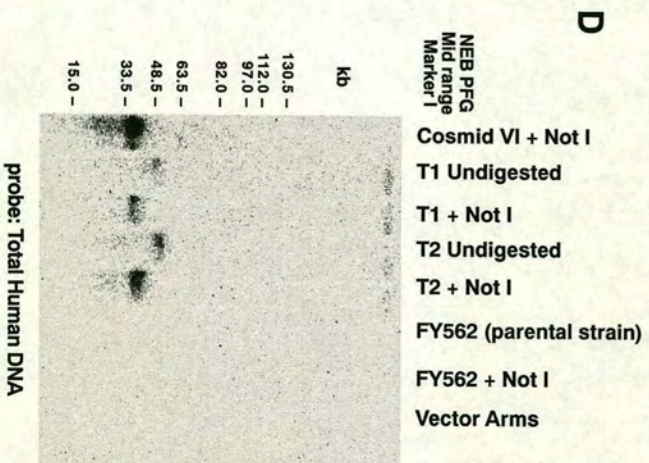
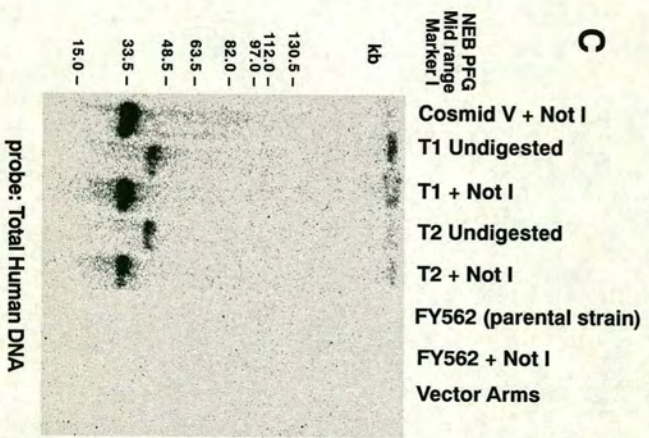
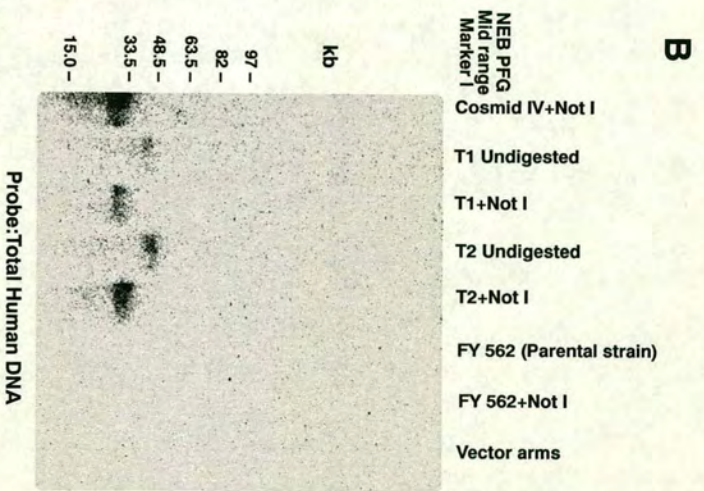
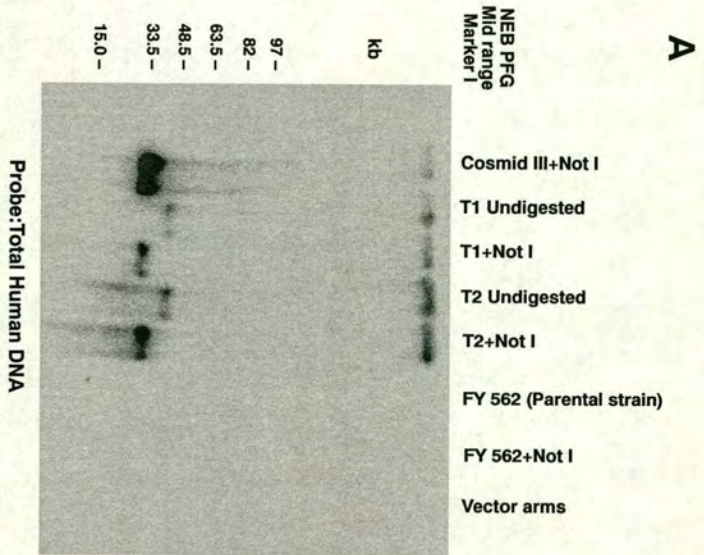
Two transformants were analysed for each of the human cosmids III to VI. A *NotI* digest of DNA from the transformants should confirm if the correct sized fragment was cloned and propagated without alteration. Figure 4.10 shows the phosphor-image results from the cosmid SPARCs run on a pulsed field gel where the undigested transformant DNA (T1 and T2) was run beside *NotI* digests of both transformants (T1/T2 + *NotI*) and compared with a *NotI* digest of each starting cosmid. The probe used for these Southern blots was ( $\alpha$ -<sup>32</sup>P)dCTP labelled total human DNA. Figure 4.10(a) shows that the human cosmid fragment cloned comigrates with the 30 kb human *NotI* fragment of cosmid III in both transformants analysed. In figure 4.10(b) both transformants contain a cloned human fragment of approximately 30 kb corresponding to the size of the cosmid IV *NotI* fragment. Cosmid V carried a human *NotI* fragment of 30 kb, clearly both of the SPARC transformants analysed produce a *NotI* band of the same size (figure 4.10(c)). Figure 4.10(d) shows that when the cosmid VI SPARC transformant DNA is digested with *NotI* then the resulting human insert is 40 kb which corresponds to the size of the cloned cosmid VI *NotI* fragment.



**Figure 4.10. *NotI* restriction digest analyses on DNA from transformants with SPARCs.**

Phosphor-images analysis of 2 SPARCs derived from cosmids III to VI. Pulsed field gels transferred to Genescreen membrane and probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled total human DNA. The smaller bands from the New England Biolabs pulsed field gel marker I have been indicated. (A) Cosmid III, T1 is transformant 1 and T2 is transformant 2. (B) Cosmid IV, T1 is transformant 1 and T2 is transformant 2. (C) Cosmid V, T1 is transformant 1 and T2 is transformant 2. (D) Cosmid VI, T1 is transformant 1 and T2 is transformant 2. The human *NotI* fragment from the cosmid were used as the positive control with the parental strain DNA used as the negative control. The pulsed field gel electrophoresis conditions were 150 volts, 2 - 20 seconds, 20 hours, 0.5xTBE, 12°C.







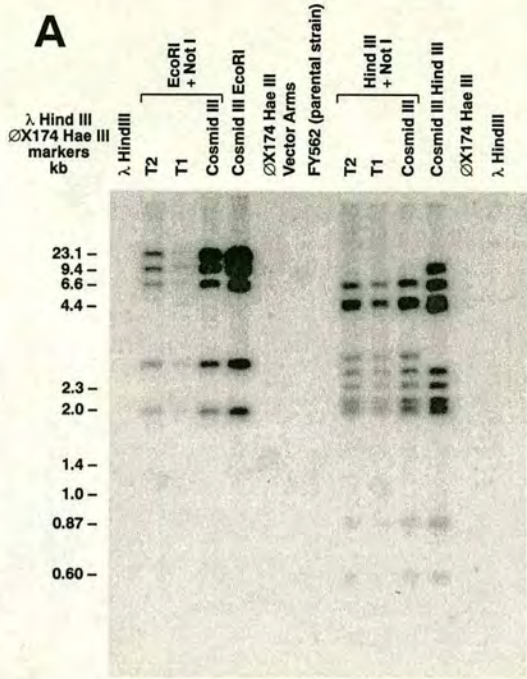
Further restriction digest analyses with *NotI* and *EcoRI* or *HinDIII* on the DNA isolated from the transformants from each cosmid was compared to the equivalent digest of that cosmid. This DNA was subject to gel electrophoresis, transferred to nylon membrane, and the resulting filters probed with the corresponding ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled human cosmid *NotI* fragment. Figure 4.11(a) shows the digest profiles of the cosmid insert of two transformants with cosmid III SPARCs as compared to the original human cosmid III. The DNA digest profiles of the transformants is the same in both digests as that of the cosmid itself. In figure 4.11(b) the same comparisons were made for the cosmid IV SPARCs with the DNA digest profiles the same for cosmid IV and the two SPARC transformants. In figure 4.11(c) the DNA digest profiles of the human cosmid inserts of the two SPARC transformants are the same as cosmid V in both digests. In figure 4.11(d) the same comparison was made between the human cosmid fragment VI and the human cosmid inserts of both of the transformants. Again the DNA digest profiles were the same. These results show that these four human *NotI* fragments can be propagated intact in *S. pombe* as linear episomes.



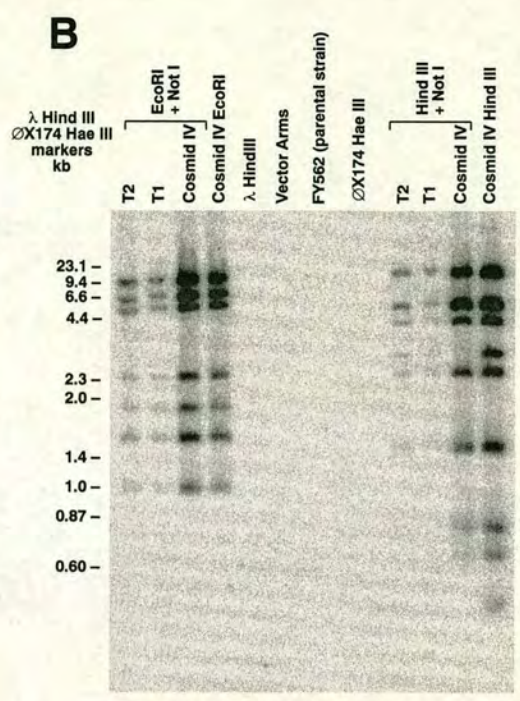
**Figure 4.11. Further restriction analyses on DNA from transformants with SPARCs.**

Phosphor-image of digested DNA from 2 transformants (T1 and T2) for each of the human cosmid fragments cloned into the SPARC system. After conventional gel electrophoresis and transfer to Genescreen membrane the filter was probed with the appropriate ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled cosmid fragment, (A) cosmid III fragment, (B) cosmid IV, (C) cosmid V and (D) cosmid VI. Each example shows the transformant DNA digest profiles with *EcoRI* plus *NotI* or *HindIII* plus *NotI* compared to the equivalent restriction digest for the cosmid.

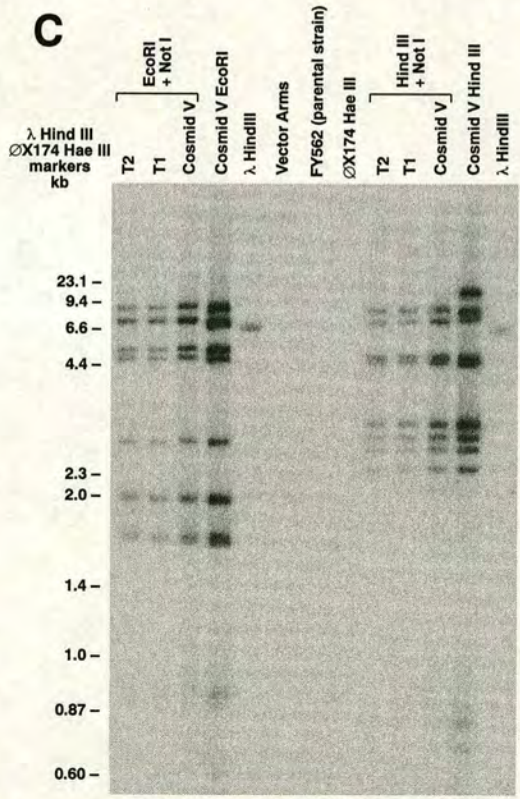




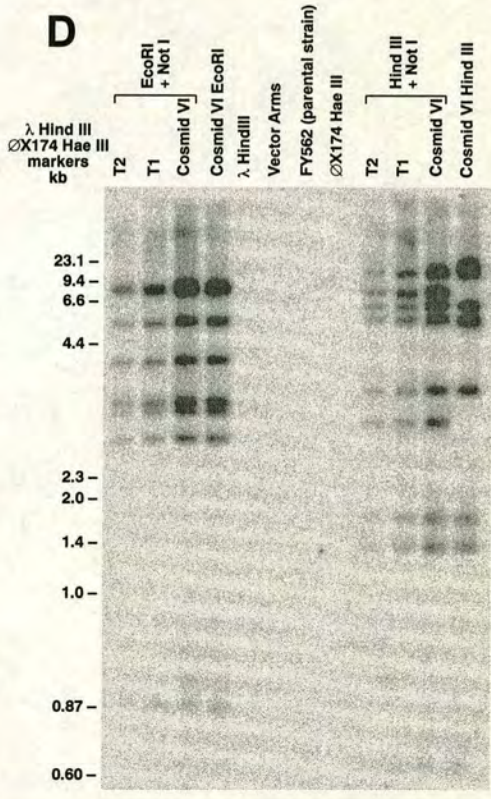
Probe: Human Not I Fragment of Cosmid III



Probe: Human Not I Fragment of Cosmid IV



Probe: Human Not I Fragment of Cosmid V



Probe: Human Not I Fragment of Cosmid VI



## 4.7 DISCUSSION

The cloning of all six human cosmid fragments into the SPARC vector system has shown that human fragments of up to 45 kb remained intact when cloned in the *S. pombe* SPARC system. There were no apparent problems with using an acentric SPARC vector, although SPARCs must be maintained under selection for presence of both vector arms. Judging by the intensity of the ethidium bromide staining of the SPARC bands on pulsed field gels, the copy number of the cosmid SPARCs did appear to be greater than one copy per cell (the copy number of human SPARCs is discussed in chapter 5). Many of the cosmid SPARCs appeared to produce multimers upon propagation in *S. pombe*. The cause of multimer formation with each of the cosmids SPARC transformants analysed was not formally identified. Multimerisation could result from telomeric associations or recombination events. However use of a recombination deficient strain did not appear to affect the occurrence of these multimers. In addition, both vector arms were clearly visible in figure 4.4(b) and no dimer or rearranged vector arms of greater than the expected 8 kb were detected. This may suggest that the multimers resulted from non covalent end to end fusion events. The six human cosmid fragments were cloned successfully into the SPARC vector system.

The next stage is to attempt to clone large fragments of human DNA with the SPARC system to determine the usefulness of the SPARC vectors for producing a library. A human SPARC library should complement the existing human DNA libraries constructed in other vectors, including *S. cerevisiae* YACs. DNA from a normal human male would provide a complete complement of human DNA for creating a fully representative library.



## **5 GENERATION OF SPARCS WITH LARGE HUMAN INSERTS**



## 5.1 INTRODUCTION

In the previous chapter, it was demonstrated that small human DNA fragments (derived from well-characterised cosmids) could be cloned successfully into the *S. pombe* vector system. This produced acentric *S. pombe* artificial chromosomes (SPARCs) which were maintained intact as linear episomes. The absence of a centromere in the SPARCs had no bearing on the integrity of the human insert. The experiments in this chapter were designed to test whether randomly selected human DNA fragments can be cloned and propagated as large intact fragments in SPARCs.

Both vector arms contain a convenient *NotI* site, therefore these experiments centred on testing if randomly sized *NotI* fragments of human DNA could be cloned between the SPARC vector arms. *NotI*, recognises an 8 bp, GC rich sequence and is therefore an infrequent cutter of mammalian genomic DNA. It gives a wide range of fragments of 50 kb up to megabase sizes for cloning. To construct a library representative of the entire human genome would require partial digestion using a more frequently cutting enzyme to produce an array of overlapping fragments for cloning. However it is appropriate to test the SPARC vector system with fragments of a known original size since this should assist in detecting rearrangements of inserts.

The human DNA utilised was derived from a male human lymphoblastoid cell line, FATO, immortalised by infection with EBV. The initial analysis presented below demonstrates that the SPARC vectors allow large fragments of DNA from the human genome to be cloned and propagated in *S. pombe*. Since the SPARCs lack centromeres which would usually serve to maintain a single copy of a chromosome per cell, the copy number of SPARCs of various sizes was estimated. *S. pombe*



protoplasts were retransformed with two of the SPARCs to determine the extent of rearrangement of the human insert during such manipulations. This experiment also allowed the efficiency of transformation with large SPARC molecules to be estimated.

## **5.2 ISOLATION OF HUMAN DNA AND LIGATION INTO SPARC VECTORS**

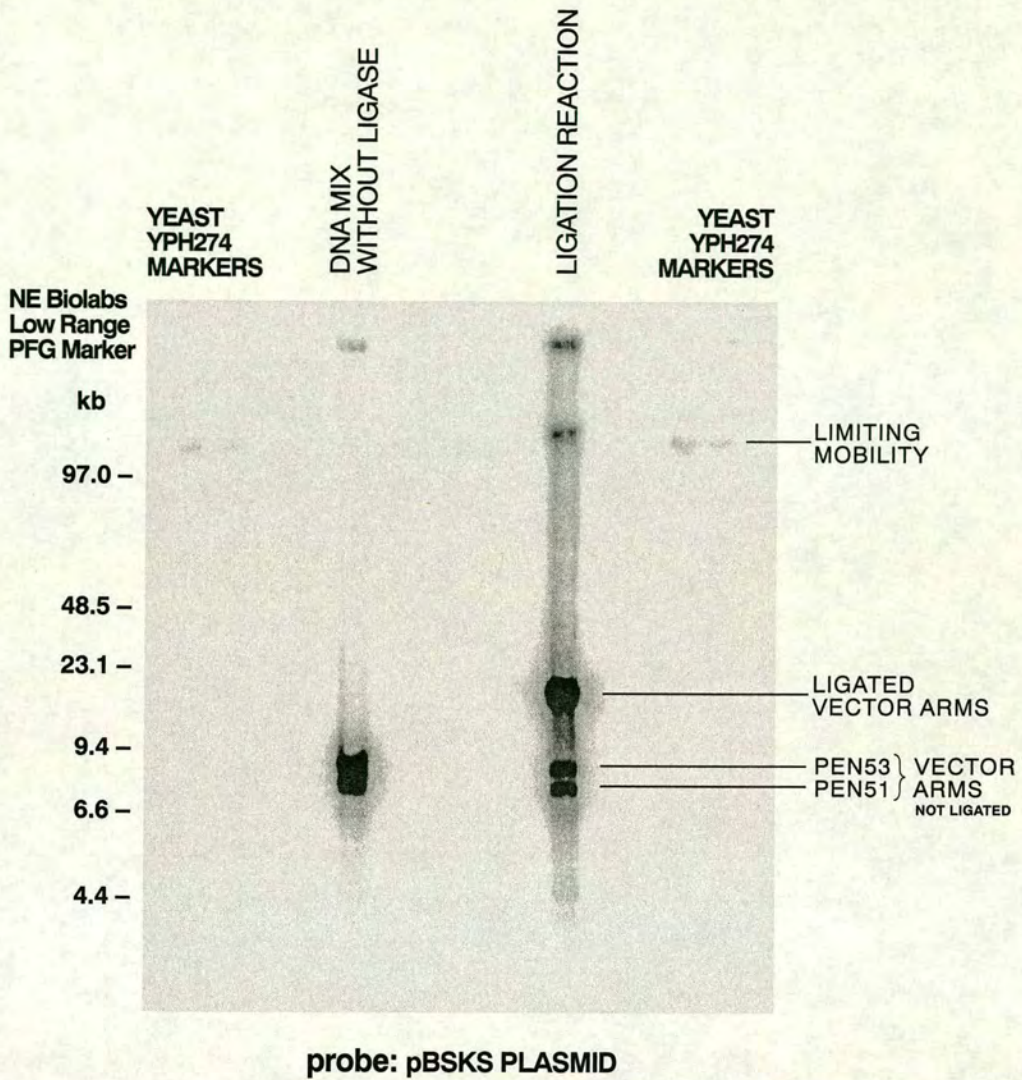
Concentrated human DNA prepared in agarose plugs from the FATO cell line was provided by W. Bickmore. These plugs were equilibrated in a high salt buffer (TE with 100 mM NaCl) after digestion with *NotI* in order to protect the DNA from degradation upon melting (refer to chapter 3). Whereas polyamines were the preferred method for protecting the YAC DNA in chapter three, this high molecular weight human DNA was protected with salt. The reason for changing the protective solutions was mainly because many of the buffers used for the ligation and digestion steps contain a high concentration of salt and combining these with polyamines could have resulted in the precipitation of large DNAmolecules (as shown for YAC2 in figure 3 (b)). After equilibration in high salt buffer the agarose plugs were melted at 65°C, cooled to around 37°C and the vector arms added. For the ligation step, the agarose containing human *NotI* DNA fragments and the vector arms was reformed as 100 µl plugs and allowed to equilibrate in ligation buffer prior to the addition of ligase. Although the ligation step was first attempted in the molten agarose, the high molecular weight DNA was more prone to degradation under these conditions (data not shown). Therefore the ligation reaction was carried out in set agarose in which the high molecular DNA was better protected. In order to check the ligation reaction had progressed, a portion of an agarose plug was subject to PFGE and compared to a



portion of this DNA mixture to which ligase was not added. Pulsed field gel conditions were used which separated the smaller DNA fragments, i.e. the vector arms, while the high molecular weight DNA was retained in the region of limiting mobility. The resulting Southern blot was probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled plasmid DNA which hybridised to the vector arms. Figure 5.1 shows the results of a phosphor-image where the ligation was compared to mock ligation control. In both tracks the vector arms were visible while additional bands were seen in the ligation, one which corresponds to the product self ligation of vector arms resulting in a 15 kb band and another band seen in the limiting mobility.

When a successful ligation reaction was observed the rest of the reaction was subjected to PFGE in a large well (around 4-5 cm in width) on a low melting point agarose gel. The pulsed field gel conditions were set in order to separate DNA molecules less than 200 kb. The limiting mobility was identified by staining of the outer edges of the lane and cutting out the limiting mobility in a gel slice from the unstained portion. This limiting mobility gel slice should contain all molecules greater than 200 kb including those where the SPARC arms had ligated to human *NotI* fragments. This gel slice was then equilibrated in a buffer containing TE plus 100 mM NaCl and melted at 65°C until the agarose was molten. Aliquots of this melted gel slice containing the DNA transformed *S. pombe* protoplasts and plated on selective minimal plates. Transformants appeared after 7-10 days at 32°C.





**Figure 5.1. Ligation of SPARC vector arms to large human DNA fragments.**

Phosphor-image of DNA subject to PFGE and transferred to genescreen membrane then probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled plasmid. The PFG conditions used were 100 volts, 2 to 5 seconds, 20 hours, 0.5xTBE, 12°C. Such conditions allowed the small fragments to migrate into the gel while the high molecular weight DNA remained at the region of limiting mobility. DNA from the yeast strain YPH274 was used to highlight the limiting mobility.



### 5.3 ANALYSIS OF SPARC TRANSFORMANTS

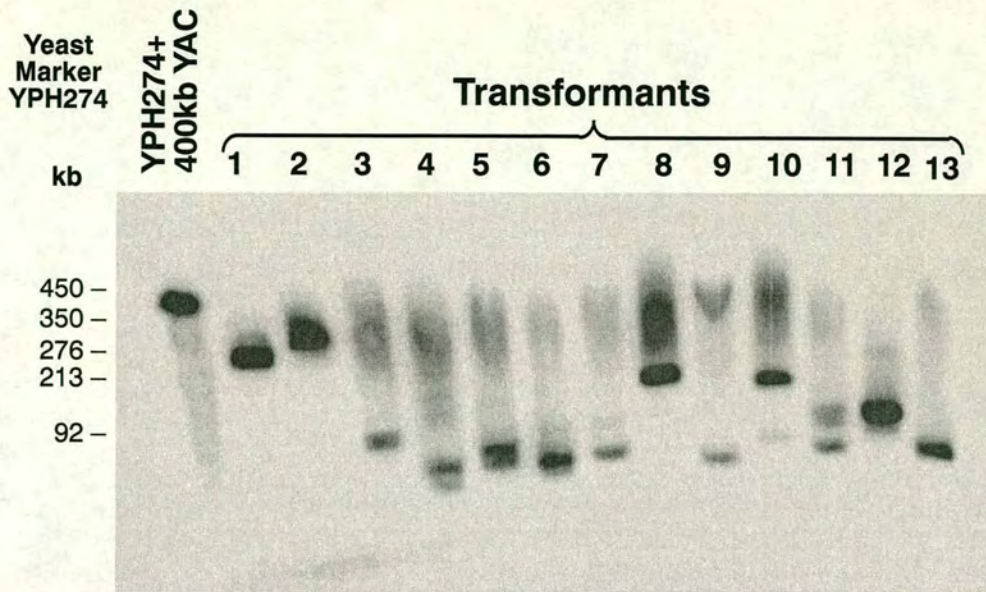
The products of the ligation reaction shown in figure 5.1 transformed *S. pombe*. Twenty five *ura<sup>+</sup>leu<sup>+</sup>* transformants were recovered and DNA was prepared from all 25 of these strains. In order to check these transformants for the presence of extrachromosomal bands, these undigested DNAs were subject to PFGE. These pulsed field gels were transferred to genescreen and probed with ( $\alpha$ -<sup>32</sup>P)dCTP labelled total human DNA. In figure 5.2 the resulting phosphor-image shows that 22 of the 25 transformants have SPARCs containing human inserts. Transformants 18, 19 and 23 did not contain any extrachromosomal bands with human inserts. Twelve of the transformants (3, 4, 5, 6, 7, 9, 11, 13, 17, 22, 24 and 25) had SPARC bands less than 100 kb in size. The other ten transformants (SPARCs 1, 2, 8, 10, 12, 14, 15, 16, 20 and 21) had human inserts of greater than 100 kb. Therefore, this set of SPARCs have human inserts ranging from 50 kb to 400 kb. The average insert size for all 25 transformants is around 150 kb which is lower than expected considering that the ligation products were size selected to be over 200 kb. This may be because some of the smaller DNA molecules can still be trapped in the region of limiting mobility. The size of the fragments used to transform *S. pombe* were checked by PFGE which used conditions to separate out DNA fragments of up to 1 Mb. The phosphor-image of the resulting Southern blot probed with ( $\alpha$ -<sup>32</sup>P)dCTP labelled human DNA shows that the range of the ligation products transformed into *S. pombe* from around 50 kb to 500 kb (data not shown). It is likely that some degradation occurred therefore reducing the amount of high molecular weight DNA for transformation. Nevertheless, twenty-two transformants were recovered with a SPARC containing human inserts ranging from approximately 50-400 kb.



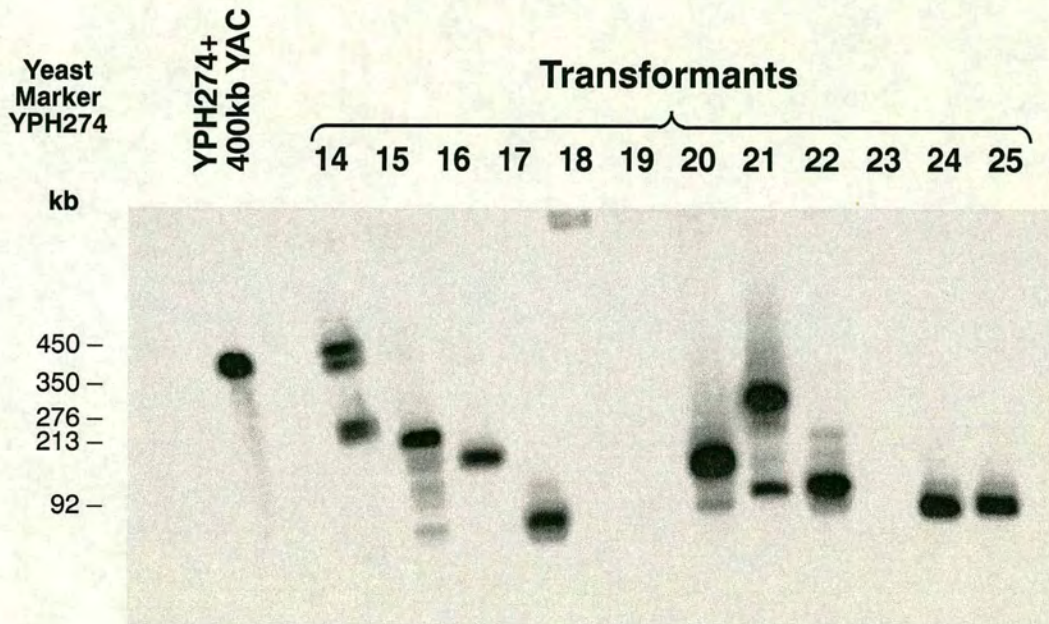
**Figure 5.2. DNA from human SPARC transformants.**

Phosphor-images of DNA from 25 human SPARC transformants subject to PFGE, transferred to genescreen and probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled total human DNA. The smaller bands of the yeast strain YPH274 are indicated. The PFG conditions were 150 volts, 20 to 70 seconds, 24 hours, 0.5xTBE, 12°C.





Probe: Total Human DNA



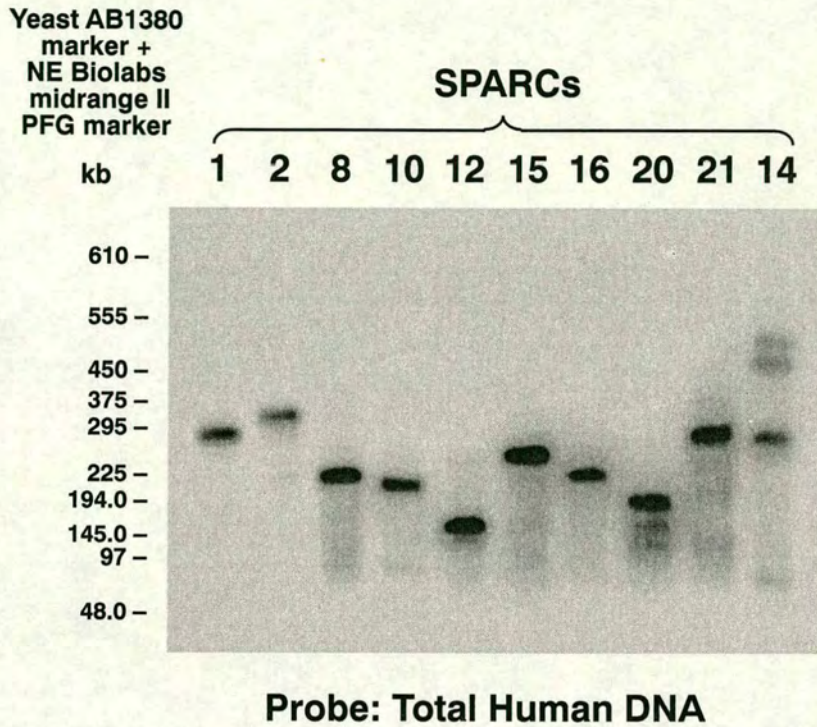
Probe: Total Human DNA



Although there was some background smearing in figure 5.2, a few of the transformants have obvious multiple bands. Transformant 22 has a strong SPARC band of around 90 kb and a second band of reduced intensity at 180 kb. This may be due to telomere association or a rearrangement as seen with the cosmid SPARCs (chapter 4). The most prominent SPARC bands in transformants 10, 20 and 21 approximately 200 kb, 150 kb and 300 kb respectively, in addition all also contain a smaller band migrating about 90 kb. This smaller band may be a deletion derivative of the larger SPARC. Transformant 14 has three apparent extrachromosomal bands with approximate sizes of 200 kb, 350 kb and 400 kb. There are several possible explanations for the presence of these multiple SPARC bands. They could be due to rearrangements occurring at the time of transformation or as a result of more than one ligation product being transformed into a single *S. pombe* protoplast. The events which lead to the generation of these multiple bands will be discussed later.

Construction of the cosmid SPARCs demonstrated that human fragments of 30-50 kb can be cloned and propagated intact in *S. pombe*. SPARCs containing small human inserts have been analysed in the previous chapter, therefore the ten transformants generated above with the largest bands were selected for further study. High molecular weight DNA prepared from these transformants (1, 2, 8, 10, 12, 14, 15, 16, 20 and 21) was subject to PFGE under conditions designed to give a more accurate size of each SPARC band. The resulting pulsed field gel was analysed by Southern transfer to genescreen and probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled total human DNA. Figure 5.3 shows the resulting phosphor-image and table 5.1 presents the sizes of the SPARCs.





**Figure 5.3. Ten human SPARCs.**

Phosphor-image of DNA from 10 human SPARC transformants with the largest extrachromosomal bands subject to PFGE, transferred to genescreen membrane and probed with  $(\alpha\text{-}^{32}\text{P})\text{dCTP}$  labelled total human DNA. The sizes of the smaller yeast strain AB1380 (225 - 610 kb) and New England Biolabs mid range marker II (48 - 194 kb) have been indicated. These transformants contained one or more SPARCs which had human inserts of 90 kb or greater. The PFGE conditions used were 150 volts, 50 -150 seconds, 24 hours, 0.5xTBE, 12°C.



**Table 5.1 SIZE OF SPARC BANDS IN EACH TRANSFORMANT**

<b>TRANSFORMANT</b>	<b>SIZE OF SPARCS (kb)</b>	<b>NO. OF SPARC BANDS</b>
1	280	1
2	300	1
8	220	1
10	210 , 90	2
12	145	1
14	260, 425, 470	3
15	240	1
16	220	1
20	170, 145, 110	3
21	270, 110	2

Six of these transformants contained a single SPARC, with a human insert, ranging from 145-300 kb. The other four transformants contained 2 or 3 SPARC bands each with human inserts. Including all the transformants with multiple bands, the SPARC bands range from around 100 kb up to 470 kb. The average insert size was estimated to be around 240 kb.

#### **5.4 ESTIMATION OF SPARC COPY NUMBER**

As this *S. pombe* large DNA cloning system uses acentric vectors, it was expected that the SPARCs would segregate randomly during mitosis and perhaps accumulate to a copy number of greater than one per cell as is found with circular ars plasmids

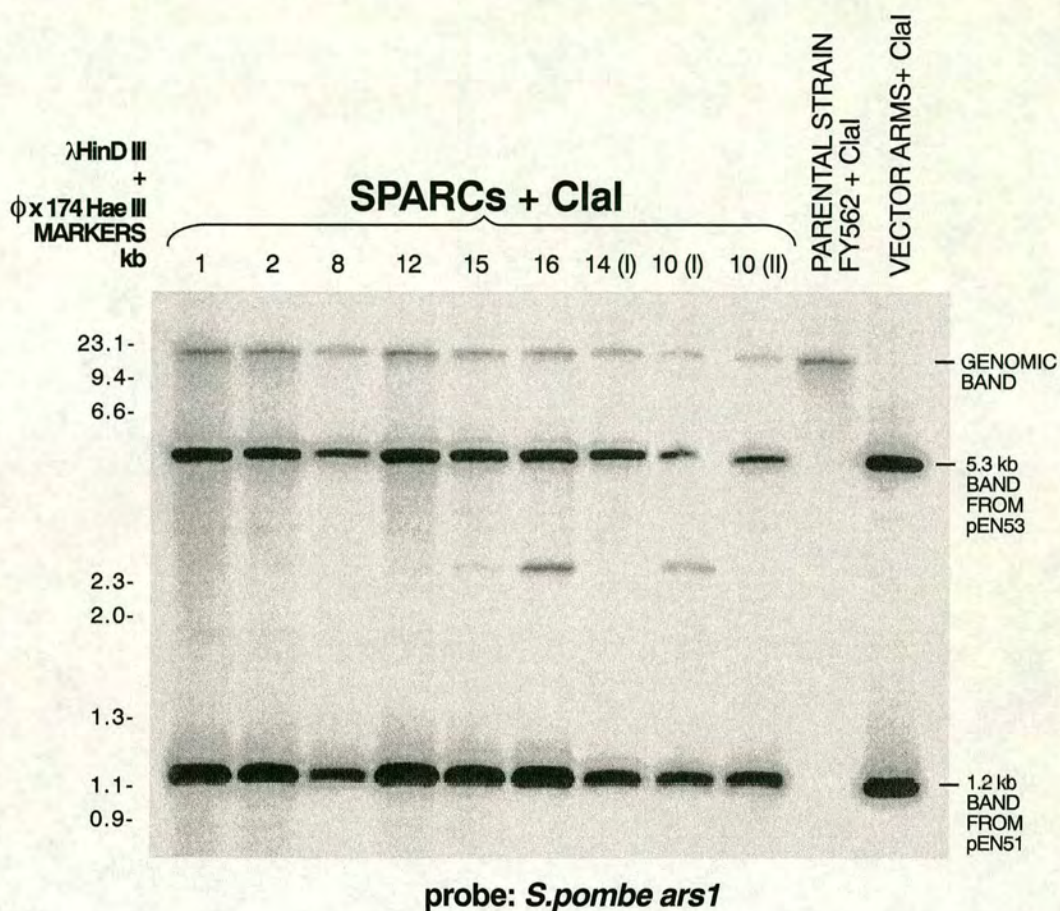


(Maundrell *et al.*, 1988). On pulsed field gels stained with ethidium bromide and viewed using an UV transilluminator, even the small SPARC bands appeared more intense than expected if present at one copy per cell (refer to figure 4.3). An estimate of the average copy number per cell can be determined by comparing the intensity of a single copy genomic sequence to the same sequence on the SPARC vectors. Since the *ura4* gene is deleted in the parental *S. pombe* strain the only common sequence between the host genome and the SPARC vectors is *ars1*. *Ars1* maps at a single site on the left arm of chromosome I.

From the published sequences, several restriction enzymes were identified which cut in the SPARC vector arms but not in the *S. pombe ars1* sequence. Restriction analysis was carried out on the plasmid and the parental strains to determine the digest most suitable for separating the vector arm bands and the genomic copy. These digests were subject to gel electrophoresis, Southern blotted and probed with ( $\alpha$ -<sup>32</sup>P)dCTP labelled *ars1* DNA (data not shown). The most convenient digest pattern was produced by the *ClaI* restriction enzyme. Using *ClaI*, *ars1* DNA was isolated on a 1.2 kb band from pEN51 vector and on a 5.3 kb band from pEN53 vector, while the genomic copy of *ars1* was on a fragment of around 20 kb. Figure 5.4 shows the results from the Southern blot used for the copy number estimation.

The right-hand lane shows the resulting bands from the vector arms digest. The adjacent lane is the DNA digest from the parental strain, FY562, where the genomic band containing *ars1* was identified. DNA digests from 9 transformants are also shown in figure 5.4.





**Figure 5.4. Copy number of several SPARC strains.**

Phosphor-image of SPARC DNA digested with *Cla*I subjected to conventional gel electrophoresis, transferred to genescreen membrane and probed with ( $\alpha$ - $^{32}$ P)dCTP labelled *S. pombe ars1*. The sizes of the larger bands for the  $\lambda$ HindIII and  $\phi$ X174HaeIII digest markers are shown. *Cla*I digests are shown for SPARCs 1, 2, 8, 12, 15, 16, 14 (I - 470 kb band), 10 (I - 210 kb band) and 10 (II - 90 kb band). The genomic *ars1* band is shown in the DNA digest of the FY562 parental strain and vector arms alone digested with *Cla*I were loaded in the lane on the right. The comparison of the densities of the 5.3 kb vector arm band and the genomic band for each SPARC was calculated to determine an approximate average copy number per *S. pombe* cell.



The vector arm bands are much more intense than the genomic bands for each transformant. The Molecular Dynamic's Imagequant program was used to determine the relative density of each band. Firstly a box was drawn over one of the bands for analysis and this was copied and positioned over each of the other bands. Each box was individually numbered therefore once processed through the imagequant software the density of each band was assigned to the given box number. It was assumed that the density of the genomic band was equivalent to one copy per cell. Therefore in any one lane, the SPARC band density can be divided by the density of the genomic band to give an estimate of the copy number.

**Table 5.2 COPY NUMBER OF SPARCS SHOWN IN FIGURE 5.4**

<b>SPARC (size)</b>	<b>pEN53 band density</b>	<b>Genomic band density</b>	<b>Copy no.</b>
1 (280 kb)	55487	8620	6.4
2 (300 kb)	45773	7617	6.0
8 (220 kb)	20741	3503	5.9
12 (145 kb)	67751	8527	7.9
15 (240 kb)	48348	6394	7.5
16 (220 kb)	48197	7456	6.5
14 (470 kb)	32813	5621	5.8
10 (210 kb)	10047	3080	3.2
10 (90 kb)	16912	3041	5.6

Using these results the average copy number was calculated to be around 6.1 copies per cell. The copy number of SPARC 10 (210 kb) was lower than the rest by this estimate however the shape of the band suggests that the DNA was inefficiently transferred to the filter. As results were consistent for all samples analysed, the size



of the SPARC does not appear to affect its copy number within this size range and therefore the copy number of these acentric SPARCs is estimated to be between 5 and 8 copies per cell. These measurements probably overestimate the SPARC copy number since the large *ars1* genomic fragment (approximately 20 kb) may have transferred to the filter less efficiently than the 5.3 kb pEN53 fragment. An improved method for determining the copy number could be competitive PCR where the SPARCs are used to transform a strain which contains a partial deletion of the *ura4* gene. Using PCR the *ura4* sequences of the vector and the deleted version of *ura4* in the *S. pombe* genome can be amplified. From a comparison of the amounts of full length and deletion product an estimate of the average copy number per cell could be deduced.

## 5.5 TRANSFORMATION EFFICIENCY

Using lipofectin cationic liposomes, the transformation efficiency into *S. pombe* protoplasts for the plasmid pUR19 was consistently around  $1 \times 10^6$  cfu per  $\mu\text{g}$  of plasmid DNA (Allshire, 1990). The main reason for improving the transformation efficiency of *S. pombe* protoplasts was in anticipation of constructing a human library in a *S. pombe* YAC-like vector system. YAC transformation into *S. cerevisiae* spheroplasts has been shown to be around  $1 \times 10^4$  cfu per  $\mu\text{g}$  of high molecular DNA (Anand *et al*, 1989). However the transformation efficiency for *S. pombe* protoplasts using larger molecules of DNA has not been measured. An estimate of the transformation efficiency was made to determine the feasibility of eventually constructing a SPARC library with human inserts.



To determine the transformation efficiency with the SPARCs, two SPARCs with single stable bands were isolated to retransform *S. pombe*. The SPARCs selected for retransformation of *S. pombe* were 2 and 16 of 300 kb and 220 kb respectively. These SPARCs apart from *ars1* have no homology to endogenous *S. pombe* chromosomes. Estimation of the amount of SPARC DNA being transformed into *S. pombe* protoplasts is difficult as the DNA was melted while suspended in a gel slice. Estimations were made by comparing the amount of DNA transformed to a known amount of DNA present in agarose plugs prepared from minichromosome containing strains. The DNA from the minichromosome strain was prepared from a known number of cells and it was assumed that the minichromosome is in only one copy per cell. This minichromosome DNA and the DNA from SPARCs 2 and 16 were subject to PFGE, Southern blotted and probed with ( $\alpha$ - $^{32}$ P)dCTP labelled *LEU2* from which the amount of SPARC DNA used in the transformation of *S. pombe* was estimated. It was estimated that 30 ng of SPARC 2 (300 kb) and 10 ng of SPARC 16 (220 kb) transformed *S. pombe* resulting in 2035 and 1021 colonies respectively. The transformation efficiencies were  $7 \times 10^4$  cfu per  $\mu$ g of SPARC 2 DNA and  $1 \times 10^5$  cfu per  $\mu$ g of SPARC 16 DNA. The transformation efficiency of high molecular weight DNA into *S. cerevisiae* has been estimated to be  $1 \times 10^4$  cfu per  $\mu$ g of YAC DNA when producing human YAC libraries (Anand *et al.*, 1991). The larger of the two SPARCs used had a slightly reduced transformation efficiency which is not unexpected as the transformation efficiency may decrease as the size of the DNA molecules increase. This may be due to some degradation of this DNA or increased difficulty of transforming *S. pombe* with larger molecules.

From both SPARC retransformations, 10 single colonies were propagated and DNA was prepared from these. The resulting DNA was subject to pulsed field gel electrophoresis and the ethidium stained gels are shown in figure 5.5. Figure 5.5(a)



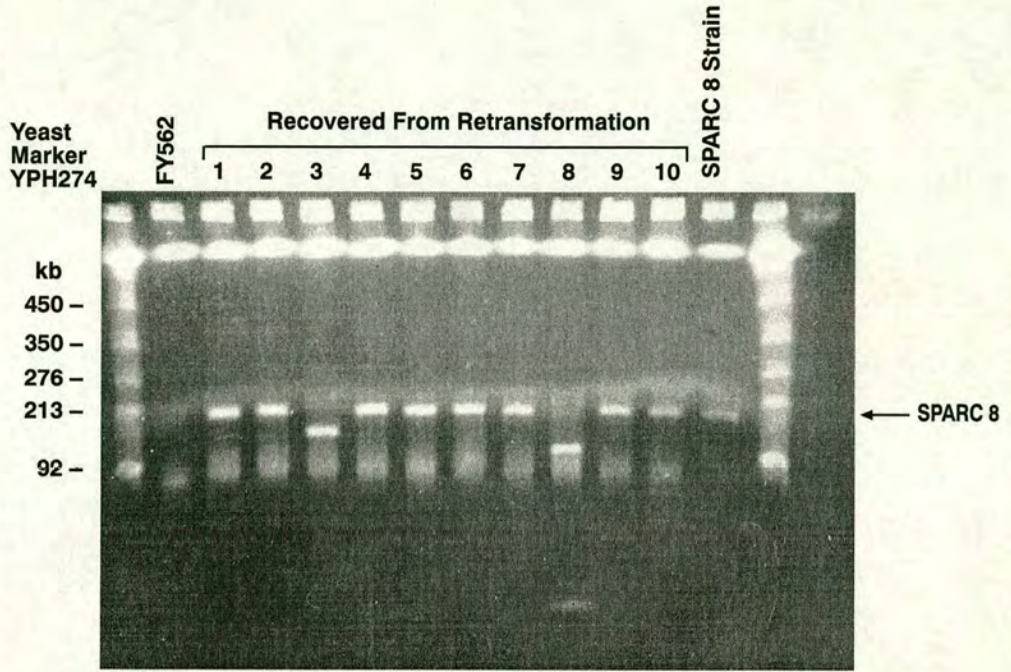
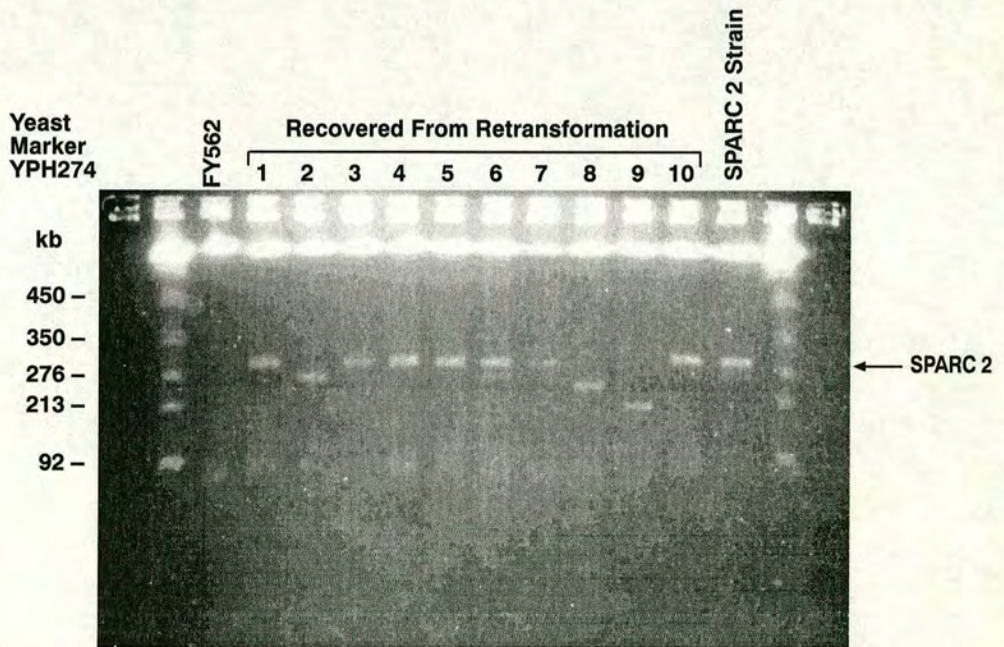
shows that two of the SPARC 8 retransformants carry a band reduced in size to approximately 150 kb (transformant 3) and 100 kb (transformant 8). In figure 5.5(b) DNA from 10 transformants resulting from retransformation of SPARC 2 into *S. pombe* was compared to DNA from the original strain. Three of these transformants had a SPARC band less than the expected size of 300 kb. Transformants 2, 8 and 9 have altered SPARC bands of approximately 270, 250 and 200 kb respectively. This shows that around 20-30 % of these transformants have undergone some rearrangements while the rest appear to remain intact. Nevertheless the SPARC system is fairly robust since SPARCs do not suffer gross rearrangements upon retransformation in most secondary transformants. The efficiency of transformation with large DNA is also high enough to potentially allow the construction of whole genomic libraries.



**Figure 5.5. DNA from *S. pombe* transformants recovered after retransforming with SPARCs.**

Photographs of DNA subject to PFGE, stained with ethidium bromide and visualised using ultraviolet transilluminator. The sizes of the smaller bands of the yeast strain YPH274 are indicated. DNA from 10 transformants recovered after retransformation of (A) SPARC 8 and (B) SPARC 2 into the *S. pombe* strain FY562 are shown. The sizes of the retransformed SPARCs are compared to the size of the original SPARC 8 and SPARC 2 bands. The PFGE conditions used were 150 volts, 20 to 70 seconds, 24 hours, 0.5xTBE, 12°C.



**A****B**



## 5.6 DISCUSSION

The experiments presented in this chapter demonstrate that SPARC vectors can be successfully used to clone *NotI* human fragments of approximately 50 to 450 kb in size. It was important to establish the capability of the *S. pombe* vector system to propagate a human library. The transformation efficiency was estimated to be in the order of  $1 \times 10^4$  to  $1 \times 10^5$  cfu per  $\mu\text{g}$  of DNA when retransforming *S. pombe* with two of the SPARCs. This is comparable to *S. cerevisiae* where the transformation efficiency has been shown to be  $1 \times 10^4$  cfu per  $\mu\text{g}$  of DNA (Anand *et al.*, 1991). The analysis of DNA from some of the resulting transformants has shown that 75 % of the SPARC molecules retransformed without rearrangement. The rest of these retransformants had reduced in size indicating that there had been loss of DNA. This demonstrates that rearrangements are not necessarily indicative of unclonable DNA but due to events which occur during the process of transformation. Such events might include DNA degradation or shearing during the manipulation of the DNA, as well as recombination during transformation.

The *S. cerevisiae* YACs are present in around one copy per cell. To increase YAC concentration for isolation, YACs can be subjected to a modification process which introduces a conditional centromere through which the YAC copy number can be increased (Smith *et al.*, 1990). As the SPARCs are acentric artificial chromosomes the copy number is already greater than one.

In *S. cerevisiae* YAC libraries, rearrangements have been found in some of the inserts (Anderson, 1993; Larinov *et al.*, 1994; Ling *et al.*, 1993; Wada *et al.*, 1994; Neil *et al.*, 1990). These rearrangements either result in chimaeric YACs (i.e. the insert has been derived from more than one chromosomal fragment) or loss of insert DNA. A



recombination deficient mutation in *S. cerevisiae* *RAD52*, has been shown to reduce chimaerism from around 50 % to 10 %. Further investigation is necessary to determine the degree of rearrangements found in the human inserts of the SPARCs. The *S. pombe* homologue of *S. cerevisiae* *RAD52*, *rad22*, has been identified. The *S. pombe* strain used in the studies presented here included a mutation which reduces mitotic recombination by 10 fold. However in the initial analysis of the first SPARCs isolated, several of them contained more than one extrachromosomal band. Such strains may have resulted from multiple SPARC molecules which transformed a single cell. This could be because lipofectin, used to assist the transformation, is thought to bind to the DNA and the cellular membrane. Therefore more than one SPARC molecule may bind to the lipofectin resulting in more than one SPARC per cell. Alternatively a rearrangement of the SPARC molecule occurred during transformation or upon replication in *S. pombe*. In the following chapter these *S. pombe* SPARCs with human DNA inserts are analysed in more detail to establish their mitotic and meiotic stability and to determine whether any rearrangements of the insert had occurred.



## **6 ANALYSIS OF SPARCS CONTAINING LARGE HUMAN INSERTS**



## 6.1 INTRODUCTION

In the previous chapter, twenty-two *S. pombe* transformants were recovered with SPARCs containing human inserts. Ten of these SPARCs were selected for further study because the size of their insert was greater than 100 kb. Analyses on these larger inserts are important to show the feasibility of cloning large fragments with the SPARC system. The efficiency of transforming *S. pombe* protoplasts with SPARC molecules was demonstrated to be approximately  $1 \times 10^5$  cfu per  $\mu\text{g}$  of large DNA molecules. The following experiments were designed to extend analyses of these transformants to determine the extent of rearrangements or DNA loss found in these SPARCs.

As these SPARCs are acentric it would be expected that they have an altered stability as they progress through mitosis and meiosis. It has been shown that these SPARCs exist at approximately 5-8 copies per cell. This increase in copy number did not appear to cause any rearrangements of the human insert. Rearrangements may occur through successive rounds of cell divisions therefore the SPARC can be analysed after different numbers of cell divisions. The SPARCs would be expected to be lost through meiosis due to their lack of a centromere. This can be tested by crossing the SPARC containing strain to strains without SPARCs and analysing the spores for transmission of the SPARCs.

In *S. cerevisiae*, the YACs are subject to rearrangement events which produce chimeric YACs or YACs with internal deletions (Anderson, 1993; Larinov *et al*, 1994; Ling *et al*, 1993; Wada *et al.*, 1994; Neil *et al.*, 1990). YAC libraries with mammalian DNA inserts can have up to 50 % chimerism (i.e. the insert contains a non contiguous fragment of DNA). Such a rearrangement can be detected using



Fluorescence *in situ* hybridisation (FISH). If a YAC contains a chimeric insert, then it may hybridise to more than one chromosomal location, whereas a YAC with a non-chimeric insert should give a signal at a single location (unless it contains low copy of sequences present at more than one location). Using FISH it would be possible to determine if the multiple banding pattern in SPARC strains 10, 14, 20 and 21 result from rearrangements. With internal deletions of the YAC insert it is more difficult to define loss of DNA as adjacent markers can still be present. Having randomly cloned *NotI* fragments from a complete human *NotI* digest into the SPARC system it should be possible to detect the equivalent sized fragments in a *NotI* digest of total human DNA by hybridisation.

## 6.2 ANALYSIS OF SPARCS THROUGH MITOSIS

It has been shown that acentric minichromosomes in *S. pombe* will go through mitosis with a loss rate of 50 % but are maintained at higher copy number per cell than minichromosomes with centromere sequences (Matsumoto *et al.*, 1990). An equivalent loss rate was demonstrated for the acentric SPARCs from *S. pombe* cells (data not shown). It is important to determine if passage of these SPARCs through successive cell divisions results in rearrangements of the human insert.

Approximately 20 to 30 cell divisions are required to produce a colony of 2 mm in diameter from a single cell. Each SPARC strain was streaked out on selective plates lacking both leucine and uracil thereby selecting for both SPARC arms to propagate single colonies. DNA was prepared from a 5 ml overnight culture derived from a single colony propagated on fresh selective media. This was the first colony generation. Another single colony was streaked out from this first selective plate onto a second plate with selective media to propagate a second generation of single



colonies. A third generation of single colonies was generated by repeating this procedure again. The DNA from all three generations for the SPARC strains was now analysed using PFGE to determine if the size of each SPARC was maintained through approximately 60 to 90 cell divisions. Figure 6.1 shows a photograph of an ethidium bromide stained pulsed field gel which shows the three generations of SPARCs 1, 8, 15 and 16, while only the first and third generations are shown for SPARC 2. SPARCs 2, 8, and 16 were maintained at the correct sizes of 360, 220, and 220 kb respectively through the 3 generations analysed here. SPARC 15 was also shown to be maintained at 240 kb, however by the third generation a smaller band of approximately 100 kb is also detected. DNA from SPARC 12 has also been demonstrated to be maintained at 145 kb throughout the three generations (data not shown). SPARC 1 was 300 kb after the first generation but decreased in size to 225 kb by the second generation. This rearranged SPARC was apparently stabilised at 225 kb in the second and third generations. This demonstrates that human DNA inserts can still be rearranged in *S. pombe* even after the transformant has been established.





**Figure 6.1. Maintenance of SPARCs through many cell divisions.**

Photograph of DNA subjected to PFGE, stained with ethidium bromide and viewed under ultraviolet transilluminator. The sizes of the smaller chromosomal bands of the yeast AB1380 marker strain are shown. DNA of SPARCs 1, 8, 15 and 16 are shown through three generations and SPARC 2 through generations 1 and 3. The PFGE conditions were 150 volts, 20 to 70 seconds, 24 hours, 0.5x TBE, 12°C.



Transformants 10, 14, 20 and 21 have been described in chapter 5. They have multiple SPARC bands therefore 10 single colonies from all four of these transformants were streaked out on fresh selective plates. From each isolate a single colony was propagated and DNA was prepared.

In figure 6.2 (a) five isolates from generation 3 of SPARCs 10 and 20 are shown. With the five isolates of SPARC 10 the two SPARC bands were separable into different strains. Four isolates (2-5) contained the 210 kb SPARC band while the other (isolate 1) contained the 110 kb SPARC band. For SPARC 20 three different isolates were derived, the first two lanes (isolates 1 and 2) contained a band of 110 kb (one of which contained additional faint bands), the third isolate had a band of 130 kb, while isolates 4 and 5 had a band of approximately 90 kb. These are substantially smaller than the 170, 145 and 110 kb bands detected in the original strain indicating that the human insert is rearranging after the transformation.

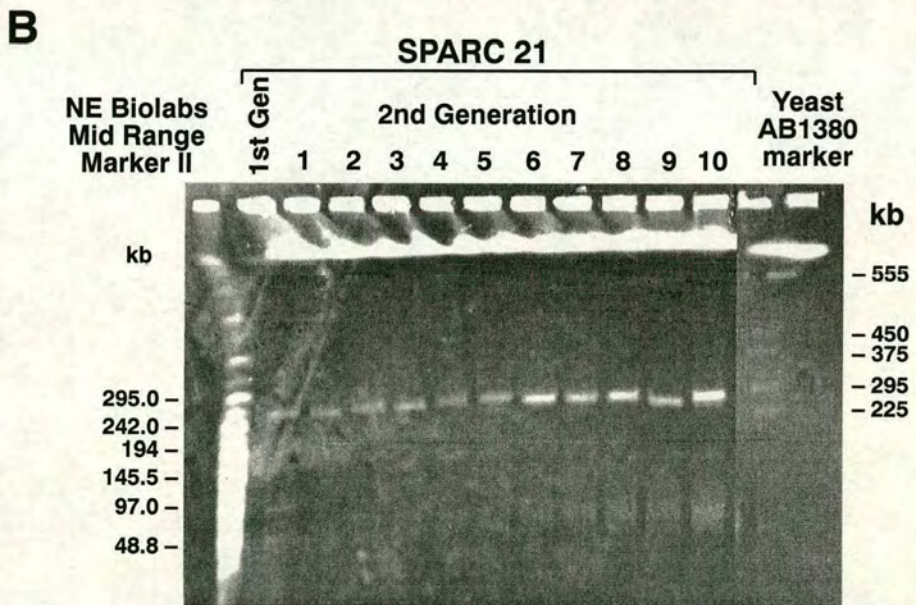
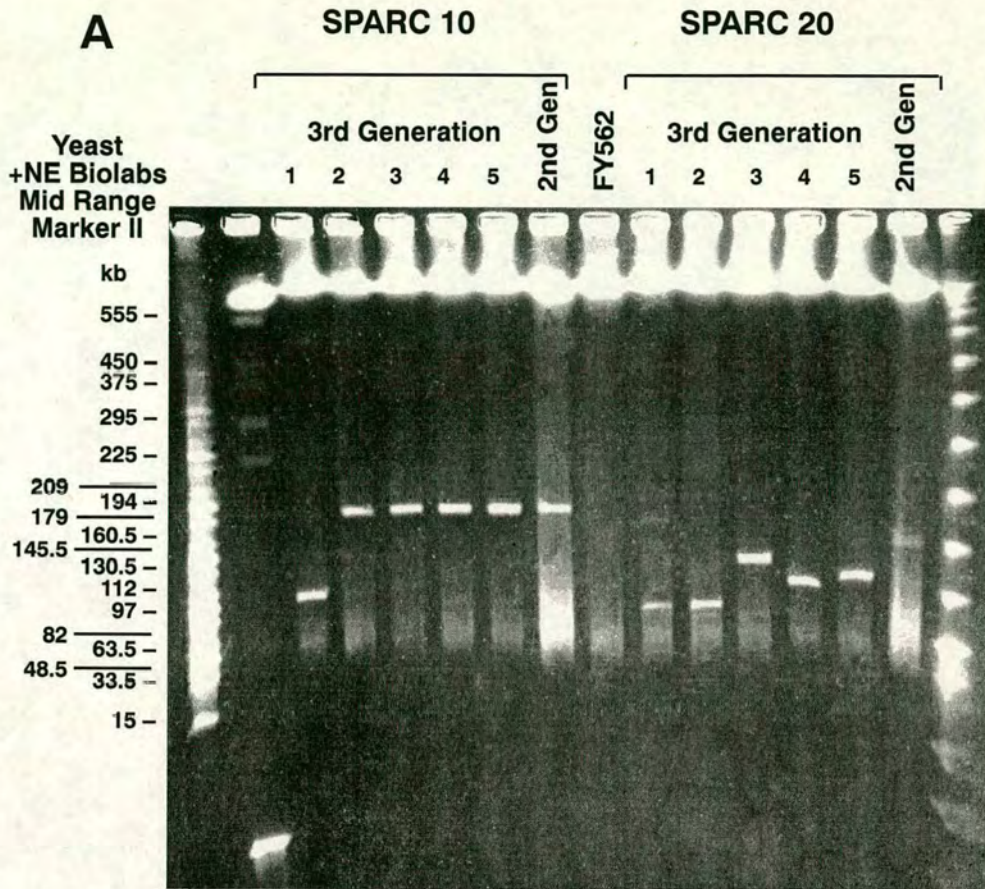
In figure 6.2 (b) 10 second generation isolates of SPARC 21 are shown containing bands of 270 kb and 90 kb. The 270 kb band is seen in all 10 isolates with only one of them retaining the smaller 90 kb band. The smaller band of SPARC 21 may have been produced by a rearrangement included the loss of vector arm sequences therefore this product cannot be from the larger SPARC band on the media lacking both uracil and leucine.



**Figure 6.2. Maintenance of SPARC bands from strains with more than one initial band.**

Photographs of DNA subjected to PFGE, stained with ethidium bromide and viewed using an ultraviolet transilluminator. The sizes of the smaller bands of yeast AB1380 strain and the New England Biolabs PFG marker II have been indicated. (A) shows five 3rd generation isolates of the SPARC 10 compared to the 2nd generation isolate in lanes 3 to 8 and SPARC 20 results in lanes 10 to 15. The DNA from the parental strain, FY562, is shown in lane 9. (B) 10 isolates derived from a second generation of SPARC 21. The PFGE conditions used were 150 volts, 20 to 70 seconds, 24 hours, 0.5xTBE, 12°C.



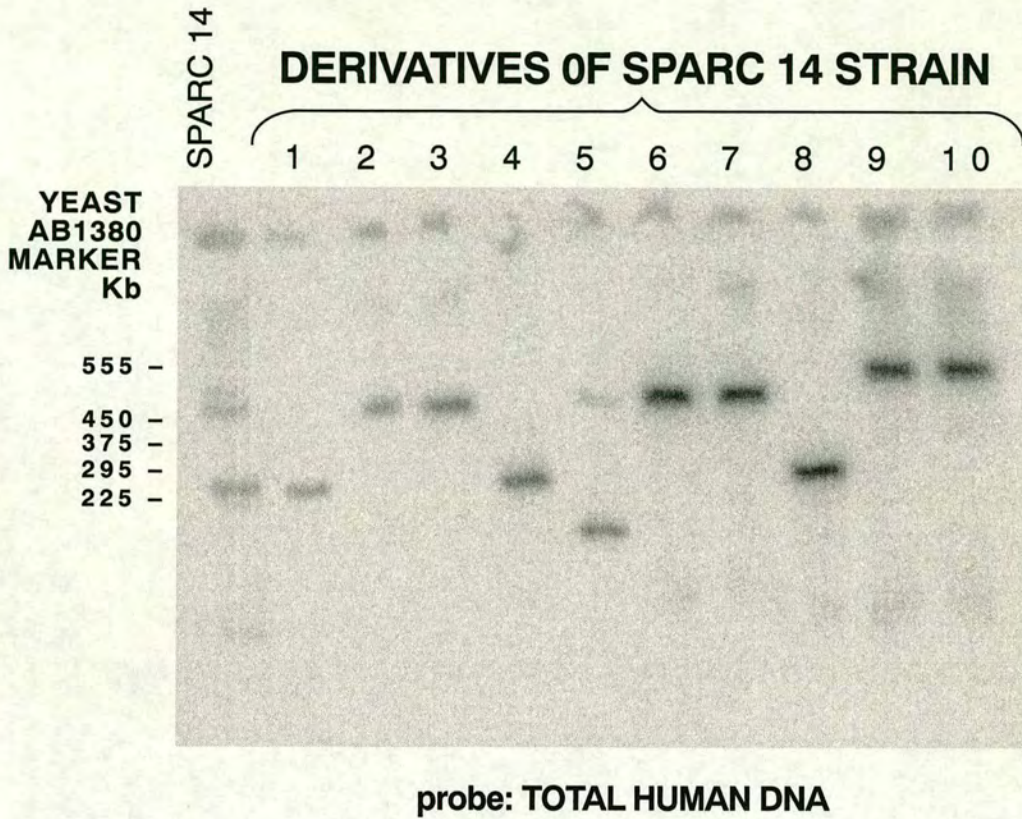




The DNA from ten isolates from the second generation of SPARC 14 are shown in figure 6.3. Isolates 1, 4 and 8 contain a single band of 260 kb, isolates 2,3, 6 and 7 have a band of 425 kb while isolates 9 and 10 have a band of 470 kb. The fifth isolate contained a faint band of 425 kb with a smaller band of around 100 kb.

These results show that from transformants containing multiple SPARCs, strains can be derived which contain a single SPARC band. These isolates usually have a SPARC band similar in size to one of the original bands from the founder strain. Occasionally strains were detected with novel bands therefore suggesting that the SPARC insert had rearranged. Most of these isolates retained the ability to be maintained on media lacking leucine and uracil therefore the SPARC bands recovered have sequences from both vector arms. The bands isolated from transformant 20 clearly have altered in size suggesting continual rearrangement of this SPARC occurs. However as the bands isolated from transformants 21, 10 and 14 appear to be stabilised as different isolates, it is possible that these could still be derived from multiple SPARCs transformed into a single cell. Alternatively, they may result from rearrangements of a single founder SPARC during the transformation or in the growth of the initial colony. FISH analysis could distinguish between these two possibilities if fluorescent probes were prepared from each isolate and used on a human metaphase spread. This will be discussed later in this chapter.





**Figure 6.3. Isolation of single SPARC bands derived from transformant 14.**

Phosphor-image of DNA subject to PFGE, transferred to genescreen membrane and probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled total human DNA. The marker bands shown are from the smaller bands of yeast strain AB1380. This shows ten 2nd generation derivatives of SPARC 14. The PFGE conditions used were 150 volts, 20 to 70 seconds, for 24 hours, in 0.5xTBE, 12°C.



### 6.3 DETERMINING LOSS OF SPARCS THROUGH MEIOSIS

A simple biological question to be asked is whether these acentric SPARCs can be transmitted through meiosis. The absence of a centromere on the SPARCs results in an increased copy number per cell, however they are otherwise stably maintained under selection. Matsumoto *et al.* (1991) have shown that partial loss of centromere sequences affects minichromosomes through meiosis, however the acentric minichromosome was not tested. It is possible that the meiotic loss of these acentric SPARCs may be greater than minichromosomes containing a centomere. To analyse the loss rate SPARC containing strains were crossed to the opposite mating type of the parental strain and resulting spores tested for the presence of the SPARC. Each cross was carried out in duplicate. The isolated spores were counted using haemocytometer and approximately 500 were plated on each plate. An equal volume of the spore suspension from each cross was spread on to non selective (YES) and selective (EMMG) plates.



**TABLE 6.1 NUMBERS OF SPORES THROUGH MEIOSIS CONTAINING SPARC**

CROSS h <sup>+</sup> parental strain with	NUMBER OF COLONIES ON YES	NUMBER OF COLONIES ON EMMG	PERCENTAGE OF SPORES WITH SPARCS
h <sup>-</sup> SPARC 2 strain	497	6	1.1
	1305	18	
h <sup>-</sup> SPARC 8 strain	533	6	1.2
	1277	17	
h <sup>-</sup> SPARC 16 strain	553	30	5.4
	ND*	ND*	
h <sup>-</sup> SPARC 14 (470 kb) strain	700	35	5.5
	632	38	
h <sup>-</sup> parental strain	1040	0	0
	590	0	

\*ND not done (the plates used for this cross became contaminated).

The data in table 6.1 shows that very few of the SPARCs go through meiosis without being lost. Several of the resultant colonies were grown up and DNA agarose plugs were prepared. These were run out on a pulse field gel and the SPARCs were of the correct size when compared to the parental SPARCs (data not shown). This data suggests that SPARCs have difficulty going through meiosis but the few that survive seem to be intact, i.e. they have both selectable markers and are the right size on pulse field gel analysis.



## 6.4 ANALYSIS OF SPARCS FOR CHIMERIC INSERTS

In some *S. cerevisiae* YAC libraries with mammalian inserts, it has been shown that approximately 50% of inserts were chimeric, i.e. inserts from more than one chromosomal location (Anderson, 1993; Larinov *et al*, 1994). YACs with chimeric inserts were thought to be caused by co-ligation events or recombination between repetitive elements of two DNA molecules which had co-transformed a *S. cerevisiae* cell. Wada *et al* (1994) have shown that suppressing co-ligation events does not significantly reduce the chimerism in YAC libraries containing mammalian inserts. In addition YAC libraries with mammalian inserts were particularly affected compared to other eukaryote libraries, such as the *C. elegans* YAC library reported to have only 5% chimerism (Coulson *et al*, 1991). This problem has been partially alleviated by using a recombination deficient strain, i.e. *RAD52* mutation, which reduces the chimeric inserts to approximately 10% (Larinov *et al*, 1994; Ling *et al*, 1993). It is important to assess if chimerism is found to a similar degree in the SPARC system. The chromosomal position of the human insert from the SPARCs can be determined using FISH to metaphase spreads of a human cell line.

DNA from several SPARCs was isolated from several agarose plugs (prepared from a culture of the transformant strain) in a 3-4 cm well on a LMP pulsed field gel. Biotinylated probes were prepared from these isolated SPARC DNA's using catch-linked PCR. This was used as a probe for FISH onto metaphase spreads of human chromosomes and was detected using anti-biotin IgG antibodies conjugated to Texas red, seen as a red signal on the FISH images. The chromosomes were visualised using a DAPI stain, seen as blue. The blue image of the chromosomes was converted to a black and white image to give a banding pattern to identify the chromosomal position of the Texas red signal. The images were captured and processed using IP-



lab Digital Scientific (Cambridge) extensions for the Iplab Spectrum software.

Approximately 20 images were analysed to determine the chromosomal position of the human insert from the SPARC. An example for each of the SPARCs analysed is shown in figure 6.4.

FISH analysis was used to determine the chromosomal position of SPARCs 2, 8, 16, 21 (270 kb band), 14 (425 kb band) and 14 (470 kb band). In figure 6.4, the chromosomal positions of the human insert of the SPARCs are as follows: (a) SPARC 8 at 3p27-29; (b) SPARC 2 at 4q33; (c) SPARC 16 at 3p14; (d) SPARC 21 (270 kb band) on the end of 18q; (e) and (f) SPARC bands from transformant 14 of 425 kb and 470 kb respectively were both at 6q23.

The results show that all but one of the human inserts of the SPARCs tested were contiguous pieces of DNA. The FISH analysis with the SPARC 21 probe produced a strong signal at the end of 18q however in approximately 40 % of metaphases analysed there was additional signal on the human chromosome 19. This suggested that the human insert of SPARC 21 (270 kb band) may be chimeric. In addition, FISH analyses of both of the largest SPARC bands (470 kb and 425 kb) from transformant 14 were derived from the same chromosomal location. This demonstrates that the SPARC bands in this strain are related indicating that there has been a rearrangement event during transformation, probably by deletion of the largest DNA molecule (470 kb) giving rise to the smaller molecules.

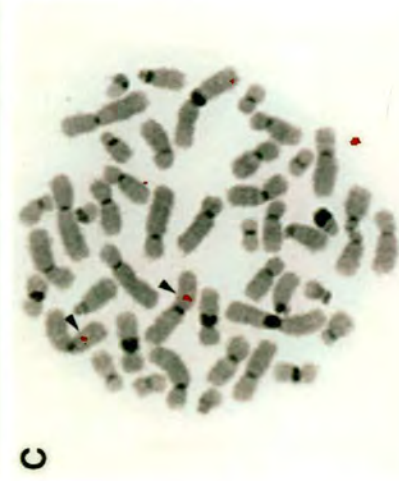
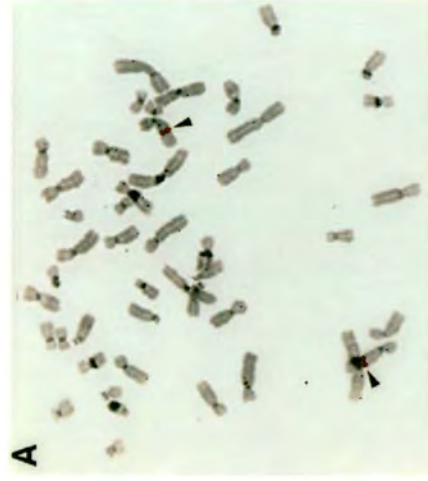


**Figure 6.4. Using FISH to determine chromosomal location of the SPARCs human insert**

Fluorescence *in situ* hybridisation of biotinylated SPARC probes plus competitor cot1 DNA onto human metaphase spreads detected using anti-biotin IgG conjugated with Texas red. The signal for each predominant signal on homologous chromosomes is indicated using a black arrowhead.

Approximately 20 FISH images were viewed to determine the chromosomal positions for each SPARC probe. One image was shown as an example for each SPARC with (A) SPARC 8, (B) SPARC 2, (C) SPARC 16, (D) SPARC 21 (210 kb band), (E) SPARC 14 (470 kb band) and (F) SPARC 14 (425 kb band).





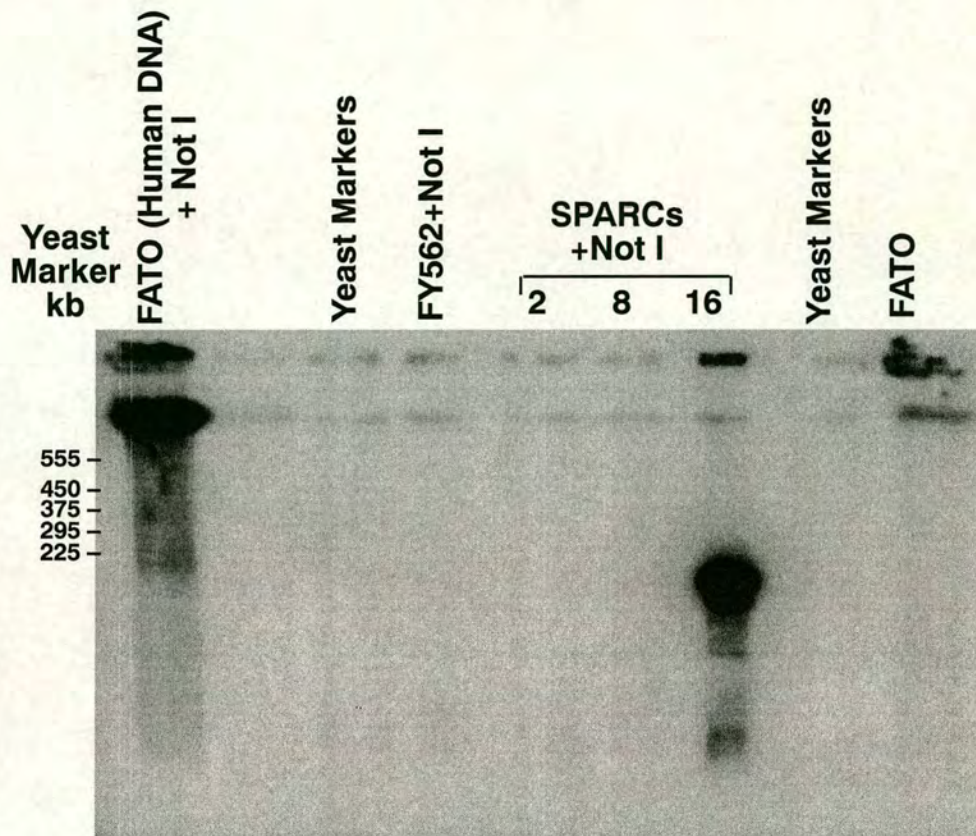


## 6.5 ANALYSIS OF SPARCS TO INVESTIGATE POSSIBLE INSERT REARRANGEMENTS

As previously described the human DNA fragments were cloned at random from a complete *NotI* digest of DNA from the FATO lymphoblastoid cell line. This should allow the insert size in each of these SPARCs to be compared to the original *NotI* fragment in a total *NotI* digest of FATO DNA. The method used to isolate unique probes from a large human insert *S. pombe* SPARCs was to utilise primers to the Alu repeats which transcribe out of the repeat into the unique intervening human sequences i.e inter Alu PCR (Arveiler and Porteous, 1992). If two inverted Alu repeats lie within several kilobases of each other the sequence between them would be amplified by PCR. From each primer used, several bands can be detected and these can be used to produce a ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled probe. A probe with a high specific activity was generated by decreasing the amount of input DNA into the reaction therefore increasing the ratio of hot to cold DNA within the probe. Any repetitive DNA which remained in these probes was competed out using 200  $\mu\text{g}$  of cold total human DNA or cotI DNA (a fraction of human DNA containing highly repetitive material). *NotI* digests of the SPARC DNA and the human DNA (FATO) were subject to pulsed field gel electrophoresis. The Southern blots from these pulsed field gels were probed with these competed probes.

The result from one blot is shown in figure 6.5 where the *NotI* insert from SPARC 16 can be seen. There is a faint band present in the FATO lane which is the same size as the 225 kb *NotI* band from SPARC 16.





**Probe:Alu PCR From SPARC 16  
(Plus Competing Human DNA)**

**Figure 6.5. Comparing the human *NotI* fragment of SPARC16 with total human DNA digest.** Phosphor-image of DNA subject to PFGE transferred to genescreen and probed with ( $\alpha$ - $^{32}\text{P}$ )dCTP labelled Alu PCR product with competing cold total human DNA. The yeast strain AB1380 smaller chromosome marker bands are shown. The results demonstrate the hybridisation of the unique probe from the SPARC 16 insert to the FATO *NotI* digest and the SPARC 16 *NotI* digest. The DNA from SPARCs 2 and 8 were used as negative controls to ensure the probe was unique to SPARC 16. The DNA from the parental strain, FY562, was also included. The PFGE conditions used were 150 volts, 20 to 70 seconds, 24 hours, 0.5xTBE, 12°C.



To produce clearer results the conditions used would have to be optimised further to conclusively determine whether any rearrangement had occurred. It is more difficult to determine the optimal conditions required to detect unique sequences on human digests from Southern blot membranes transferred from pulsed field gels than those from conventional gels because the size of the fragments being transferred are larger. Pulsed field gel electrophoresis enables large DNA fragments to be separated by switching the direction of the current thereby allowing large fragments of DNA to be separated from the total DNA sample (Schartz & Canter, 1984). The time lapse between each pulse switch is important to the length of fragments isolated. This process may not efficiently separate large DNA fragments to the same degree as the DNA will be in varying states of entanglement. The FATO DNA digest with many fragments within the size range studied may not separate out as well as the SPARC sample. With this and the possibility of some degradation of DNA in the FATO sample may account for the background signal in the FATO DNA digest.



## 6.6 DISCUSSION

Five of the largest SPARCs analysed were stably maintained as linear episomes with no insert rearrangements detectable through three generations of colony formation, which is equivalent to at least 90 cell divisions. SPARC 1 reduced in size from 300 kb to 225 kb which was then stabilised, presumably due to a post transformation rearrangement. All of the SPARCs were isolated from each of the strains containing multiple bands as single band derivative strains except the 110 kb band from SPARC 21. Therefore all of these single band derivatives must have retained the vector arms to be maintained extrachromosomally on plates lacking uracil and leucine. FISH analysis with probes derived from the 425 kb and 470 kb bands from SPARC 14 on human chromosome metaphase spreads show that both inserts were derived from the same chromosomal location. Thus in this strain which contained three SPARC bands, at least the 425 kb and 470 kb bands are related, and probably result from insert rearrangements during transformation. The multiple bands of SPARC 20 also undergo further rearrangements. Recombination has already been shown to be the principal cause of rearrangements resulting in chimeric inserts of *S. cerevisiae* YACs which is alleviated using the recombination deficient mutant, *rad52* (Larinov *et al.*, 1994; Ling *et al.*, 1993). The *RAD52* mutation also proved effective at stabilising insert rearrangements of some repetitive sequences from the human Y chromosome (Neil *et al.*, 1990). The strain used here, *rec55-36*, had a 10 fold reduction in mitotic recombination (Gysler-Junker *et al.*, 1991), however this recombination mutation alone may not inhibit all SPARC insert rearrangements in *S. pombe*. The discovery of the *S. pombe* *RAD52* homologue, *rad22* (Ostermann *et al.*, 1993) may prove useful for suppressing some of the SPARC insert rearrangements.



It is encouraging that there is little evidence for chimerism in the SPARC inserts analysed, however these inserts were relatively small when compared to some of the current YACs. Experiments to determine whether many internal deletions had occurred during transformation by comparing the size of the SPARC *NotI* insert to the original fragment in a FATO digest proved to be more difficult than expected. Analysis indicated that the human insert from SPARC 16 possibly had been cloned without insert rearrangement. If the conditions for these Southern blots were further optimised more accurate results may be produced for the other SPARCs. Producing a unique probe by another method may give results. Unique probes can be isolated from end clones from random fragments. Internal deletions may be detected if the restriction digest profiles of a fragmented SPARC molecule were compared to the same restriction digest of FATO DNA using these probes. Differences between the banding patterns of the FATO DNA digest and the SPARC DNA digest would indicate insert rearrangements.

Most of the SPARCs have been shown to be maintained without rearrangements and even through meiosis, when a high percentage of the SPARCs were lost, those SPARCs recovered were intact. It was apparent that some SPARC insert rearrangements occur during transformation or during initial colony growth and as a result transformants with more than one SPARC band were produced. Further major insert rearrangements were limited to isolates of two SPARCs. It has been previously demonstrated that there are similar events in some *S. cerevisiae* YACs with human inserts (as shown on figure 3.2; Anderson, 1993). Rearrangements are thought to occur more in YACs with mammalian inserts due to recombination between repetitive sequences which are more abundant on mammalian DNA (Larinov *et al*, 1994; Ling *et al*, 1993; Wada *et al.*, 1994; Neil *et al.*, 1990). By comparison only 5 % chimerism was reported for a *C. elegans* YAC library (Coulson



*et al*, 1991). SPARCs are present at several copies per cell allowing both intra and inter molecular recombination to occur. It is therefore perhaps not surprising that rearrangements were detected in the SPARCs.



## 7 DISCUSSION



## 7.1 ANALYSES OF THE SPARC VECTOR SYSTEM

During the course of this project an artificial chromosome vector system for cloning large fragments of heterologous DNA in *S. pombe* has been developed. This *S. pombe* system has similarities to the YAC system in *S. cerevisiae* (Burke *et al.*, 1987). *S. pombe* was tested for its ability to carry human sequences on extrachromosomal fragments. These analyses were useful for determining the feasibility of constructing a human library in *S. pombe*.

The investigations described in the preceding chapters have shown that the *S. pombe* system can clone human fragments from 50 kb to several hundred kilobases. The *S. pombe* artificial chromosome (SPARC) vector arms were derived from plasmids which had been constructed to test that cloned *S. pombe* telomeres were functional (Nimmo *et al.*, 1994). Each vector arm is similar in structure and included a selectable marker (either *S. pombe* *ura4* or *S. cerevisiae* *LEU2*), an *S. pombe* replication origin (*ars1*) and an array of *S. pombe* telomere repeat sequences. *S. pombe* centromeres, with an inverted repeat structure of at least 35 kb, are considerably larger and more complex than the simpler structure of the *S. cerevisiae* centromere (reviewed Clarke, 1990; Carbon and Clarke, 1990). To include *S. pombe* centromeric sequences as a component of the SPARC vectors is complicated because of their size and inability to be cloned intact in bacterial vectors. As previously shown with acentric minichromosomes (Niwa *et al.*, 1989) a centromere is not necessary for the maintenance of linear episomes under selective growth conditions. The mitotic and meiotic stability of SPARCs has been shown to be affected by the lack of a centromere. However rearrangements of the human insert itself were apparently not caused by this SPARC segregational instability. Acentric vectors have a copy number of greater than one per cell. For manipulations and analyses of



YACs, high concentrations of YAC DNA are preferable. The detection of these sequences is improved if the DNA is more concentrated and small volumes of DNA in solution are used for transfer into mammalian cells. *S. cerevisiae* YAC vectors have been constructed with conditional centromeres which when functionally inactivated result in an increased copy number of YAC molecules per cells (Smith *et al.*, 1990). Upon inactivation, additional selective pressures can be applied to produce 10 to 20 copies of the YAC per cell. Instead of preparing new libraries with new uncharacterised vectors, well studied YACs of particular interest can be modified by replacing the vector arms by homologous recombination and selecting for this alteration (Smith *et al.*, 1991; Gobin *et al.*, 1995; Taylor *et al.*, 1994). This process, known as retrofitting, is time consuming and requires analysis to ensure the modified YAC structure is correct. The results described here indicate that SPARCs are easily maintained as linear episomes when selecting for both selectable markers. In selective medium, approximately 50 % of the population of *S. pombe* cells within a culture contain the SPARC (data not shown). Therefore a centromere is not essential for the propagation of SPARCs provided selective pressure is maintained. The resulting cells contain an estimated average of 5 copies of each SPARC per cell.

Initial analyses demonstrated that human fragments of DNA of 30 to 50 kb can be cloned successfully without rearrangement. By cloning random fragments of human DNA into the SPARC system, several transformants were recovered with SPARCs containing human inserts of greater than 100 kb. Several of the SPARCs were tested to determine if the human inserts were chimeric. These results have shown that all of the SPARCs analysed appeared to contain contiguous human DNA inserts. By comparison, it has been reported that some mammalian YAC libraries have up to 50% chimerism (Larinov *et al.*, 1994; Ling *et al.*, 1993) although the RAD52 mutant strain effectively reduced this percentage. The data, shown in chapter 6, suggested



that internal deletions may be a problem in some SPARCs, even in the SPARCs with no obvious rearrangements small deletions may remain undetected. Some of the transformants had multiple extrachromosomal bands which can be propagated individually. Using FISH onto human metaphase spreads it has been shown that in one of these strains the human DNA insert of two individual bands originated from the same chromosomal location, suggesting that a rearrangement of some sort occurred. The rearrangements which led to these multiple bands was probably due to recombination events occurring during the transformation, as in subsequent analyses no further rearrangements were detected. Additional analysis is required on the remaining SPARCs to ensure that no other rearrangements are detected. Such analyses should give an indication of what insert rearrangements to expect when producing a human library. Many of the problems associated with YACs were discovered when the YAC libraries were in extensive use (Anderson, 1993). Whether these results shown in this study are representative of all SPARCs is uncertain and will require the construction and analysis of a human library in the *S. pombe* system. It is likely that similar problems will be encountered in SPARC libraries to those found in YAC libraries.

The eventual construction of a human library in the SPARC system requires a good transformation efficiency of large recombinant DNA molecules into *S. pombe*. The addition of lipofectin to the *S. pombe* transformation protocol increased the transformation efficiency by up to 50 fold for plasmid transformation (Allshire, 1992). The transformation efficiency of SPARC DNA into *S. pombe* protoplasts was estimated to be approximately  $1 \times 10^5$  cfu per  $\mu\text{g}$  of SPARC DNA. This is roughly equivalent to the transformation efficiency of YAC DNA of a similar size transformed into *S. cerevisiae* (Anand *et al.*, 1991), therefore it should be feasible to construct a human library in the *S. pombe* SPARC vector system. When these



SPARC molecules were retransformed into *S. pombe* it was shown that around 75% remained intact, while the rest had undergone minor rearrangements producing linear SPARCs of reduced size. Therefore SPARCs can be easily transformed into other *S. pombe* strains with little rearrangement. Transferring SPARCs to another *S. pombe* strain may be required when modifying the vector arms to include other selectable markers. When a YAC carrying an insert with an introduced *HIS3* selectable marker was retransformed into *S. cerevisiae* it was found that 33% of resulting transformants had undergone a deletion of 80 to 260 kb encompassing the *HIS3* gene (Kouprina *et al.*, 1994). Retransforming this same *HIS3* YAC into an *S. cerevisiae rad52* background did not reduce the frequency of these rearrangements. Therefore with the retransformation of stable SPARCs into *S. pombe*, some rearrangements appear to occur.

Another inherent advantage of the *S. pombe* system is that all three chromosomes are greater than 3.5 Mb (Fan *et al.*, 1988; Kohli *et al.*, 1977; Robinow, 1977) so that SPARCs can be easily isolated without contamination by *S. pombe* DNA. This was demonstrated when the DNA from the transformants was subject to pulsed field gel electrophoresis where the SPARC bands were clearly separated from the host chromosomes (retained within the limiting mobility). By comparison, *S. cerevisiae* has 16 chromosomes which range in size from 225 kb to 1.6 Mb (Mortimer *et al.*, 1989). The result is that YACs tend to co-migrate with the endogenous chromosomes therefore isolated YAC DNA is frequently contaminated with yeast sequences. Clearly, as SPARCs can be well separated from endogenous *S. pombe* chromosomes using PFGE, there will be fewer problems of contamination of isolated SPARC DNA with *S. pombe* DNA.



## 7.2 FUTURE INVESTIGATIONS OF THE SPARC SYSTEM

To determine if SPARCs had suffered internal deletions, unsuccessful attempts were made to compare the size of the SPARC human *NotI* inserts to the original *NotI* fragment in total human DNA. Several modifications and improvements to these analyses are required. Unique probes from the human inserts of these SPARCs were derived using PCR to amplify an inter Alu fragment (Arweiler and Porteous, 1992). Suppression of the repetitive sequences which may remain in the probe by preannealing with cot1 DNA or total human DNA was not always effective and produced a high background on Southern blots. The result shown in figure 6.5 demonstrates that results can be attained with the Alu PCR probe. This background in the human track may be due to incomplete digestion or to DNA degradation. Inclusion of multiple human DNA samples prepared separately may corroborate any results. Alternatively rearrangements of the human inserts may be detected by comparing the restriction digest profile of the SPARC DNA to the human digest profile using the unique probes. Rearrangements of the human inserts would be detected when there was a difference in restriction digest patterns.

Some areas of the human genome are not represented in YAC or cosmid libraries (Anderson, 1993). This may be due to an inherent instability of insert DNA as a result of an abundance of repetitive sequences. One notable region which is particularly unstable when cloned in YACs are mammalian centromeric repeats (Neil *et al.*, 1990). While YACs libraries may be constructed in a recombination deficient *S. cerevisiae* strain or a combination of YAC, cosmid, BAC and PAC libraries may prove useful, the addition of a human library in the *S. pombe* SPARC system will serve to complement these vector systems. This may enable any existing gaps in the human genome contigs to be filled. Therefore the construction of a human library in



the SPARC system may aid the development of contiguous maps for regions where other vectors appear to fail. Given that SPARCs can be manipulated using similar methods to those used for *S. cerevisiae* YACs then the *S. pombe* system lends itself to the functional analyses of inserts.

The longest mammalian centromere sequences cloned into *S. cerevisiae* YACs are around 150-200 kb (Neil *et al.*, 1990). Such sequences would have to be cloned and analysed in *S. pombe* SPARCs as a direct test of *S. pombe*'s ability to cope with tandem arrays of alphoid repeats. Since *S. pombe* has more complex repetitive centromere arranged in a long inverted repeat structure around a central core (reviewed Clarke, 1990; Carbon and Clarke, 1990), it is a possibility that *S. pombe* may have the ability to propagate arrays of mammalian centromeric repeats. In *S. cerevisiae*, mammalian centromeric sequences are thought to be rearranged due to their highly repetitive nature, however it has been hypothesised that incomplete replication may also play a role. It is thought that there are sequences throughout the mammalian genome that can act as replication origins which when transform yeast cells allowed artificial chromosomes to be replicated once per cell cycle. In the absence of any of these sequences replication may be restricted to 150 to 200 kb using the *ars* sequences from the vector arms (Alfano and Davis, 1995). Therefore, if mammalian centromere repeats lack *ars* activity then the only YACs recovered may be deletion derivatives of larger fragments from mammalian centromeres which are able to replicate completely in *S. pombe*. There are differences between the *ars* sequence requirements of *S. cerevisiae* (Murray & Szostak, 1983) and *S. pombe* (Maundrell *et al.*, 1988; Dubey *et al.*, 1996), therefore determining if sequences residing within mammalian centromere repeats can act as *ars* elements would indicate the suitability of *S. pombe* for propagating arrays of mammalian centromeric repeats. Pilot studies were carried out by cloning mouse minor satellite repeat and



human alphoid repeat into plasmids which lack a *S. pombe* replication origin. These plasmids were used to transform *S. pombe* and it was found that neither centromeric repeat contained a *S. pombe* replication origin based on the frequency of transformants (data not shown). This experiment suggests that arrays of alphoid and minor satellite sequences may not have the ability to replicate in *S. pombe*. To conclusively show if larger arrays of mammalian centromeric repeats can be propagated in *S. pombe* requires an attempt at cloning of such sequences into the SPARC vectors.

It is possible to identify chromosomal structures, such as centromeres and telomeres, using FISH on fixed *S. pombe* cells (Uzawa and Yanagida, 1992; Funabiki *et al.*, 1993). Using this technique, with a probe to the human sequence, SPARCs may be detected at various points through the cell cycle. These investigations may show how acentric molecules segregate in the absence of a centromere. In addition, it may be interesting to view the behaviour of the human sequences in *S. pombe*.

### **7.3 MODIFICATIONS TO SPARC VECTORS**

Human DNA fragments of several hundred kilobases can be cloned between SPARC vector arms. Around half of these SPARCs were shown to have rearranged but it was thought this occurred mainly during the transformation. Therefore it is feasible to proceed with the construction of a human library in the *S. pombe* system. Several modifications to the current system should be considered in order to construct a fully representative library with arrays of overlapping clones.



The plasmids used in this study to create the SPARC vector arms had been originally constructed to demonstrate that telomere repeats were able to seed new telomere growth in *S. pombe* (Nimmo *et al.*, 1994). The plasmid included a telomere cassette with two of the *S. pombe* telomere repeat sequences inverted around a kanomycin resistance gene and cloned within the pBLUESCRIPT plasmid polylinker. To simplify the SPARC plasmids a single *S. pombe* telomere repeat sequence need only be included in the vectors. Through cloning of the telomere cassette some restriction enzyme sites were introduced between the telomeres and the plasmid sequences in an inappropriate position to be used as cloning sites. This restricted the available cloning sites to *NotI* and *BamHI*. A polylinker could be constructed for the SPARC vectors with appropriate restriction enzyme sites, cloned between the tip of the telomere repeats and the plasmid sequences. This would produce vectors which contained all the necessary elements for the creation of SPARCs containing large DNA inserts with a variety of restriction enzymes from heterologous sources. From the data produced insert sizes of several hundred kilobases are possible. Given that YACs of greater than 1 Mb have been constructed, it should be feasible to clone fragments of an equivalent size into SPARCs.

The inclusion of *S. pombe* centromere sequences on the SPARC vectors may be advantageous where mitotic stability may be important for the maintenance of particular SPARC inserts. Although it has been demonstrated that the SPARCs can be maintained as linear extrachromosomal elements, there was no indication whether the stability of some human sequences may be compromised due to the high copy number per cell of SPARCs.

YAC insert rearrangements are thought to occur during transformation by recombination between repeats within the insert and such rearrangements are reduced



in *S. cerevisiae* by a recombination deficient mutation, *RAD52* (Larinov *et al.*, 1994; Ling *et al.*, 1993; Wada *et al.*, 1994). *RAD52* mutant strains of *S. cerevisiae* have been shown to reduce chimerism of mammalian inserts from 50 % to 10 % in a YAC library (Larinov *et al.*, 1994). Other *S. cerevisiae* recombination deficient mutants i.e. *RAD1* have been shown not to be as effective. Analysis of the inserts in the SPARCs constructed in this study do not give any indication of a high percentage of chimeric inserts. Of the 10 SPARCs studied, five had undergone a rearrangement, four of which occurred at the time of transformation and resulted in multiple bands being formed. Further investigation of one of these multiple SPARC strains has shown that two of the isolated bands were derived from the same chromosomal location. One of these strains containing multiple SPARC bands has gone through some further rearrangements after transformation. The other rearrangement which occurred resulted in the loss of insert DNA. It appears the SPARC system has few problems with chimerism, however the insert sizes were smaller than those currently found in the human YAC libraries and the larger YACs are known to be prone to more rearrangements (Anderson, 1993). In the investigations of SPARCs presented in chapters 3 onwards, a recombination deficient mutant, *rec55-36*, was used which has a 10 fold reduction in mitotic recombination between two tandemly arranged *ade6* genes (Gysler-Junker *et al.*, 1991). This may have reduced the number of rearrangements although no evidence is available. Cloning other recombination genes in *S. pombe* should allow the isolation of other recombination deficient mutants which can be incorporated in the strain used for creating the SPARC libraries. Recently, a homologue of the *S. cerevisiae* *RAD52* gene, known as *rad22* in *S. pombe* has been reported (Ostermann *et al.*, 1993). As the *RAD52* mutation has proved effective in reducing rearrangements of YACs then inclusion of a *rad22* mutation in future *S. pombe* strains may prove useful for reducing rearrangements when constructing a human library in *S. pombe* SPARCs.



Modification of SPARCs may be carried out using a process like retrofitting in YACs (Smith *et al.*, 1991; Gobin *et al.*, 1995; Taylor *et al.*, 1994). These modified vector arms include mammalian selectable markers (such as HPRT, thymidine kinase and neomycin) and mammalian telomere repeats. These modifications are useful for allowing studies on the function of the mammalian genes in their normal chromosomal context. For such investigations additional SPARC vector arms may be required to construct mammalian libraries or to allow modifications to existing SPARCs from a human library. Other *S. pombe* selectable markers may be required in order to alter the vector arms and select for the new vector arms. Selectable markers such as the *S. cerevisiae* *URA3* and *LEU2* genes are used to complement the corresponding mutations in *S. pombe* in *ura4* and *leu1* genes respectively (Russell, 1989). As shown in the last chapter the *S. cerevisiae* YACs were selected with *S. cerevisiae* *URA3* to complement an *S. pombe* strain with a deletion of the *ura4* gene. Alternatively there are several *S. pombe* selectable markers such as *ura4*, *ade6*, *leu1* and *his3* genes which could be incorporated into the *S. pombe* vectors (Russell, 1989; Waddell & Jenkin, 1995; Burke & Gould, 1994).

## 7.4 POTENTIAL OF SPARC VECTORS

The modifications described above will help to produce a more versatile SPARC system which can be adapted to suit the requirements of any investigation. This will be essential if the SPARC vector system proves to have significant advantages over other vector systems.

Other groups have already begun to plan the construction of libraries in the SPARC system. The SPARC system has been included in a proposal for use in the Malaria



Genome Project. The objectives are to sequence the genomes of the *Plasmodium* species which cause malaria therefore allowing the identification of proteins which may be targeted with drugs (Hoffmann and Adams, personal communication). *Plasmodium falciparum* and *P. yoelii* high molecular weight DNA may be cloned in the SPARC vector system. The ability to isolate the SPARCs of several hundred kilobases would prove useful since, as discussed above, SPARCs can be isolated in the absence of contamination from *S. pombe* genomic DNA. This project aims to use a combination of technology including YACs, BACs and SPARCs in order to fully sequence the entire genome. Other projects similar to this may reveal any added advantages of producing libraries in the *S. pombe* SPARC system and integrating it with existing libraries and cloning systems to produce contiguous maps of whole genomes.

Current human genome mapping investigations are isolating an increasing number of sequences derived from many genes. As sequencing of the human genome moves closer to completion, the emphasis will be to determine the function of all new genes isolated. Some mammalian genes are known to cover large genomic regions, like muscular dystrophy gene stretches over 1.4 Mb (Koenig *et al.*, 1987). Other genes have distant regulatory elements, such as the locus control regions (LCR) of  $\beta$ -globin and CD2 genes or the downstream regulatory elements of the PAX6 gene (Kawasaki *et al.*, 1995; Craddock *et al.*, 1995; Festenstein *et al.*, 1996; Plaza *et al.*, 1995; Huxley, 1994). To study genes in their chromosomal context with all regulatory elements will require large fragments of DNA to be transformed into mammalian cells. The transfer of YACs to mammalian cell lines (Brown *et al.*, 1994; Huxley *et al.*, 1991; Tyler-Smith, 1994) and into ES cells to produce transgenic mice has been demonstrated with a view to establishing methods to determine the function of a gene (Paven, *et al.*, 1990; Schedl *et al.*, 1993; Schedl *et al.*, 1996). SPARCs are easily



isolated without contamination with *S. pombe* genomic sequences therefore this should prove advantageous when transforming mammalian cells with SPARC DNA. Such sequences may affect the analyses of the SPARC insert as it has previously been demonstrated that large fragments of the *S. pombe* genome can be transferred into a mouse cell line and these *S. pombe* genomic sequences integrated into the mouse genome. In some cell lines, this DNA was able to autonomously replicate as unstable visible elements (Allshire *et al.*, 1987).

If longer stretches of mammalian sequences can be cloned in *S. pombe* than is currently possible in *S. cerevisiae* then the SPARC system may assist the progression towards the construction of a mammalian artificial chromosome (MAC). The information gained from such experiments will lead to a greater understanding of the mammalian centromeres.

The SPARC system may also be used to investigate the degree of conservation between centromeres of other fission yeast species. It may be possible to clone a functional centromere into the SPARC system from *Schizosaccharomyces octosporus* and *Schizosaccharomyces japonicus* which so far have not been studied extensively. Any centromeric elements which are conserved may be important in analysis of the fission yeast centromeres and may also allow extrapolations to the centromeres of higher eukaryotes. The modification of existing SPARC vectors by inclusion of the *ade6* marker on one of the arms would provide a direct indicator of mitotic stability and therefore a direct assay for centromeric function. Such vectors would allow sequences from any species to be tested for centromere function.



## 7.5 CONCLUSIONS

Initial analysis of the SPARC system suggests that it is feasible to construct a human library with inserts of at least several hundred kilobases. The possible advantages of the SPARC vector system have not been fully investigated although it has been demonstrated that SPARC DNA can be isolated with little or no contamination with *S. pombe* genomic DNA. This will enable functional analyses of SPARC insert to be analysed without interference from yeast sequences. Ultimately construction of the SPARC library with human inserts will demonstrate the potential of the *S. pombe* system as an alternative cloning system for large DNA molecules.



## 8 BIBLIOGRAPHY



Albertsen, H.M., Abderrahum, H., Cann, H.M., Dausset, J., Lepaslieu, D., Cohen, D. (1990): Construction and characterisation of a yeast artificial chromosome library containing seven haploid genome equivalents. *Proc. Natl. Acad. Sci. USA* **87**, 4256-4260.

Alfano, C., Davis, R.W. (1995): Use of yeast artificial chromosomes to study DNA replication and its regulation within the cell cycle. *J. Cell. Biochemistry* **S21A**, 302.

Allshire, R.C. (1995): Elements of chromosome structure and function in fission yeast. *Seminars in Cell Biology* **6**(2), 55-64.

Allshire, R.C., Nimmo, E.R., Ekwall, K., Javerzat, J.P., Cranston, G. (1995): Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes & Development* **9**(2), 218-233

Allshire, R.C., Redhead, N.J., Cranston, G. (1994): Position effect variegation at fission yeast centromeres. *Cell* **76**, 157-169.

Allshire, R.C. (1992): Manipulation of large minichromosomes in *S. pombe* with liposome enhanced transformation. *Methods in Enzymology* **216**, 614-631.

Allshire, R.C. (1990): Introduction of large linear minichromosomes into *Schizosaccharomyces pombe* by an improved transformation procedure. *Proc. Natl. Acad. Sci.* **87**, 4043-4047.

Anand, R., Villasante, A., Tyler-Smith, C. (1989): Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed field gel electrophoresis. *Nucleic Acids Res.* **17**(9), 3425-3433.

Anderson, C.A. (1993): Genome shortcut leads to problems. *Science* **259**, 1684-1687.

Aparicio, O.M., Billington, B.L., Gottschling, D.E. (1991): Modifiers of position effect are shared between telomeric and silent mating type loci in *Saccharomyces cerevisiae*. *Cell* **66**(6), 1279-1287.

Arshad, M.F., Dunn, F.J., Vega, R., Valvano, J.N., Serwer, P. (1993): Progress in developing improved programs for pulse field gel electrophoresis of DNA. *Electrophoresis* **14**, 342-348.

Arveiler, B., Porteous, D. (1992): Distribution of Alu and L1 repeat in human YAC recombinants. *Mamm. Genome* **3**, 661-668.

Ashworth, L.K., Alegriahartman, M., Burgin, M., Devlin, L., Carrano, A.V., Batzer, M.A. (1995): Assembly of high resolution bacterial artificial chromosome, P1-derived artificial chromosome, and cosmid contigs. *Analytical Biochemistry* **224**(2), 564-571.

Barbet, N., Muriel, W.J., Carr, A.M. (1992): Versatile shuttle vectors and genomic libraries for use with *Schizosaccharomyces pombe*. *Gene* **114**, 59-66.



- Baum, M., Ngan, V.K., Clarke, L. (1994): The centromeric K-type repeat and the central core are together sufficient to establish a functional *Schizosaccharomyces pombe* centromere. *Mol. Biol. Of Cell* **5**(7), 747-761.
- Beach, D., Nurse, P. (1981): High frequency transformation of fission yeast *Schizosaccharomyces pombe*. *Nature* **290**, 140-142.
- Beach, D., Piper, M., Nurse, P. (1982): Construction of a *Schizosaccharomyces pombe* gene bank in a bacterial shuttle vector and its use to isolate genes by complementation. *Mol. Gen. Genetics* **187**, 326-329.
- Bell, S.P. (1995): Eukaryote replicators and associated protein complexes. *Curr. opin. Genet. Development* **5**, 162-167.
- Blackburn, E.H. (1985): Artificial chromosomes in yeast. *Trends in Genetics* **3**, 8-12.
- Blackburn, E.H., Szostak, J.W. (1984): The molecular structure of centromeres and telomeres. *Annu. Rev. Biochem.* **53**, 163-194.
- Brewer, B.J., Fangman, W.L. (1987): The localization of replication origin on *ARS* plasmids in *S. cerevisiae*. *Cell* **51**, 437-471.
- Brown, K.E., Badnett, M.A., Burgtorf, C., Shaw, P., Buckle, G.J., Brown, W.R.A. (1994): Dissecting the centromere of the human Y chromosome with cloned telomere DNA. *Human Mol. Genetics* **3**(8), 1227-1237.
- Burke, J.D., Gould, K.L. (1994): Molecular cloning and characterisation of the *Schizosaccharomyces pombe his3* gene for use as a selectable marker. *Molecular and General Genetics* **242**(2), 169-176.
- Burke,DT, Carle,GF, Olson,MV (1987): Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**, 806-812.
- Carbon,J, Clarke,L (1990): Centromere structure and function in budding and fission yeasts. *The New Biologist* **2**(1), 10-14.
- Chartier, F.L., Keer, J.T., Sutcliffe, M.J., Henriques, D.A., Mileham, P., Brown, S.D.M. (1992): Construction of a mouse yeast artificial chromosome library in a recombination deficient strain of yeast. *Nature Genetics* **1**, 132-136.
- Chikashige,Y, Kinoshita,N, Nakaseko,Y, Matsumoto,T, Murakami,S, Niwa,O, Yanagida,M (1989): Composite Motifs and Repeat Symmetry in *S. pombe* centromeres: direct analysis by integration of *NotI* restriction sites. *Cell* **57**, 739-751.
- Cohen, D., Chumakov, I., Weissenbach, J.(1993): A 1<sup>st</sup> generation physical map of the human genome. *Nature* **366**, 698-701.
- Chumakov, I., Rigault, P., Guillou, S., Ougen, P., Billaut, A., Guasconi, G., Gervy, P., LeGall, I., Soularue, P., Grinas, L., Bougueleret, L., Bellanne-Chantelot, C., Lacroix, B., Barillot, E., Gesnouin, P., Pook, S., Vaysseix, G., Frelat, G., Schmitz, A., Sambucy, J., Bosch, A., Estivill, X., Weissenbach, J., Vignal, A., Riethman, H., Cox, D., Patterson, D., Gardiner, K., Hattori, M., Sakaki, Y., Ichikawa, H., Ohki, M., Le Paslier, D., Heilig, R., Antonarakis, S., Cohen, D. (1992): Continuum of



overlapping clones spanning the entire human chromosome 21q. *Nature* **359**, 380-387.

Church, G.M., Gilbert, W. (1984): Genomic Sequencing. *Proc. Natl. Acad. Sci. USA* **81**, 1991-1995.

Clarke, L. (1990): Centromeres of budding and fission yeasts. *Trends in Genetics* **6**(5), 150-154.

Clarke, L., Amstutz, H., Fishel, B., Carbon, J. (1986): Analysis of centromeric DNA in the fission yeast *Schizosaccharomyces pombe*. *Proc. Natl. Acad. Sci. USA* **83**, 8253-8257.

Clarke, L., Baum, M.P. (1990): Functional analysis of a centromere from fission yeast: a role for centromere-specific repeated DNA sequences. *Mol. Cell. Biol.* **10**(5), 1863-1872.

Coulson, A., Kozono, Y., Lutterbach, B., Shownkeen, R., Sulston, J., Waterston, R. (1991): YACs and the *C. elegans* genome. *Bioessays* **13**(8), 413-418.

Coulson, A., Waterson, R., Kiff, J., Sulston, J., Kohara, Y. (1988): Genome linking with yeast artificial chromosome. *Nature* **335**, 184-186.

Craddock, C.F., Vyas, P., Sharpe, J.A., Ayyub, H., Wood, W.G., Higgs, D.R. (1995): Contrasting effects of alpha-globin and beta-globin regulatory elements on chromatin structure may be related to their different chromosomal environments. *EMBO J.* **14**(8), 1718-1726.

Cranston, G., Gosden, J.R., Maule, J.C., Hastie, N.D., Fantes, P.A. (1987): A fission yeast chromosome can replicate autonomously in mouse cells. *Cell* **50**, 391-403.

Dubey, D.D., Kim, S.M., Todorov, I.T., Huberman, J.A. (1996): Large, complex modular structure of a fission yeast DNA-replication origin. *Current Biology* **6**(4), 467-473.

Egal, R., Willer, M., Nielson, O. (1989): Unblocking of meiotic crossing over between the silent mating type cassettes of fission yeast, conditioned by the recessive, pleiotropic mutant *rik1*. *Current Genetics* **15**(6), 407-410.

Ekwall, K., Nimmo, E.R., Javerat, J.P., Borgstrom, B., Egel, R., Cranston, G., Allshire, R.C. (1996): Mutations in the fission yeast silencing factors *clr4<sup>+</sup>* and *rik1<sup>+</sup>* disrupt the localization of the chromo domain protein *swi6p* and impair centromere function. *J. Cell Science.* **109**, 2637-2648.

Ekwall, K., Ruusala, T. (1994): Mutations in *rik1*, *clr2*, *clr3*, *clr4* genes asymmetrically derepress the silent mating type loci in fission yeast. *Genetics* **136**, 53-64.

Evans, G.A. (1993): MegaYAC library. *Science* **260**, 877.

Fan, J.B., Chikashige, Y., Smith, C.L., Niwa, O., Yanagida, M., Cantor, C.R. (1988): Construction of a *NotI* restriction map of the fission yeast *Schizosaccharomyces pombe* genome. *Nucleic Acids Res.* **17**, 2801-2818.



Fantes, J.A., Bickmore, W.A., Fletcher, J.M., Ballesta, F., Hanson, I.M., Van Heyningen, V. (1992): Non radioactive *in situ* hybridisation for the rapid analysis of submicroscopic deletions at the WAGR locus. *Am. J. Human Genetics*. **51**, 1286-1294.

Farr, C.J., Bayne, R.A.L., Kipling, D., Mills, W., Critcher, R., Cooke, H.J. (1995): Generation of a human X-derived minichromosome using telomere-associated chromosome fragmentation. *EMBO J*. **14**(21), 5444-5454.

Farr, C.J., Fantes, J., Goodfellow, P., Cooke, H. (1991): Functional reintroduction of human telomeres into mammalian cells. *Proc. Natl. Acad. Sci. USA* **88**, 7006-7010.

Felgner, P.L., Gadek, T.R., Holm, M., Roman, R., Chan, H.W., Wenz, M., Northrop, J.P., Ringold, G.M., Danielson, M. (1987): Lipofectin - A highly efficient lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci. USA* **84**, 7413-7417.

Felgner, P.L., Ringold, G.M. (1989): Cationic liposome-mediated transfection. *Nature* **337**, 387-388.

Festenstein, R., Tolaini, M., Corbella, P., Mamalaki, C., Parrington, J., Fox, M., Miliou, A., Jones, M., Kioussis, D. *Science* **271**, 1123-1125.

Fishel, B., Amstutz, H., Baum, M., Carbon, J., Clarke, L. (1988): Structural organisation and functional analysis of centromeric DNA in the fission yeast *Schizosaccharomyces pombe*. *Mol. Cell. Biol.* **8**(2), 754-763.

Foote, S., Vollrath, D., Hilton, A., Page, D.C. (1992): The human Y chromosome: Overlapping DNA clones spanning the euchromatic region. *Science* **258**, 60-66.

Funabiki, H., Hagan, I., Uzawa, S., Yanagida, M. (1993): Cell cycle-dependent specific positioning and clustering of centromeres and telomeres in fission yeast. *J. Cell. Biol.* **121**(5), 961-976.

Gnrke, A., Huxley, C., Peterson, K., Olson, M.V. (1993): Microinjection of intact 200 - 500 kb fragments of YAC DNA into mammalian cells. *Genomics* **15**, 659-667.

Gobin, S.J.P., Alcaideloridan, C., Bono, M.R., Ottone, C., Chumakov, I, Rothstein, R. Fellous, M. (1995): Transfer of yeast artificial chromosomes into mammalian cells and comparative study of their integrity. *Gene* **163**, 27-33.

Green, E.D., Riethman, H.C., Dutchik, J.E., Olson, M.V. (1991): Systematic generation of sequence tagged sites for physical mapping of human chromosomes - application to the mapping of the chromosome 7 using yeast artificial chromosomes. *Genomics* **11**, 658-669.

Grosveld, F., Van Assendelft, G.B., Greaves, D.R., Kollias, G. (1987): Position-independent, high-level expression of the human  $\beta$ -globin gene in transgenic mice. *Cell* **51**, 975-985.



- Gysler-Junker, A., Bodi, Z., Kohli, J. (1991): Isolation and characterization of *Schizosaccharomyces pombe* mutants affected in mitotic recombination. *Genetics* **128**, 495-504.
- Hahnenberger, K.M., Baum, M.P., Pollizzi, C.M., Carbon, J., Clarke, L. (1989): Construction of functional artificial minichromosomes in the fission yeast *Schizosaccharomyces pombe*. *Proc. Natl. Acad. Sci. USA* **86**, 577-581.
- Hahnenberger, K.M., Carbon, J., Clarke, L. (1991): Identification of DNA regions required for mitotic and meiotic function within the centromere of *Schizosaccharomyces pombe* chromosome I. *Mol. Cell. Biol.* **11**(4), 2206-2215.
- Haldi, M., Perrot, V., Saumier, M., Desai, T., Cohen, D., Cherif, D., Ward, D., Lander, E.S. (1994): Large human YACs constructed in a *RAD52* strain show a reduced rate of chimerism. *Genomics* **24**(3), 478-484.
- Hanish, J.P., Yanowitz, J.L., DeLange, T. (1994): Stringent sequence requirements for the formation of human telomeres. *Proc. Natl. Acad. Sci. USA* **91**(19), 8861-8865.
- Heller, R., Brown, K.E., Burgtorf, C., Brown, W.R.A. (1996): Minichromosomes derived from the human Y-chromosome by telomere directed chromosome breakage. *Proc. Natl. Acad. Sci. USA* **93**(14), 7125-7130.
- Hermanston, G.G., Lichter, P., Selleri, L., Ward, D.C., Evans, G.A. (1992): Cosmid linking clones localised to the long arm of human chromosome 11. *Genomics* **13**, 134-143.
- Holm, C., Meekswagner, D.W., Fangman, W.L., Botstein, D. (1986): A rapid efficient method for isolating DNA from yeast. *Gene* **42**(2), 169-173.
- Hultman, T., Bergh, S., Moks, T., Uhlen, M. (1991): Bidirectional solid phase sequencing of *in vitro* amplified plasmid DNA. *Biotechniques* **10**, 84-93.
- Huysmans, E., Dewachter, R. (1983): Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Research* **14**(S), R73-R118.
- Huxley, C. (1994): Mammalian artificial chromosomes: a new tool for gene therapy. *Gene Therapy* **1**, 7-12.
- Huxley, C., Gnirke, A. (1991): Transfer of yeast artificial chromosomes from yeast to mammalian cells. *Bioessays* **13**(10), 545-550.
- Huxley, C., Hagino, Y., Schlessinger, D., Olson, M.V. (1991): The human HPRT gene on a yeast artificial chromosome is functional when transferred to mouse cells by cell fusion. *Genomics* **9**, 742-750.
- Ioannou, P.A., Amemiya, C.T., Garnes, J., Kroisel, P.M., Shizuya, H., Chen, C., Batzer, M.A., de Jong, P.J. (1994): A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genetics* **6**, 84-89.



- Kaufers, N.F., Simanis, V., Nurse, P. (1985): Fission yeast *Schizosaccharomyces pombe* correctly excises a mammalian RNA transcript intervening sequence. *Nature* **318**, 78-80.
- Kawasaki, K., Minoshima, S., Schooler, K., Kudoh, J., Asakawa, S., DeJong, P.J., Shimizu, N. (1995): The organisation of the human immunoglobulin lambda gene locus. *Genome Research* **5**(2), 125-135.
- Kerem, B.S., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M., Tsui, L.C. (1989): Identification of the cystic fibrosis gene - genetic analysis. *Science* **245**, 1073-1080.
- Kim, U.J., Shizula, H., Chen, X.N., Deaven, L., Speicher, S., Solomon, J., Korenberg, J., Simon, M.I. (1995): Characterisation of a human chromosome 22 enriched bacterial artificial chromosome sublibrary. *Genetic analysis - Biomolecular engineering* **12** (2), 73-79.
- Kim, U.J., Shizuya, H., DeJong, P.J., Birren, B., Simon, M.I. (1992): Stable propagation of cosmid sized human DNA inserts in an F-factor based vector. *Nucleic Acids Research* **20**(5), 1083-1085.
- Kipling, D. (1995): *The telomere*. Oxford University Press.
- Klein, H.L. (1988): Different types of recombination events are controlled the *RAD1* and *RAD52* genes of *Saccharomyces cerevisiae*. *Genetics* **120**, 367-377.
- Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C., Kunkel, L.M. (1987): Complete cloning of the Duchenne Muscular Dystrophy (DMD) cDNA and preliminary genomic organisation of the DMD gene in normal and affected individuals. *Cell* **50**, 509-517.
- Kohli, J., Hottinger, H., Munz, P., Strauss, A., Theriaux, P. (1977): Genetic mapping in *Schizosaccharomyces pombe* by mitotic and meiotic analyses and induced haploidization. *Genetics* **87**, 471-489.
- Kouprina, N., Eldarov, M., Moyzis, R., Resnick, M., Larionov, V. (1994): A model system to assess the integrity of mammalian YACs during transformation and propagation in yeast. *Genomics* **21** (1), 7-17.
- Kuhn, R.M., Clarke, L., Carbon, J. (1991): Clustered tRNA genes in *Schizosaccharomyces pombe* centromeric DNA sequence repeats. *Proc. Natl. Acad. Sci. USA* **88**, 1306-1310.
- Larin, Z., Monaco, A.P., Lehrach, H. (1992): Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. *Proc. Natl. Acad. Sci. USA* **88**, 4123-4127.
- Larionov, V., Kouprina, N., Nikolaishvili, N., Resnick, M.A. (1994): Recombination during transformation as a source of chimeric mammalian artificial chromosomes in yeast (YACs). *Nucleic Acid Research* **22** (20), 4154-4162.
- Laurenson, P., Rine, J. (1992): Silencers, silencing and heritable transcriptional states. *Microbiological Reviews* **56**(4), 543-560.



- Leach, D.R.F., Stahl, F.W. (1983): Viability of the  $\lambda$  phages carrying a perfect palindrome in absence of recombination nuclease. *Nature* **305**, 448-451.
- Ling, L.L., Ma, N.S.F., Smith, D.R., Miller, D.D., Moir, D.T. (1993): Reduced occurrence of chimeric YACs in recombination deficient hosts. *Nucliec Acids Research* **21**(25), 6045-6046.
- Lohe, A.R., Hilliker, A.J. (1995): Return of the H-word (heterochromatin). *Current Opinion in Genetics and Development* **5**(6), 746-755.
- Losson, R., Lacroute, F. (1983): Plasmids carrying the yeast OMP decarboxylase structural and regulatory genes: transcription regulation in a foreign enviroment. *Cell* **32**, 371-377.
- Louis, E.J. (1995): The chromosome ends of *Saccharomyces cerevisiae*. *Yeast* **11**(16), 1553-1573.
- Marshall, L.G., Clarke, L. (1995): A novel cis-acting centromeric DNA element affects *S. pombe* centromeric chromatin structure at a distance. *J. Cell. Biol.* **128**(4), 445-454.
- Matsumoto, T., Fukui, K., Niwa, O., Sugawara, N., Szostak, J.W., Yanagida, M. (1987): Indentification of healed terminal DNA fragments in linear minichromosomes of *Schizosaccharomyces pombe*. *Mol. Cell. Biol.* **7**(12), 4424-4430.
- Matsumoto, T., Murakami, S., Niwa, O., Yanagida, M. (1990): Construction and characterisation of centric circular and acentric linear chromosomes in fission yeast. *Curr. Genetics* **18**, 323-330.
- Maundrell, K., Hutchinson, A., Shall, S. (1988): Sequence analysis of ARS elements in fission yeast. *EMBO J.* **7**, 2203-2207.
- McCormick, M.K., Shero, J.H., Connelly, C.J., Antonarakis, S.E., Hieter, P. (1990): Methods for cloning large DNA segments as artificial chromosomes in *S. cerevisiae*. *Technique-A Journal of Methods in Cell and Molecular Biology* **2**(2), 65-71.
- Mizakami, T., Chang, W.I., Garkavtsev, I., Kaplan, N., Lombardi, D., Matsumoto, T., Niwa, O., Kounosu, A., Yanagida, M., Marr, T.G., Beach, D. (1993): A 13 kb resolution cosmid map of the 14 kb fission yeast genome by non random sequence tagged site mapping. *Cell* **73**, 121-132.
- Monaco, A.P., Larin, Z. (1994): YACs, BACs, PACs, and MACs: artificial chromosomes as research tools. *Trends Biotechnol.* **12**, 280-286.
- Moreno, S., Klar, A., Nurse, P.(1991): Molecular analysis of fission yeast *Schizosaccharomyces pombe*. *Methods Enzymol.* **194**, 795-823.
- Mortimer, R.K., Schild, D., Contopoulou, C.R., Kans, J.A. (1989): Genetic map of *Saccharomyces cerevisiae* edition-10. *Yeast* **5**(5), 321



- Murakani, S., Matsumoto, T., Niwa, O., Yanagida, M. (1991): Structure of the fission yeast centromere *cen3*: direct analysis of the reiterated inverted region. *Chromosoma* **101**, 214-221.
- Murray, A.W., Szostak, J.W. (1983): Construction of artificial chromosomes in yeast. *Nature* **305**, 189-193.
- Nakaseko, Y., Adachi, Y., Funahashi, S., Niwa, O., Yanagida, M. (1986): Chromosome walking shows a highly homologous repetitive sequence present in all the centromere regions of fission yeast. *EMBO J.* **5**(5), 1011-1021.
- Neil, O.L., Villasante, A., Fisher, R.B., Vetrie, D., Cox, B., Tyler-Smith, C. (1990): Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome vectors. *Nucleic Acids Res.* **18**(6), 1421-1428.
- Nimmo, E.R., Cranston, G., Allshire, R.C. (1994): Telomere-associated chromosome breakage in fission yeast results in variegated expression of adjacent genes. *EMBO J.* **13**, 3801-3811.
- Niwa, O., Matsumoto, T., Chikashige, Y., Yanagida, M. (1989): Characterization of *Schizosaccharomyces pombe* minichromosome deletion derivatives and a functional allocation of their centromere. *EMBO J.* **8**(10), 3045-3052.
- Niwa, O., Matsumoto, T., Yanagida, M. (1986): Construction of a mini-chromosome by deletion and its mitotic and meiotic behaviour in fission yeast. *Mol. Gen. Genet.* **203**, 395-405.
- Ostermann, K., Lorentz, A., Schmidt, H. (1993): The fission yeast *rad22* gene, having a function in mating type switching and repair of DNA damages, encodes a protein homolog to *RAD52* of *Saccharomyces cerevisiae*. *Nucleic Acids Research* **21**(25), 5940-5944.
- Pavan, W.J., Hieter, P., Reeves, R.H. (1990): Modification and transfer into an embryonic carcinoma cell line of a 360-kilobase human-derived yeast artificial chromosome. *Mol. Cell. Biol.* **10**(8), 4163-4169.
- Pierce, J.C., Sternberg, N., Sauer, B. (1992): A mouse genomic library in the bacteriophage P1 cloning system: organisation and characterization. *Mamm. Genome* **3**, 550-558.
- Plaza, S., Turque, N., Dozier, C., Bailly, M., Saule, S. (1995): C-myb acts as transcriptional activator of the Quail Pax6 (Pax-QNR) promoter through 2 different mechanisms. *Oncogene* **10**, 329-340.
- Pluta, A.F., MacKay, A.M., Ainsztein, A.M., Goldberg, I.G., Earnshaw, W.C. (1995): Centromere - Hub of chromosomal activities. *Science* **270**, 1591-1594.
- Rattner, J.B. (1991): The structure of mammalian centromere. *Bioessays* **13**(12), 51-56.



Riordan, J.R., Rommens, J.M., Kerem, B.S., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.L., Drumm, M.L., Iannuzzi, M.C., Collins, F.S., Tsui, L.C. (1989): Identification of the cystic fibrosis gene - cloning and characterisation of complementary DNA. *Science* **245**, 1066-1072.

Robinow, C.F. (1977): The number of chromosomes in *Schizosaccharomyces pombe*: light microscopy of stained preparations. *Genetics* **87**, 491-497.

Rommens, J.M., Iannuzzi, M.C., Kerem, B.S., Drumm, M.L., Melmer, G., Dean, M., Rozmahel, R., Cole, J.L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J.R., Tsui, L.C., Collins, F.S. (1989): Identification of the cystic fibrosis gene - chromosome walking and jumping. *Science* **245**, 1066-1072.

Russell, P. (1989): Molecular biology of the fission yeast. Academic press. Inc.,(Chap 7)

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989): Molecular cloning: A laboratory Manual. 2nd ed. 3 vols. Cold Spring Harbour Laboratory Press, New York.

Schedl, A., Larin, Z., Montoliu, L., Thies, E., Kelsey, G., Lehrach, H., Schutz, G. (1993): A method for the generation of YAC transgenic mice by pronuclear microinjection. *Nucleic Acids Research* **21**(20), 4783-4787.

Schedl, A., Ross, A., Lee, M., Engelkamp, D., Rashbass, P., Van Heyningen, V., Hastie, N.D. (1996): Influence of *pax6* gene dosage on development - overexpression causes severe eye abnormalities. *Cell* **86**, 71-82.

Schmitt, H., Kim, U.J., Slepak, T., Blin, N., Simon, M.I., Shizuya, H. (1996): Framework for a physical map of the human 22q13 region using Bacterial artificial chromosomes (BACs). *Genomics* **33** (1), 9-20.

Schulman, I., Bloom, K.S. (1991): Centromeres: an integrated protein/DNA complex required for chromosome movement. *Annu. Rev. Cell Biol.* **7**, 311-336.

Schwartz, D.C., Cantor, C.R. (1984): Separation of yeast chromosome sized DNA's by pulsed field gradient gel electrophoresis. *Cell* **37**, 67-75.

Shibasaki, Y., Maule, J.C., Devon, R.S., Slorach, E.M., Gosden, J.R., Porteous, D.J., Brookes, A.J. (1995): Catch-linker plus PCR labelling - A simple method to generate Fluorescence *in situ* hybridisation probes from yeast artificial chromosomes. *PCR Methods and Applications* **4**, 209-211.

Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiri, Y., Simon, M.(1992): Cloning and stable maintenance of 300 kilobase pair fragments of human DNA in *Escherichia coli* using an F-factor based vector. *Proc. Natl. Acad. Sci. USA* **89**(18), 8794-8797.

Singer, M.F. (1982): Highly repeated sequences in mammalian genomes. *Int. Rev. Cytol.* **76**, 67-112.

Smith, D.R., Smyth, A.P., Moir, D.T. (1990): Amplification of large artificial chromosomes. *Proc. Natl. Acad. Sci. USA* **87**, 8242-8246.



- Smoller, D.A., Petrov, D., Hartl, D.L. (1991): Characterization of a bacteriophage P1 library containing inserts of *Drosophila* DNA of 75-100 kilobase pairs. *Chromosoma* **100**, 487-494.
- Steiner, N.C., Hahnenberger, K.M., Clarke, L. (1993): Centromeres of the fission yeast *Schizosaccharomyces pombe* are highly variable genetic loci. *Mol. Cell. Biol.* **13**(8), 4578-4587.
- Sternberg, N. (1990): Bacteriophage P1 cloning system for the isolation, amplification and recovery of DNA fragments as large as 100Kb pairs. *Proc. Natl. Acad. Sci. USA* **87**, 103-107.
- Stinchcomb, D.T., Struhl, K., Davis, R.W. (1979): Isolation and characterisation of a yeast chromosomal replicator. *Nature* **282**, 39-43.
- Sugawara, N. (1989): PhD thesis, Harvard University.
- Szostak, J.W., Blackburn, E.H. (1982): Cloning yeast telomeres on linear plasmid vectors. *Cell* **29**: 245-255.
- Szostak, J.W., Sugawara, N. (1986): Telomeres in *Schizosaccharomyces pombe*. *Yeast* **2**(Suppl.), 373. (Abstract)
- Takahashi, K., Murakami, S., Chikashige, Y., Funabiki, H., Niwa, O., Yanagida, M. (1992): A low copy number central sequence with strict symmetry and unusual chromatin is essential for fission yeast centromere. *Mol. Biol. Cell* **3**: 819-835.
- Takahashi, K., Murakami, S., Chikashige, Y., Niwa, O., Yanagida, M. (1991): A large number of tRNA genes are symmetrically located in fission yeast centromeres. *J Mol. Biol.* **218**, 13-17.
- Taylor, S. S., Larin, Z., Smith, C.T. (1994): Addition of functional human telomeres to YACs. *Human Molecular Genetics* **3** (8), 1383-1386.
- Thon, G., Klar, A.J.S. (1994): The *clr1* locus regulates the expression of the sryptic mating-type loci of fission yeast. *Genetics* **131**, 287-296.
- Tyler-Smith, C., Oakey, R.T., Larin, Z., Fisher, R.B., Crocker, M., Affera, N.A., Ferguson-Smith, M.A., Muenke, M., Zuffardi, O., Jobley, M.A. (1993): Localisation of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes. *Nature Genetics* **5**, 368-375.
- Uzawa, S., Yanagida, M. (1992): Visualisation of centromeric and nucleolar DNA in fission yeast by fluorescence *in situ* hybridisation. *J. Cell Science* **101**, 267-275.
- Vogt, P. (1992): Code domains in tandem repetitive DNA sequence structures. *Chromosoma* **101**, 585-589.
- Wada, M., Abe, K., Okumura, K., Taguchi, H., Kohno, K., Imamoto, F., Schlessinger, D., Kuwano, M. (1994): Chimeric YACs were generated at unreduced rates in conditions that suppress coligation. *Nucleic Acids Research* **22**(9), 1651-1654.



- Waddell, S. and Jenkins, J.F. (1995): *Arg*<sup>3+</sup>, A new selection marker system for *Schizosaccharomyces pombe* - application of *ura*<sup>4+</sup> as a removable integration marker. *Nucleic Acids Research* **23**(10), 1836-1837.
- Wang, M., Shouse, S., Lipes, B., Kim, U.J., Shizuya, H., Lai, E. (1996): Bacterial artificial chromosome cloning and mapping of a 630 kb human extrachromosomal structure. *Genome Research* **6** (7), 612-619.
- Wang, G.L., Holsten, T.E., Song, W.Y., Wang, H.P., Ronald, P.C. (1995): Construction of a rice bacterial artificial chromosome library and identification of clones linked to the XA-21 disease resistance locus. *Plant Journal* **7**(3), 525-533.
- Wang, M., Chen, X.N., Shouse, S., Manson, J., Wu, Q.Z., Li, R.M., Wrestler, J., Noya, D., Sun, Z.G., Korenberg, J., Lai, E. (1994): Construction and characterisation of a human chromosome 2 specific BAC library. *Genomics* **24**(3), 527-534.
- Willard, H.F. (1990): Centromeres of mammalian chromosomes. *Trends in Genetics* **6**(12), 410-416.
- Wyman, A.R., Wolfe, L.B., Botstein, D. (1985): Propagation of some human DNA sequences in bacteriophage  $\lambda$  vectors requires mutant *Escherichia coli* hosts. *Proc. Natl. Acad. Sci. USA* **82**, 2880-2884.
- Yokobata, K., Trenchak, B., DeJong, P.J.(1991): Rescue of unstable cosmids by *in vitro* packaging. *Nucleic Acids Research* **19**(2), 403-404.