ISOLATION AND CHARACTERISATION OF HUMAN TELOMERES

SALLY H. CROSS

PhD
University of Edinburgh
1989

DECLARATION

I declare

(a) that this thesis has been composed by myself, and

(b) that the work is my own except where stated.

ABSTRACT

The telomere is the structure found at the end of a linear chromosome. It is essential for the stability and for the complete replication of the chromosome end. Telomeres can associate transiently with each other and with the nuclear envelope and these associations may be important for chromosome segregation or in the spatial organisation of the chromosomes in the nucleus. Eukaryotic telomeres have been well-studied in species such as ciliates and yeast and a number of common features have emerged. All telomeres studied so far consist of an array of short tandem repeats which have a G-rich and a C-rich strand. The G-rich strand runs 5'-3' towards the terminus and this strand extends at the end.

That human telomeres shared these properties was suggested by analysis with a molecular probe from close to one human telomere which showed that variable numbers of tandem repeats were likely to be present at the end and that the terminus was a 3'overhang. Sperm chromosome telomeres appear to have an additional 10kb of the terminal repeats when compared to blood chromosome telomeres. Cross-hybridization experiments described here show that the sequence of the terminal repeats is closely related to the simple repeat found at Trypanosoma telomeres, TTAGGG. As this sequence would be expected to provide telomere function in yeast, a yeast artificial chromosome vector, pYAC4NEO, was modified to provide a specialised vector designed to clone human telomeres and flanking sequences by complementation of telomere function in yeast. When human DNA enriched for telomeres was ligated to linearised vector short linear molecules of the predicted structure were obtained.

Analysis of the human sequences from the end of the new chromosomes showed that an array of terminal repeats of the expected sequence was found. Pro-terminal sequences were isolated and shown to detect Bal-31 sensitive restriction fragments in human DNA in addition to other fragments suggesting that these sequences must be pro-terminal on at least one human chromosome end. Further analysis showed that these pro-terminal sequences are shared between different chromosome ends. When the hybridization patterns with these pro-terminal sequences of blood and sperm DNA digests from the same individual were compared differences were found with several enzymes, some of which are accounted for by differences in methylation between blood and sperm DNA.

The terminal repeats, which are found at all telomeres, are probably responsible for the stability and complete replication of the chromosome end and the pro-terminal sequences may be responsible for the various telomere interactions that have been observed.

ABBREVIATIONS

A	Adenine
Amp	Ampicillin
ARS	Autonomously replicating sequence
ATP	Adenosine triphosphate
BAP	Bacterial alkaline phosphatase
bp	base pair
BSA	Bovine serum albumin
C	Cytosine
CMGT	Chromosome mediated gene transfer
cpm	Counts per minute
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanidine triphosphate
dTTP	Deoxythymidine triphosphate
ddNTP	Dideoxynucleotide triphosphate
dNTP	Deoxynucleotide triphoshpate
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
ds	Double stranded
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FACS	Fluorescence-activated cell sorter
FCS	Fetal calf serum
G	Guanine
HAT	Hypoxanthine aminopterin thymidine
HPRT	Hypoxanthine guanine phosphoribosyl transferase

hrs	hours
IPTG	Isopropyl-β-D-thio-galactopyranoside
K	1000 revolutions per minute
kb	Kilobase pairs
LGT	Low gelling temperature
Mb	Megabase pairs
mins	Minutes
nt	Nucleotide
OD	Optical density
p	Short chromosome arm
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
PFG	Pulsed field gel
q	Long chromosome arm
RNA	Ribonucleic acid
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
ss	Single stranded
SSC	Sodium saline citrate
T	Thymidine
TE	10mM Tris pH 7.5, 1mM EDTA pH 8.0
TEMED	N,N,N',N'-tetramethylethylenediamine
TPB	Trypto phosphate broth
Tris	Tris (hydroxymethyl) aminomethan
Tet-4	oligonucleotide $(TTGGGG)_4$
Try-4	oligonucleotide $(TTAGGG)_4$
UV	Ultraviolet
vol	Volume

w/v	Weight/volume
X-Gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside
XP1	XhoI - PstI fragment of pHutel1
XP22	XhoI - PstI fragment of pHutel22
YAC	Yeast artificial chromosome

ACKNOWLEDGEMENTS

I owe a large number of people thanks for their help, advice and encouragement during my PhD, without whom this thesis would not have been written.

First of all many thanks to Howard who devised the project in the first place and was always there to supervise and guide me when I needed it. Special thanks to Robin Allshire who made the crucial observation that the Tetrahymena telomeric repeat cross-hybridized to human telomeres and was always around to have long conversations with about TTAGGG and TTGGGG. The advice (and reagents) provided by Veronica van Heyningen and John Ansell were largely responsible for the success of the cell fusion experiments, the teaching of Stewart in the ways of tissue culture essential. Thanks also to Harris Morrison for FACS analysis, Judy Fantes for doing the in situs and John Inglis for making the oligos. The members of the MGU/Structural Studies Section, Nio, Stewart, Mary, Donald and Janet all provided much help and encouragement in many different ways. Thanks also to Nick Hastie for the lively interest he took in my project and the suggestions he made.

This thesis would definately not have been produced without the expertise of Sandy, Norman and Douglas in Photography whose considerable skill transformed my scruffy auto-rads into presentable data. Thanks to Sheila for running the best library I have ever used and to Helen for photocopying vast numbers of papers for me. Thanks to Professor Evans for encouragement and support during my time at HGU and for finding travel money. I thank the SERC for financial support during my studentship.

Other people I have had the good fortune to meet here contributed the good time I have had in Edinburgh. The members of the RP group especially Helen and Sylvia for innovative Friday Clubs, the Irish Trip flozzies: Paula, Sinéad, Anne, Wendy and Marie-Pierre, (joint motto - I'd rather be in Fermanagh), Kathy and Bruno, John, Val and the kids, Robin, Richard (undoubtly the best squash player in the world), John Mc, Niki (who will always remember to put her career first), Mark who introduced me to Farmer Jack's and Mrs Junis spicy naan bread, and to all the other people at HGU who took an interest in my work.

The person I owe most to is Ian who precipitated my move to Edinburgh and encouraged me to do a PhD. Unfortunately for him he then had to live with a manic, broke student for the next three years and had the unenviable task of reading my thesis for mistakes. Thank you for everything. I also want to thank Peter Little who gave me my first job in molecular biology and who taught me how to do it, showing great patience while doing so and to Keith Dudley who first encouraged me to do a PhD.

This thesis is dedicated to my father, a truly wonderful man whose best piece of advice to me was to to do what I wanted to do, not necessarily what seemed the most sensible.

You can get it if you really want,

but you must try, try and try, try and try.

Jimmy Cliff

The Harder They Come

CONTENTS

CHAPTER 1

INTRODUCTION

1 INTRODUCTION

Telomeres are defined cytologically as the structure found at the end of a linear chromosome. This structure confers stability on the chromosome and allows the complete replication of the chromosome end. In terms of the DNA there must be specific sequences which fulfil these functions.

Telomeres maintain chromosomes intact by preventing them degrading and undergoing fusion events. Ends of broken chromosomes and molecules terminating in non-telomeric sequences are recombinogenic, can fuse with each other and are subject to degradation (McClintock, 1941, 1942). As DNA polymerization only proceeds from a primer in a 5'-3' direction there must be a specialised mechanism which acts at telomeres to ensure that ends are completely replicated otherwise material would continually be lost from the ends and chromosomes would become progressively shorter. This would be a lethal event either because of loss of the telomeric sequences which cap the chromosome leading to chromosomal instability or because essential genes are deleted. In addition to these functions various telomere-telomere associations have been observed. In this Introduction I will describe what is known about telomeres and how they fulfil their role in chromosome stability, structure and replication.

1.1 Cytology of Telomeres

Telomeric regions are often, but not always, heterochromatic in appearance and occasionally characteristic telomeric "knobs" are observed, for example in maize (McClintock, 1941). Various telomere-telomere associations have been observed at both mitosis and meiosis (for example Moses, 1977). At leptotene, the first meiotic stage at which chromosomes are visible cytologically, the telomeres are attached to the nuclear membrane. The telomeres then move towards each other while still attached to the nuclear membrane and the chromosomes take up a characteristic "bouquet" conformation largely as a consequence of the telomere grouping (Moens, 1987). Pairing and synapsis of chromosomes at zygotene usually starts at the telomeres (Chandley, 1986). At mitosis telomeres are paired or associated to some degree and association of telomeres with each other and the nuclear membrane during interphase has been observed (Moens, 1987). Studies of intact Drosophila salivary gland nuclei have revealed that telomeres are associated with the nuclear membrane and with each other, congregating around the opposite nuclear pole to the chromocentre (Agard and Sedat, 1983; Mathog et al., 1984). Vimentin, an intermediate filament protein, can bind to oligonucleotides corresponding to telomeric DNA sequences (Shoeman et al., 1988). As intermediate filaments are found at the nuclear periphery they may mediate associations between telomeres and the nuclear envelope.

Occasionally telomeres of acrocentric chromosomes fuse and form a new stable metacentric chromosome, a "Robertsonian" fusion (White, 1978). These fusions are usually stable and can account for karyotype differences between related species. For example human chromosome 2 is

thought to have been formed by fusion of two acrocenrtric ape chromosomes by telomere fusion (Lejeune *et al.*, 1973; Yunis and Prakash, 1982).

1.2 Telomere Sequences

Telomeres have been extensively studied in yeast and in those organisms such as *Tetrahymena* which fragment their genome during macronuclear development to generate a large number of small linear chromosomes and therefore telomeres. The molecular structure of the telomeric region found in these organisms is similar. In general very simple tandem repeats, usually only a few hundred base pairs long, are found at the extreme ends. Sub-telomeric regions include arrays of complex repeats which may extend for many kilobases. These sub-telomeric repeats may be responsible for telomere associations and the evidence suggests that the terminal repeats confer stability on the chromosome and allow for the complete replication of the end.

Sequences found at the telomeres of a number of species are shown in Table 1.1. The telomeric repeats share many common features. The sequence composition of the repeats is strikingly similar in that the G's and C's are segregated onto different strands. The G-strand always runs 5'-3' towards the telomere and this strand extends at the end (Henderson and Blackburn, 1989). This terminal 3' extension can be either single-stranded as in *Oxytricha* where there is a 16 base overhang of sequence $T_4G_4T_4G_4$ (Pluta *et al.*, 1982), or a gapped hairpin as in *Tetrahymena*, as judged by the inaccessibility of the terminus to T4 DNA polymerase end labelling, (Blackburn and Gall, 1978).

TABLE 1.1 Telomeric Repeat Sequences

Organism	5'-3' Sequence[1]	Reference
Tetrahymena	TTGGGG	Blackburn and Gall (1978)
Oxytricha Euplotes	TTTTGGGG	Klobutcher et al. (1981)
Paramecium	TT(T/G)GGG	Baroin et al. (1987)
Physarum Didymium Trypanosoma	TTAGGG	Johnson (1980) Forney et al. (1987) Van der Ploeg et al. (1984)
Dictyostelium	AG_{1-8}	Emery and Weiner (1981)
Plasmodium	TT(T/C)AGGG	Ponzi et al. (1985)
Arabidopsis	TTTAG(G/A)G	Richards and Ausubel (1988)
S. cerevisiae	TG_{1-3}	Shampay et al. (1984)
S. pombe	$T_{1-2}ACA_{0-1}C_{0-1}G_{1-6}$	Sugawara and Szostak (1986)

Notes

1. Tandem arrays of these repeats are found at the telomeres of these organisms. The strand shown runs 5' - 3' towards the telomere.

In some organisms the number of repeats on a chromosome varies from cell to cell so that terminal restriction fragments are heterogenous in length (Blackburn and Szostak, 1982). In the ciliate Oxytricha there are differences between the micronuclear and macronuclear telomeres in the number of telomeric repeats present. The micronuclear telomeres are heterogeneous consisting of 3 - 6 kb of the telomeric repeat (Dawson and Herrick, 1984; Prescott, pers. comm.), whereas the macronuclear telomeres all have the same structure consisting of a duplex region of 2.5 telomeric repeats and a 16 nucleotide 3' extension (Pluta et al., 1982). In some species copies of the telomeric repeat sequence are found elsewhere in the genome, for example S. cerevisiae (Walmsley et al., 1984).

Often species-specific non-ligatable single-strand breaks are found in the distal portion of the tandem array. In Tetrahymena the gaps are one nucleotide long and are found on the CCCCAA strand after the second A of the repeating unit every two to four repeats, although at least one gap has been found on the TTGGGG strand just internal to the C-rich strand gaps (Blackburn and Gall,1978). Other organisms in which such gaps have been found are yeast (Szostak and Blackburn, 1982), Physarum (Johnson, 1980), Plasmodium (Ponzi et al., 1985) and Trypanosoma (Blackburn and Challoner, 1986). In all these organisms the gaps are principally found on the C-rich strand close to the end, as in Tetrahymena.

More complex sub-telomeric repeats are often found. In S.cerevisiae two repeat sequences are the Y' and X sequences (Chan and Tye, 1983a). The Y' sequence is a highly conserved, 6.7 kb element found repeated up to four times adjacent to the telomeric repeats. The X sequence (0.3 - 3.75 kb) is a less well conserved element found on

the centromere side of the Y' elements (Chan and Tye, 1983b). Tracts of TG_{1-3} are sometimes found interspersed between these elements (Walmsley et al., 1984). As the distribution of these sub-telomeric repeats is highly strain dependent and they are found on many but not all yeast chromosomes they are unlikely to be essential for telomere function (Zakian and Blanton, 1988; Olson, pers. comm.). A deleted version of yeast chromosome III lacking both X and Y' elements appears to be as mitotically stable as a normal chromosome III (Murray and Szostak, 1986) adding to the evidence that these elements are not essential.

Complex tandem repeats at or near telomeres have been found in other species. In Drosphila melanogaster there are several. A 3 kb repeat (Rubin, 1978) hybridizes to all telomeric regions and a 12 kb repeat is found in pericentric heterochromatin in addition to telomeres (Young et al.,1983).

Although none of these repeats have been shown to be essential for telomere function their presence may be important in the healing of telomeres after loss of telomere sequences. Linear plasmids terminating in Y' elements introduced into yeast cells can acquire yeast telomeres and telomere-adjacent Y' elements through RAD52 dependent recombination (Dunn et al., 1984). This demonstation of transfer of Y' and telomeric sequences to plasmids from chromosomes by homologous recombination may reflect a mechanism by which broken chromosomes are healed in yeast. One method by which dicentic chromosomes in yeast, produced artifically in this case, are stabilised is by chromosome breakage followed by healing of the ends by telomere addition (Jäger and Philippsen, 1989). This is a rare event and is probably mediated by Y' element homologous recombination.

Telomeric regions are known to be highly polymorphic, recombinogenic and to undergo frequent rearrangements (Horowitz et al., 1984), so it is likely that recombination plays a part in telomere healing.

Both the Y' and X elements contain replication origins (Chan and Tye, 1983b) and there is evidence that suggests that one site of yeast chromosomal initiation is near telomeres (Newlon and Burke, 1980). Recently two-dimensional agarose gel electrophoresis techniques have been developed which can identify restriction fragments that contain active replication origins (Brewer and Fangman, 1987). Analysis using this technique suggests that the ARS present in the Y' element is active (Brewer, pers. comm.). As telomeric regions in yeast replicate late in S phase (McCarroll and Fangman, 1988) the replication origins present in the X and Y' elements would have to be coordinately activated late in S phase. Experiments have shown that it is chromosomal position which is an important factor in determining replication origin activation time, rather than any intrinsic property of an origin itself (Ferguson and Fangman, pers. comm.). Proximity to a chromosomal end confers late activation on adjacent origins of replication. Interestingly telomeric repeats alone, in the absence of a free end, have only minimal effect on replication time. The property of free ends which is responsible for this effect is not clear at present. One possible explanation is that association of telomeres with the nuclear membrane makes the adjacent replication origins inaccessible to the replication machinery until late in S phase.

1.3 Structural Properties of Telomeres

The structural properties of the DNA sequences found at telomeres have been examined and some interesting observations have emerged from these studies.

Recently Henderson *et al.* (1987) showed that single-stranded DNA oligonucleotides corresponding in sequence to the G-rich strand of the telomeres from several species can form novel intramolecular structures containing guanine.guanine base pairs *in vitro*. That the telomeric sequences tested could form these structures may indicate that this G.G base-pairing is crucial for telomere function. Guanine-rich single-stranded DNA can self-associate to form four-stranded structures in which the strands run in parallel fashion (Sen and Gilbert, 1988). During meiosis this self-association may be important to bring together and to zip up in register the four homologous chromatids. Purified macronuclear telomeres of *Oxytricha* can cohere to form large aggregates, albeit slowly and in the presence of high concentrations of sodium ions (Oka and Thomas, 1987).

The rate of digestion of *Tetrahymena* and yeast telomeres by the exonuclease Bal-31 decreases once the terminal repeats have been degraded (Henderson *et al.*, 1988; Shampay *et al.*, 1984). This result is surprising as telomere-adjacent sequences are A-T rich in both species (Spangler *et al.*, 1988; Shampay *et al.*, 1984) and Bal-31 usually degrades A-T rich DNA faster than G-C rich DNA. Duplex DNA consisting of *Tetrahymena* telomeric repeats is hypersensitive to cleavage by the nuclease S1 when torsionally stressed (Burdarf and Blackburn, 1987).

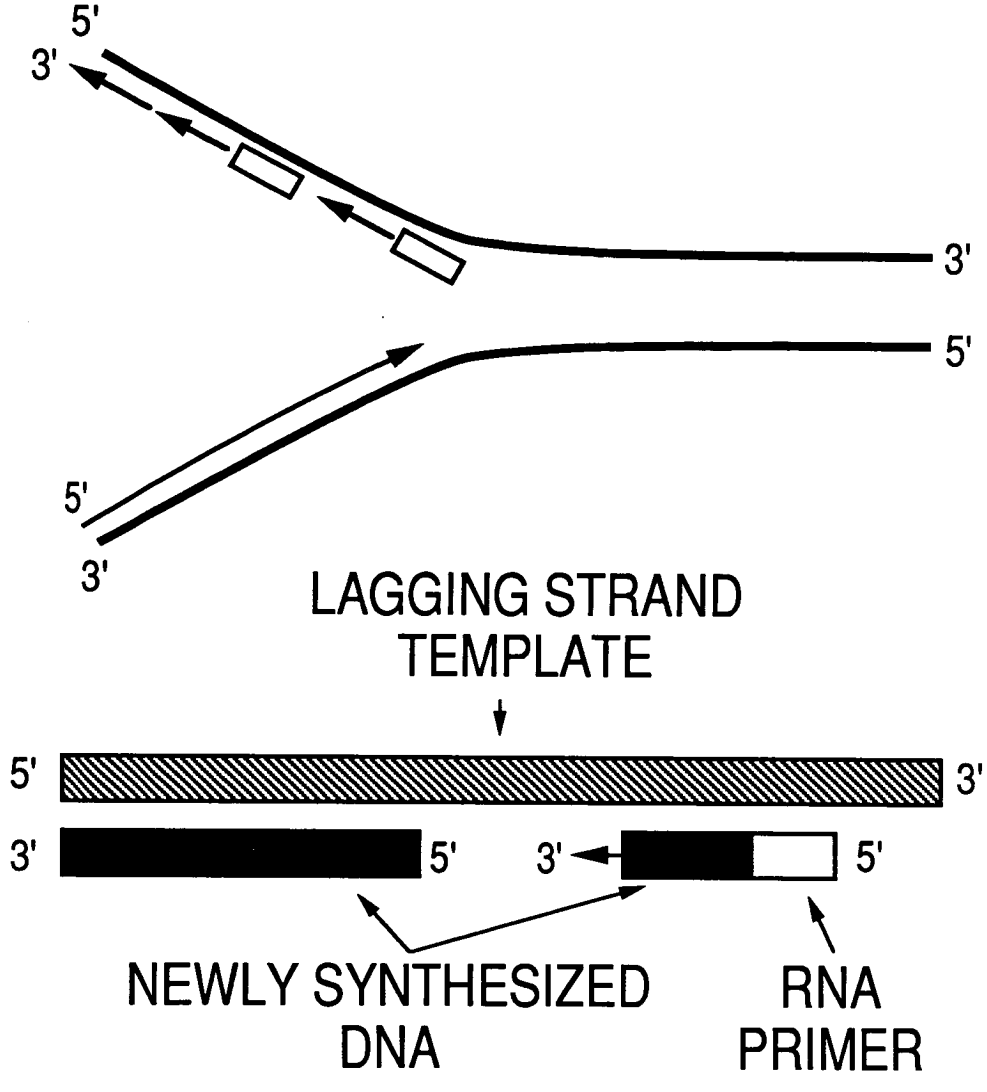
Primer extension experiments using $(CCCCAA)_4$ oligonucleotide and Klenow enzyme demonstrate that this oligonucleotide can anneal to the 3' extension of the G-rich strand found at telomeres and act as a primer for extension in the presence of dCTP and dATP (Henderson et al., 1988). Klenow lacks the 5'-3' exonuclease activity of DNA polymerase I necessary for nick-translation reactions on duplex DNA. Thus the DNA duplex of the terminal repeats must have an unusual structure to allow Klenow to strand-displace as the reaction proceeds in the absence of this exonuclease. Interestingly even in the presence of all four dNTP's the primer extension reaction cannot proceed past the terminal repeats indicating that the nature of the DNA duplex in the telomere-adjacent DNA is very different. However duplex telomeric repeats do not have an unusual helical repeat so changes in pitch in the double-helix do not account for the unusual characteristics conferred by the telomeric arrays of G-C rich repeats. As yet the molecular basis of the above observations is obscure.

1.4 The Problem of Telomere Replication

All known methods of DNA replication require a primer and only proceed in a 5'-3' direction. The structure of the telomere must be such as to allow the complete replication of the end, otherwise, at each round of replication material would be lost from the 5' end once the RNA primer had been removed and the chromosome would become progressively shorter (Watson, 1972). This problem is illustrated in Figure 1.1.

Figure 1.1

Diagram illustrating the problem of replication at the end of a linear DNA molecule. The lagging strand template is shown in the lower half of the diagram. Synthesis of new DNA uses one of the old strands as a template initiating from small RNA primers. These RNA primers are spaced about 200 nucleotides apart so the new DNA is synthesised in short fragments. The RNA is degraded and replaced by DNA as the new fragment reaches the adjacent fragment. The 3' end cannot be replicated as there is no DNA template for the RNA primer.



The telomeres of some linear eukaryotic virus genomes such as Vaccinia are covalently closed A-T rich hairpin structures which are incompletely base-paired (Baroudy et al., 1982). Replication of the hairpin results in the formation of concatemers which are then resolved by nicking of the daughter strands to separate the two molecules followed by religation (Merchlinsky and Moss, 1986). The telomeres of the adenovirus genome consist of inverted repeats containing the origin of replication. A protein is covalently attached to each telomere and acts as a primer for DNA synthesis thereby circumventing the need for an RNA primer (Stillman, 1983). Studies on the behaviour of eukaryotic chromosomal telomeres have indicated that a very different strategy has been employed to ensure that the ends are completely replicated. In the following section I will outline some of the observations which have lead to the proposal of a method of replication of eukaryotic chromosomal telomeres and the evidence which supports that proposal.

1.5 Growth and Change in Telomeres

As discussed in Section 1.2 in many species the number of telomeric terminal repeats is variable with the result that terminal restriction fragments have a characteristic "fuzzy" appearance on agarose gels. This is a feature of most chromosomal telomeres studied and reflects variation in length of the terminal array between all the telomeres in the population.

The dynamics of telomere length variation have been extensively studied in Tetrahymena. Telomeres coordinately decrease or increase in length in response to culture conditions. Experiments have showed that

during prolonged vegetative growth the macronuclear telomere length varied greatly such that the average length of terminal restriction fragments increased by 3-10 bp per generation until a strain-specific maximum is reached, whereas in stationary phase the telomere length decreased to a base-line level of 50 - 70 repeats (Larson et al., 1987). The length variations are attributable solely to variation in number of the terminal repeats (Henderson et al., 1988). Changes in telomere length have also been observed in trypanosomes (Bernards et al., 1983). It was shown that the telomere length increases by about 10 base pairs per generation in multiplying trypanosomes and under heat stress large deletions were observed. It was suggested that the increase is due to an increase in the number of the terminal tandem repeat TTAGGG by one repeat unit per generation (Van der Ploeg et al., 1984).

The yeast telomere is also a dynamic structure increasing and decreasing in length, but although telomere length is heterogeneous it does not increase significantly over many rounds of replication. Telomere length varies between yeast strains but it is fairly constant within a strain, the mean length being controlled by a number of genes (Walmsley and Petes, 1985). It has also been shown that the length of any one telomere in a cell is independent of the other telomeres and that the length of a particular telomere in a clonal population is determined by the length of that telomere in the starting cell (Shampay and Blackburn, 1988). The same study showed that in contrast to the results obtained with Tetrahymena telomere length was fairly constant during exponential vegetative growth but it increased when in stationary phase. All these length variations are due to changes in the length of the TG_{1-3} tract. These findings indicate that the yeast

telomere is maintained by a balance between addition and loss of the terminal repeat sequence, resulting in telomeres with heterogeneous terminal fragments which fall within a fairly narrow size range.

When linear or circular plasmids containing long tracts of the yeast telomeric repeat TG_{1-3} are introduced into yeast the length of the telomeric array of TG_{1-3} on the endogenous chromosomes increases (Runge and Zakian, 1989). The increase is gradual until a new steady state level is reached after about 20 generations and it is proportional to the amount of the introduced TG_{1-3} on the plasmids. This suggests that putative factor(s) which inhibit telomere elongation may be be competed out by introduced TG_{1-3}, so allowing the chromosomal telomeres to lengthen.

All the above results indicate that telomere length depends on a dynamic equilibrium between cellular activities which lengthen and shorten the terminal array of repeats. The action of nucleases and incomplete conventional replication would lead to a shortening of the array which could be counter-balanced by an activity which adds additional repeats onto the ends.

1.6 Complementation of Telomere Function

When Tetrahymena telomeres on linear plasmid vectors are introduced into S. cerevisiae they stabilise the plasmids and allow their successful replication and maintenance even though the terminal repeat sequence of Tetrahymena is different from that of yeast (Szostak and Blackburn, 1982). The plasmids terminating in TTGGGG were found to have acquired a tandem array of the yeast telomere sequence TG_{1-3} and it was proposed that the Tetrahymena repeat had

acted as a primer for the addition of the yeast sequence in a non-templated directed manner during replication (Shampay *et al.*, 1984). A similar result was found when *Oxytricha* telomeres were used (Pluta *et al.*, 1984). These experiments show that the simple repeat sequences are all that is required for telomere function and that structural features essential for telomere function but not sequence *per se* have been highly conserved through evolution. They also demonstrate that additional telomeric repeats can be added *de novo*. The orientation of the terminal repeats is crucial, if they are orientated such that the C-A rich strand runs 5'-3' towards the end there is no complementation of telomere function. It was speculated that a terminal transferase activity was extending the G-T rich strand and then the C-A rich strand was synthesied using this as a template from a primer. Incomplete sealing of nicks after removal of the primer or incomplete synthesis would account for the single-strand breaks which have been observed on the C-A rich strand (Shampay *et al.*, 1984).

1.7 Telomere Terminal Transferase

A terminal transferase activity, telomerase, has been found in *Tetrahymena* extracts (Greider and Blackburn, 1985). Synthetic single-stranded DNA oligonucleotides corresponding to the G-rich strand of *Tetrahymena* and yeast telomeric repeats can act as primers for addition of up to 30 tandem TTGGGG repeats onto their 3' ends whereas other primers tested, for example an oligonucleotide consisting of CCCCAA repeats, would not. The activity is greater during macronuclear development when a large number of telomeres are formed and replicated

than during vegetative cell growth. The addition of each repeat occurs one base at a time, but the *in vitro* reaction has a strong six-base periodicity involving pausing before addition of T residues.

Further experiments have shown that different oligonucleotides, corresponding to the telomeric G-rich strand from five species, can act as primers for the addition of TTGGGG repeats *in vitro* by *Tetrahymena* telomerase. This suggested that primer recognition must involve some feature of these oligonucleotides other than DNA sequence (Greider and Blackburn, 1987). The 3' end of the oligonucleotides specified the first nucleotide added. For example if the oligonucleotide used was $(TTAGGG)_3$ the added sequence began with GTTGGGG so that four G's preceded the first two T's as in the *Tetrahymena* repeat pattern.

Telomerase was shown to be a ribonucleoprotein complex in which both the RNA and protein component are essential for activity. Recently the cDNA for the essential 159 nucleotide RNA component of the telomerase has been cloned (Greider and Blackburn, 1989). The RNA contains the sequence CAACCCCAA which is complementary to one and a half TTGGGG repeats. This sequence appears to be the template for synthesis of TTGGGG repeats. Telomerase activity is abolished by cleavage of the RNA within this sequence. Oligonucleotides complementary to the putative template also abolish activity if added to the reaction before the primer, indicating that there is competition between the two for binding to the template sequence. An oligonucleotide complementary to the sequence just 3' to the template sequence is capable of priming the elongation reaction in spite of the fact that it does not resemble any known telomere sequence, adding further weight to the argument that the RNA is acting as an internal

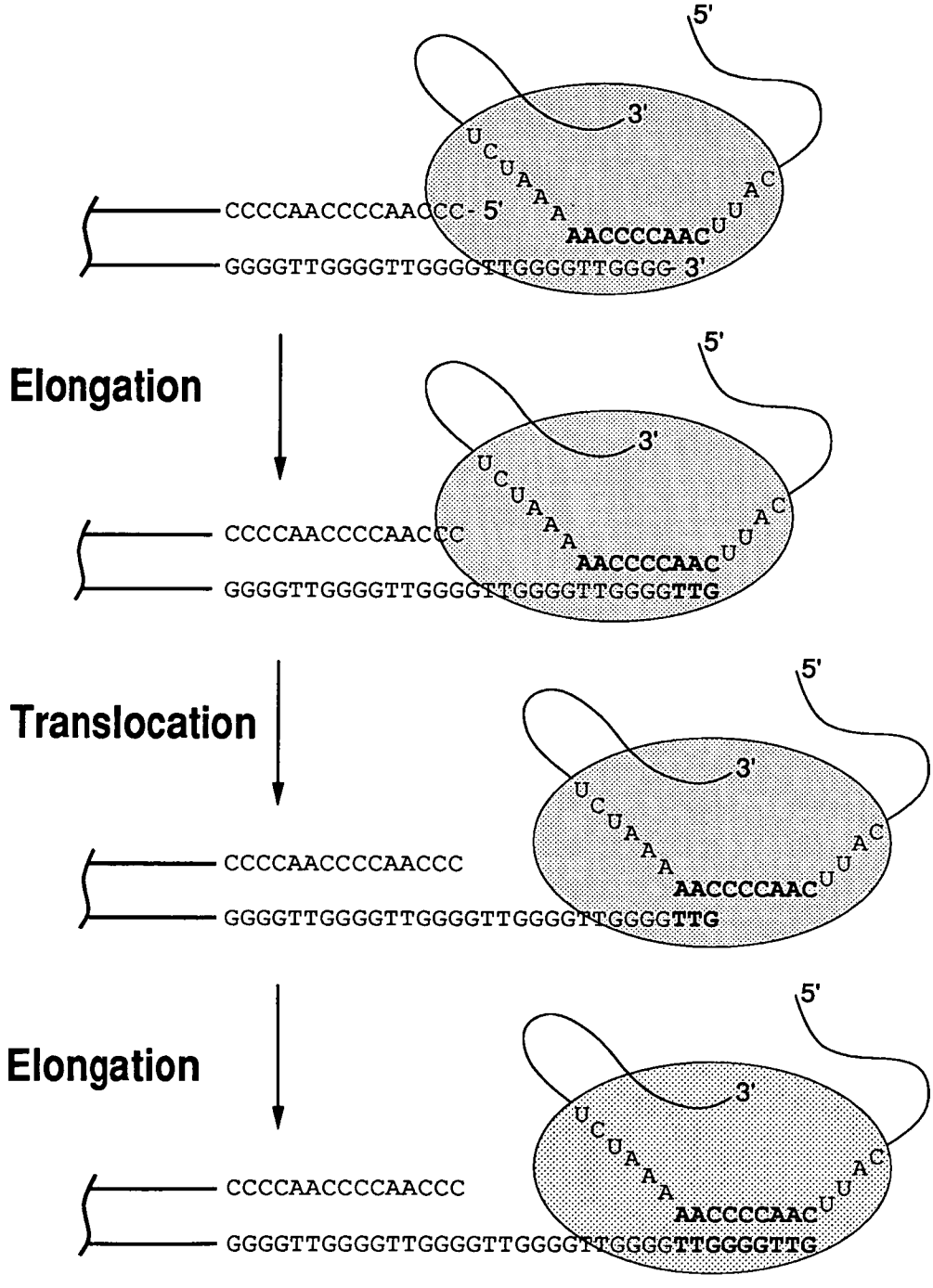
template. As the evidence suggests that the addition reaction is processive and many tandem repeats of TTGGGG are added onto primers a model based on an elongation-translocation mechanism has been proposed. This is illustrated in Figure 1.2. In this model, after primer recognition the most 3' part of the primer hybridizes to the most 3' part of the RNA template and the primer is extended to complete the last full repeat. After translocation the primer is repositioned so that its most 3' end is at the most 3' end of the template again ready for the next extension. Consistent with this model a pause after the addition of the nucleotides TTG has been observed (Greider and Blackburn, 1989). If this model is correct telomerase is an unusual polymerase in having a template as an integral part. By using an RNA template for DNA synthesis telomerase is unique among eukaryotic ribonucleoproteins.

Telomere terminal transferase activity has been found in two other ciliates, Oxytricha (Zahler and Prescott, 1988) and Euplotes (Shippen-Lentz and Blackburn, 1989). Repeats of TTTTGGGG are added onto the 3' end of primers corresponding to the G-rich strand of the telomeric repeats of these species. The addition reaction tends to pause after the addition of T residues. In contrast to the Tetrahymena activity the Oxytricha activity was found to be high in extracts from vegetatively growing cells and added only five to seven repeats to telomeric primers. A nuclease activity in the extracts appears to be in equilibrium with the telomere terminal transferase activity.

These differences may reflect the differences in organization of telomeres in hypotrichous ciliates compared to Tetrahymena. Telomeres of Tetrahymena typically consist of 50 to 70 tandem repeats of TTGGGG and the length can fluctuate under different growth condition as

Figure 1.2

Model for elongation of telomeres by telomerase using the elongation-translocation mechanism discussed in the text. Figure taken from Greider and Blackburn (1989).



discussed in section 1.5. The macronuclear telomeres of hypotrichous cilates such as Oxytricha or Euplotes all have the same species-specific structure which does not vary. Oxytricha telomeres consist of a duplex region with 2.5 repeats of the sequence TTTTGGGG with a 16 nucleotide extension of sequence $T_4G_4T_4G_4$ and Euplotes telomeres consist of a duplex region of 3.5 repeats of the sequence TTTTGGGG with a 14 nucleotide extension of sequence $T_4G_4T_4G_2$ (Klobuther et al., 1981; Pluta et al., 1982). Each Oxytricha macronucleus contains 4.8 x 10^7 telomeres whereas each Tetrahymena macronucleus contains 5 x 10^4 telomeres. Therefore during each cell cycle Oxytricha must synthesize far more telomeres than Tetrahymena. This may account for the differences found in telomerase activity between extracts made from vegetative cells of Oxytricha and Tetrahymena.

During macronuclear development in Euplotes over-sized telomeres are added to the gene-sized molecules which are then subsequently shortened to the length seen in vegetatively growing cells (Roth and Prescott, 1985). Extracts from mated Euplotes usually added approximately 10 repeats onto primers. This is the same number as that added in vivo showing that the telomerase is tightly regulated. New repeats were added one base at a time and the 3' end sequence of the primer was recognised as the last full repeat was completed before addition of new repeats. Primers corresponding to other telomeric repeats were recognised. Like Tetrahymena telomerase the Oxytricha and Euplotes terminal transferases are sensitive to RNase indicating that they too are ribonucleoproteins.

It is remarkable that the telomeric repeat of Tetrahymena can be recognised as a substrate for telomere elongation in the phylogenetically diverse organism S. cerevisiae and vica versa. It is

a striking feature of telomeres in these lower eukaryotes that the requirements essential for telomere function have been extremely highly conserved. Whether human telomeres share any of the properties so far observed for these telomeres will only be answered by cloning and analyzing the terminus.

1.8 Replication of Yeast Telomeres

Terminal transferase-like activities of the kind described in Section 1.7 have only been found in ciliates. As ciliates have such a high ratio of telomeres to total genomic DNA and a need to generate telomeres de novo telomerase activity might be expected to be high. In other organisms with a low number of chromosomes and therefore telomeres the activity may not be sufficient to be detected using the present assays. It is also possible that the ciliates represent a special case and that different mechanisms for telomere maintenance are employed by other organisms.

Some experiments have lent support to a proposal that telomere elongation in yeast may be based on recombination (Pluta and Zakian, 1989). In these experiments small linear molecules terminating in TTTTGGGG repeats at one end and TTGGGG repeats at the other end were constructed and transformed into yeast. Analysis of the resulting plasmids after propogation in yeast showed that in many cases TTGGGG sequences had been transferred to the TTTTGGGG end and vice versa before elongation of the ends by yeast TG_{1-3} repeats. It was suggested that recombination between the ends was responsible for this transfer and that similar recombination events between yeast telomeres could

account for telomere elongation. The transfer of sequences between ends did not require the RAD52 gene product and so may represent a telomere-specific recombination pathway. Other experiments in yeast showing that TG_{1-3} repeats can be added onto a stretch of non-telomeric DNA separating an array of TTGGGG repeats from the end of a transforming molecule by up to 100 bp argue against this recombination model, as it would be difficult to account for the retention of the non-telomeric DNA (Murray et al., 1988). These experiments showed that the elongation reaction only occurs when the G-rich strand of the telomeric repeat runs 5'-3' towards the end of the molecule. In the same study it was also shown that circular plasmids containing inverted head-to-head repeats of telomeric sequences are converted to linear molecules by a resolution reaction at a frequency of 10^{-2} per cell division and that these linear molecules are then efficiently elongated by the addition of TG_{1-3} repeats. Such resolution reactions are therefore unlikely to be crucial for telomere replication. Although there is no direct evidence for a telomere terminal transferase activity in yeast its presence would explain most of the results of experiments on telomere replication in yeast better than models involving recombination or inverted repeats as intermediates.

Recently a novel mutation has been described in yeast, est1-1 (for ever shorter telomeres), (Lundblad and Szostak, 1989). Null alleles of this gene have a senescence phenotype. The telomeres become progressively shorter over several generations and this is accompanied by decrease in chromosome stability. Mutant alleles of a gene necessary for telomere elongation might be expected to have a senescence phenotype as the consequences of an inability to maintain telomeres would only become apparant once essential information had

been lost owing to degradation of the chromosome from the end. Judging by the sequence of the gene the EST-1 protein is only poorly expressed in the cell. This, and the phenotype of the EST-1 gene, is consistent with it coding for a gene that is involved in telomere elongation, perhaps a component of a yeast telomerase. Confirmation of this awaits further investigation.

Other yeast mutants have been isolated with altered telomere structure, although null alleles of these genes are inviable rather than having a senescence phenotype that would be predicted for a gene involved solely in telomere replication. There are two genes for which mutants have been identified which result in shorter than normal telomeres (Lustig and Petes, 1986). The mutations are recessive and have a long phenotypic lag of about 150 generations. The telomeres do not shorten indefinately but reach a new steady-state level so the mutations might be changing the balance between elongation and degradation resulting in a new equilibrium length. A temperature-sensitive recessive lethal mutation of the cell-cycle gene CDC17 results in longer than normal telomeres at both the permissive and restrictive temperature, although the effect is more marked at the restrictive temperature (Carson and Hartwell, 1985). As temperature-sensitive mutants grown at the restrictive temperature arrest before nuclear division it was suggested that the wild-type gene might be necessary for complete replication which had the effect of preventing excessive telomere elongation. The product of the CDC17 gene has now been identified as DNA polymerase I (Carson and Hartwell, unpublished observations) so the alterations in telomere length seen with the

mutant allele are probably a secondary effect of the mutation. The est1-1 mutant described above segregates independently of all these mutations.

1.9 Chromatin Structure at Telomeres

The chromatin structure of the ends of macronuclear chromosomes in ciliates has been investigated. Generally the telomeric region, including the terminal repeats, is protected by a protein complex and phased nucleosomes start on the centromere-proximal side. This type of arrangement has been found in Oxytricha (Gottschling and Cech, 1984). The terminal 100-150 bp is protected by a protein complex and nucleosomes are phased in the proximal direction starting adjacent to this complex. A preferred site of micrococcal nuclease digestion is found at the junction between the single-stranded tail and the double-stranded part of the terminal repeats. The macronuclear telomeres of Tetrahymena (Blackburn and Chiou, 1981) appear to be tightly associated with protein complexes and not to be organised into nucleosomes judging by the pattern of micrococcal nuclease digestion seen. This is also the case for the telomeres of the RNA genes of Physarum (Cheung et al., 1981) and Dictyostelium (Edwards and Firtel, 1984). The chromatin structure of Tetrahymena ribosomal RNA genes has been further characterised (Budarf and Blackburn, 1986). These genes are present on linear 21 kb molecules at 10^4 copies/cell. The terminal TTGGGG repeats are packaged into a non-nucleosomal complex and the DNA just internal to the terminal repeats is organised into three phased nucleosomes. Recently the chromatin organisation of the Euplotes 5S gene minichromosome has been determined (Robertson et al., 1989). The

chromosome is only 930 bp long. Four nucleosomes are specifically positioned over the DNA and telomeric complexes protect the terminal 100 bp. Hypersensitive sites are found within the telomeric complexes. This terminal organisation is the same as that found on Oxytricha macronuclear chromosomes. The nucleosome phasing in all these telomeric regions is probably a consequence of a boundary imposed by the protein complex binding the terminal repeats.

1.10 Telomere Binding Proteins

The above studies indicate that proteins which bind to telomeres are likely to be universal and important for telomere function. Their role may be to protect the end from degradation or they may be necessary to confer topological constraints on the end of the chromosome so that super-coils can be introduced near the ends. They might also mediate the various telomere-telomere interactions that have been observed.

Proteins which bind specifically to telomeres have been isolated. Two distinct proteins which bind in a non-covalent way to the telomeres of Oxytricha macronuclear chromosomes and protect them from Bal-31 exonuclease have been identified (Gottschling and Zakian, 1986). They appear to be specific for both the sequence and structure of the telomere, although the 80-130 bp of telomere-adjacent DNA which is protected does not have any conserved features. To bind efficently these proteins require the duplex region of the G_4T_4 repeat, but a more important requirement is the 16 nucleotide 3' overhang. The internal structure of these telomeric complexes has been investigated (Price and Cech, 1987). The data suggest that the telomeric complex

falls into two structural domains with very different DNA-protein interactions. Salt-stable, sequence specific DNA-protein interactions characterise the terminal part of the complex. Specific guanine residues in the $T_4G_4T_4G_4$ terminal extension are protected from methylation when treated with dimethyl sulphate. In contrast, in the region 45-135 bp from the end there are no salt-stable or sequence specific DNA-protein interactions and the DNase I cleavage pattern suggests that the DNA is positioned on the outside of the complex.

Recently it has been shown that the two proteins are components of a heterodimer which is very difficult to dissociate (Price and Cech, 1989). It is tightly bound to the 3' telomeric extension, which it can protect from extensive nuclease digestion. The affinity of this heterodimer for a variety of oligonucleotides has been assessed (Raghuraman et al., 1989). Single-stranded oligonucleotides as short as 16 nucleotides of the T_4G_4 sequence are efficiently bound. The sequence at the 3' end of the single-stranded oligonucleotides is the important factor in the binding reaction. An oligonucleotide consisting of a stretch of T_4G_4 followed by a stretch of CA at the 3' end is not bound by the protein. This protein heterodimer probably determines the length of the 3' extension by binding once 16 nucleotides have been added by the telomerase, thereby preventing the addition of more repeats and the degradation of the telomere by nucleases.

A protein which will bind to the TG_{1-3} repeats of yeast, whether internally or terminally located, has been described. The protein binds to both cloned telomeres and to genomic yeast DNA (Berman et al., 1986). This protein is likely to be RAP1 (repressor/activator site binding protein) (Shore and Nasmyth, 1987; Buchman et al.,

1988a). This protein was originally characterised because it appears to be a transcriptional regulator and can bind both to yeast mating type silencers and to upstream activation sites *in vitro*. On examination of the sequences recognized by the protein it became apparant that it might be able to bind to yeast telomeric repeat, and indeed this was found to be the case. Further work suggests that RAP1 does act *in vivo* at the sequences recognised by the factor *in vitro* (Buchman *et al.*, 1988b). Elegant experiments have shown that RAP1 is required for DNA loop formation at the silent mating type locus HML *in vitro* (Hofmann *et al.*, 1989). RAP1 fractionates with scaffold proteins and it may be that sequences such as silencers bound by RAP1 represent scaffold attachment sites where chromatin loops and transcriptional control loops coincide. RAP1 may perform different functions depending on the chromosomal environment in which it is bound. It is possible that RAP1 is the factor responsible for inhibition of telomere elongation suggested by the experiments of Runge *et al.* (1989) described in Section 1.5.

All of these telomere-binding proteins are sequence specific and will not bind to different telomeric repeats from other lower eukaryotes. The presence of such telomere-binding proteins might explain the inaccessibility of chromosome termini to end-labelling observed in some species (Blackburn and Gall, 1978). The inaccessibility could also be a consequence of the unique telomeric DNA sequence, for example *Tetrahymena* ends may have a fold-back structure owing to self-association of the guanine residues in the 3'extension (Henderson *et al.*, 1987).

1.11 Telomere Healing

When chromosomes are broken the ends are sometimes "healed" by the acquisition of new telomeres. This phenomenon has been observed in many different species. One of the first was maize where usually broken chromosomes have very reactive ends and enter "breakage-fusion-bridge" cycles (McClintock, 1938). By contrast if a chromosome is broken during meiotic anaphase, and the gamete containing that chromosome is fertilised, the broken end of that chromosome will heal and behave like a normal telomere once the zygote is formed (McClintock, 1941).

In Plasmodium, chromosomal polymorphisms between strains, characterised by terminal deletions, have been found (Pologe and Ravetch, 1988). These deletions often result in the loss of terminally located antigen genes and therefore lead to antigenic variation between strains. Chromosome breakage followed by healing of the ends by the addition of telomeric repeats best explains the generation of these polymorphisms.

The fate of broken chromosomes has been studied in Drosophila. A spontaneously opened ring chromosome has been examined (Traverse and Pardue, 1988). The resulting linear chromosome is stable and the new ends have acquired He-T DNA sequences, which are sequences found in the telomeric and pericentric heterochromatin. The method by which the broken ends have acquired these sequences is not clear and whether the extreme ends resemble genuine Drosophila telomeres is unknown as Drosophila telomeres have not yet been isolated. Recently viable terminal deletions from Drosophila chromosomes have been described. These viable deletions have been generated in two ways, either by X-

irradiation of Drosophila females carrying the mu-2 mutation, which is defective in repairing X-ray induced chromosome breaks, (Biessmann and Mason, 1988), or by destabilization of a P element transposon inserted close to a telomere (Levis, 1989). In both cases deleted chromosomes lose sequences from the broken end at a rate of approximately 75 bp per generation, probably due to incomplete replication and the action of nucleases. In spite of the fact that the terminally deleted chromosomes lack telomeres they are not unstable and the broken ends do not appear to be liable to fusion events or to be recombinogenic. In most cases healing of the broken chromosomes by addition of new sequences was not observed. Whether the broken ends generated in these experiments are stable because of some property not shared by other broken ends which have been studied requires further investigation.

In cilates during the generation of the macronucleus from the micronucleus the genome is fragmented and new telomeres are added to the resulting macronuclear chromosomes in order to stabilize them. The addition of telomeric repeats at this stage of the life-cycle can be regarded as a healing reaction as the ends of the fragments added to do not bear any resemblance to G-rich telomeric sequences. This developmentally regulated telomere addition reaction has been studied in Paramecium (Baroin et al., 1987) and in Tetrahymena (Spangler et al., 1988). In several examples of the same telomere in Paramecium the position at which the terminal repeats were added varied by up to 0.8 kb. In Tetrahymena, telomere-adjacent sequences from different macronuclear chromosomes are not similar and two examples of the same macronuclear telomere have a 5 bp difference at the junction between the terminal repeats and the unique DNA. Oligonucleotides corresponding to telomere-adjacent sequences do not act as primers for

repeat addition by Tetrahymena telomerase indicating that the nature of the reaction resulting in the initial addition of telomeric repeats is very different from that which maintains telomere length.

In S.cerevisiae telomere healing reactions may be mediated by homologous recombination events involving Y' elements of internally-located tracts of TG_{1-3} repeat as discussed in Section 1.2. De novo addition of telomeric repeats onto broken ends has not been demonstrated. However such events best explain the structure of the ends of healed chromosomes in the fission yeast Schizosaccharomyces pombe after chromosome breakage by gamma irradiation (Matsumoto et al., 1987).

These telomere healing reactions are only poorly understood at present, but it seems that that the mechanisms involved may well be different from those involved in telomere elongation.

1.12 The Human Pseudoautosomal Region

The first cloned DNA sequence shown to be near the telomere of a human chromosome was isolated from the pseudoautosomal region of the Y chromosome (Cooke et al., 1985). As the pseudoautosomal region has some unique properties I will outline here some of it's essential features.

Generally at least one chiasmata is required for the successful segregation and disjunction of bivalents at meiosis I metaphase (Murray and Szostak, 1985). There are, however, a few examples of organisms in which correct segregation occurs in the absence of recombination. There is no recombination during gametogenesis in male Drosophila but there is orderly segregation. The mechanisms which

ensure correct segregation are not completely understood but at least in the case of the sex chromosomes processes hold the homologues togeather at special pairing sites (Baker et al., 1975). In female Drosophila segregation of the smallest chromosome depends on a distributive system which segregates chromosomes on the basis of size (Baker et al., 1976). In yeast fidelity of segregation at meiosis I principally depends on cross-over, but experiments with artificial chromosomes have shown that additional mechanisms exist which ensure correct segregation (Dawson et al., 1986).

The presence of a least one chiasmata appears to be a requirement for orderly segregation of human chromosomes. The human X and Y chromosomes are very different in both size and DNA content but the existence of an homologous pairing region was postulated because of this requirement for a chiasma (Koller and Darlington, 1934). Pairing of the X and Y chromosomes at the tips of the short arms has been seen at meiosis I (Chandley et al., 1984) and now DNA markers have been isolated from this region and their "pseudoautosomal" nature demonstrated. In spite of the fact that these markers are located on the sex chromosomes they show varying degrees of sex linkage, that is, they behave as if they were autosomally located hence the term pseudoautosomal.

The DNA probe 29C1 defines the locus DXYS14 and detects hypervariable DNA fragments in many restriction digests (Cooke et al., 1985). 29C1 is not sex-linked as it recombines at a frequency of 50% with the testis determining factor. This is consistent with it having a terminal location and there being an obligate cross-over which in male meiosis can only take place in the homologous pseudoautosomal region of the X and Y chromosomes. Other markers isolated from the

pseudoautosomal region exhibit a gradient of sex-linkage, the more proximal the marker the greater degree of sex-linkage (Rouyer et al., 1986a). 29C1 detects either one or two blocks of homology in all the chromosomes so far examined, all blocks being located in the terminal region of the X/Y pseudoautosomal region. The homology consists of an irregular tandem array of a 31 base-pair repeat which varies in length between individuals (Inglehearn, pers. comm.). The array is reminiscent of "minisatellite" sequences which can contract and expand (Jeffreys et al., 1985).

With some restriction enzyme digests 29C1 detects heterogenous fragments which become smaller if the DNA is pre-treated with Bal-31 exonuclease before digestion. This indicated that 29C1 was located very close to the telomere (Cooke et al., 1985). DNA from blood, sperm and a lymphoblastoid cell line from one individual were analysed and terminal heterogeneity detected in all samples. Restriction maps were identical up to 1 kb from the telomere but the size of restriction fragments detected from sperm were about 5 kb longer than the other samples. DNA from other individuals gave the same result (Cooke and Smith, 1986). Terminal restriction fragments detected by 29C1 in the lymphoblastoid cell line PES, although heterogeneous, were smaller and varied less than than those of blood and sperm telomeres from the same individual. These results strongly suggest that arrays of terminal repeats are present at the ends of human chromosomes.

29C1 has been used to investigate the nature of the terminus indirectly by monitoring the effect of various nucleases. These experiments suggest that the terminus is a 3' overhang, probably a non-covalently closed hairpin (Cooke, 1987).

The recombination frequency of pseudoautosomal markers in male meiosis is ten times greater than that in female meiosis. This is to is to be expected as the obligate crossover in female meiosis would not be confined to the pseudoautosomal region but to a much larger region of the whole chromosome. In male meiosis only one cross-over in the pseudoautosomal region has been observed to date (Rouyer et al., 1986b). Using rare-cutting restriction enzymes and pulsed-field gel electrophoresis 29C1 has been used to establish a long-range restriction map from the telomere by indirect end-labelling (Brown, 1988; Rappold and Lehrach, 1988; Petit et al., 1988). Estimates of the size of the region range from 2.3 Mb to 3 Mb. The genetic and physical maps are in agreement which infers that recombination frequency is uniform throughout the region. The boundary between the psuedoautosomal region and sex-specific DNA has recently been cloned and characterised (Ellis et al., 1989). The boundary appears to be defined by an Alu sequence inserted on the Y chromosome which is absent from the X chromosome.

1.13 Human Telomeres

The studies on the pseudoautosomal region telomere show that the telomeres of human chromosomes share many properties with those of lower eukaryotes. Human telomeres, like those of lower eukaryotes, appear to be dynamic structures, there is variability in the length of terminal restriction fragments and the terminus appears to be a 3' overhang. Whether the structure of human telomeres bears any resemblence to those of the eukaryotes already studied is one of the questions which will be addressed in this thesis. In addition the

implications that the deduced structure has for the mechanisms involved in the maintenance of human telomeres will be discussed with reference to those known mechanisms employed by lower eukaryotes.

CHAPTER 2

MATERIALS AND METHODS

2 MATERIALS AND METHODS

2.1 MAMMALIAN CELL CULTURE

2.1.1 Maintenance of cells and cell lines in culture

Cells were grown in RPMI 1640 medium (Flow Laboratories) supplemented with 10% foetal calf serum (FCS) (Gibco Bio-cult) which had been inactivated for 30 mins. Penicillin and streptomycin (Gibco Bio-cult) were added to cultures at 100 units/ml and 100µg/ml respectively. All cultures were grown at 37°C in 25, 80, 175 cm^2 plastic flasks or 96 well dishes (Nunc) in a 10% CO_2 atmosphere.

Cells were routinely fed or sub-cultured twice weekly. Cells to be split or harvested were first washed with Dulbecco's phosphate buffered saline (PBS) and were detached from the culture vessel with versene containing 10% trypsin.

Frozen stocks were made by resuspending harvested cells in 90% FCS/10% dimethylsulpoxide (DMSO) at not less than 10^6 cells/ml and freezing slowly overnight in vials at -70°C in a polystrene box. Vials were stored in liquid nitrogen. Cell lines were recovered from frozen by resuspending rapidly thawed cells in 10 mls of medium and plating in a fresh tissue culture flask.

2.1.2 Cell Lines

PES — A human male EBV transformed lymphoblastoid cell line. Grown as a suspension culture in medium supplemented with 10% trypto phosphate broth (TPB).

X63/NS1 — Local name for mouse myeloma line P3-NS1/1-Ag4-1 which is $HPRT^-$ (Köhler et al., 1976). Grown as a semi-

attached culture in medium supplemented with 0.2 international unit of porcine insulin/ml, 1mM oxoloacetic acid and 0.45 mM sodium pyruvate.

X-MGU1-9 Hybrid cell lines between PES and X63/NS1. Grown in the same medium as X63/NS1 supplemented with HAT (10^{-4}M hypoxanthine, 10^{-6}M aminopterin, 1.6 x 10^{-5}M thymidine) which maintains selection for cells carrying the human X-linked HPRT gene.

2.1.3 Cell Fusions

The cell fusion protocol used was that described in Claflin and Williams (1978) but with RPMI 1640 medium and with 35% polyethylene glycol (PEG) 1500 (BDH) made up in serum-free medium containing 5% v/v DMSO and 5 x 10^{-5}M β-mercaptoethanol. The fusion products were plated out in X-MGU medium onto four 96 well plates which contained 10^4 activated mouse macrophages/well. Ouabain to a concentation of 10^{-6}M was added the next day and ouabain selection was maintained for 10 days. After four weeks successful fusions were transferred to culture flasks and grown in X-MGU medium.

2.1.4 FACS Analysis

The monoclonal antibody used for fluorescence-activated cell sorting (FACS) analysis of the somatic cell hybrids was 12E7 (gift of Peter Goodfellow) which recognizes the pseudoautosomal marker MIC2 (Darling et al., 1986). The negative control antibody was AFP, which recognizes human α-fetoprotein (van Heyningen et al., 1982).

Cells were prepared for FACS analysis as described by Seawright et al. (1988). 10^6 cells were incubated with saturating amounts of 12E7, and fluorescein isothiocyanate-labelled sheep anti-mouse IgG, at a 1:40 dilution, was used as a second antibody. Cells were analyzed for the distribution of fluorescence intensity by using a FACS IV cell sorter. Approximately 10,000 cells were counted and data were obtained as a histogram of cell number plotted against relative fluorescence.

2.2 EXTRACTION OF DNA FROM MAMMALIAN CELLS

2.2.1 Extraction of DNA From Cells Grown in Culture

Harvested cells were pelleted at 800 rpm for 5 mins and washed twice in PBS. The number of cells was estimated by using a Coulter cell counter or Neubauer heamocytometer. Cells were resuspended in SE (75mM NaCl, 25mM EDTA pH 8.0) at 10^7 cells/ml and proteinase K was added to 100µg/ml. After thorough resuspension of the cells SDS was added to 1% and the lysate incubated for 2-4 hrs at 37°C or overnight at room temperature. Then an equal volume of SE was added and the lysate extracted three times with an equal volume of 25:24:1 phenol/chloroform/octan-2-ol (the phenol was buffered with 0.1M Tris HCl pH 8.0 and contained 0.1% 8-hydroxyquinoline) by mixing gently until homogeneous and then separating the phases by spinning for 10 mins at 1K. The organic phase, and any flocullence at the interface, were discarded after each extraction. To precipitate the DNA 1/30 volume of 3M NaOAc pH 5.5 and 1 volume of propan-2-ol were added and mixed gently at room temperature. The DNA was spooled out, rinsed in 70% ethanol and dissolved overnight at 4°C in TE (10mM Tris-HCl pH

7.5, 1mM EDTA pH 8.0). The DNA concentration was estimated by measuring the optical density (OD) at 260nm (an OD of 1 at 260nm is equivalent to a DNA concentration of 50µg/ml).

2.2.2 Extraction of DNA From White Blood Cells

Blood (~10mls) was collected into EDTA (1mg/ml) to prevent clotting. 3 vols 155mM NH_4Cl, 10mM $KHCO_3$, 0.1mM EDTA pH 7.4 were added and the blood left at 4°C for 15 mins with gentle mixing to lyse the red blood cells. The white blood cells were pelleted by spinning at 1K for 10 mins and resuspended in 10mls of SE. The rest of the extraction was as described in 2.2.1 for tissue culture cells.

2.2.3 Extraction of DNA From Sperm

10 mls 50mM EDTA pH 8.0, 1% SDS was added to a sperm sample followed by 5mg Proteinase K and it was left overnight at room temperature. Then dithiothreitol (DTT) was added to 50mM and Proteinase K to 100µg/ml and the lysate was incubated overnight at 50°C. The DNA was extracted twice with phenol, once with chloroform and then precipitated and resuspended in TE as described in 2.2.1.

2.3 BACTERIAL CELL CULTURE

2.3.1 Media and Additives

All media was sterilised by autoclaving.

L-Broth and Agar

Per litre : 10g tryptone (Difco), 5g yeast extract (Difco), 10g NaCl, 2.46g $MgSO_4$ pH 7.2. L agar contains in addition 15g agar/litre and L top agar contains 6.5g/litre.

H Agar

Per litre : 10g tryptone, 8g NaCl, 12g agar pH 7.0 (H top agar contains 6.5g/litre). This agar reduces the number of satellite colonies when plating out bacterial transformations where the drug used is bacteriostatic.

2 x TY Broth

Per litre : 16g Na_2HPO_4, 10g yeast extract, 5g NaCl, pH 7.0.

"Terrific" Broth (BRL Focus 1987)

Per 900ml : 12g tryptone, 24g yeast extract, 4mls glycerol. 100ml of 0.1M KH_2PO_4/0.72M K_2HPO_4 added after sterilising both solutions.

SOB and SOC Broth

Basic media per litre : 20g tryptone, 5g yeast extract, 0.59g NaCl, 1.86g KCl.

SOB : Before use filter sterilised $MgCl_2$ and $MgSO_4$ were added to a final concentration of 10mM.

SOC : As SOB except made 20mM glucose from 2M filter sterilised stock.

ψBroth

Per 990ml : 20g tryptone, 5g yeast extract pH 7.6 with KOH. Just before use 10ml 100 x salts added.

100 x ψSalts

1M $MgCl_2$, 1M $MgSO_4$, 1M NaCl, 0.25M KCl.

10 x M9 Salts

Per litre : 60g $NaHPO_4$, 30g KH_2PO_4, 10g NH_4Cl, 5g NaCl.

Glucose/Minimal Agar

Per litre : 15g agar in 900ml H_2O. Autoclave and whilst still liquid add 100ml 10 x M9 salts, 1ml 1M $MgSO_4$, 1 ml 0.1M $CaCl_2$, 1ml 1M thiamine and 10ml 20% glucose. All these are sterile solutions.

Media Additives

When appropriate antibiotics were added to media and agar. Ampicillin was used at a concentration of 50μg/ml. X-Gal (5-Bromo-4-chloro-3-indoyl-β-galactopyranoside) stock solution was 2% in dimethylformamide. IPTG (isopropyl-β-D-thio-galactopyranoside) stock solution was 100mM.

Bacteria were grown at 37°C, with good aeration for liquid cultures. Bacterial stocks were kept at -70°C in 20% glycerol.

2.3.2 Bacterial Strains Used

All the bacterial strains used were *E. coli* K-12.

JM83 : *ara*, Δ(*lac-proAB*), *rspL*, F80, *lacZ*ΔM15, (rk^+,mk^+) (Vieira and Messing, 1982). A host for pUC plasmids. The lacZ M15 gene is integrated into the host chromosome.

JM101 : Δ(*lac-proAB*), *thi*, *supE*,{F'*traD36*, *proAB*, *lac* $1^qZ\Delta M15$} (Yanisch-Peron *et al.*, 1985). Strain used for the propogation of M13. Since JM101 contains the *lacI*^{q} mutation which overproduces the *lac* repressor, IPTG must be added to induce β-galactosidase synthesis (Müller-Hill *et al.*, 1968).

DH5α : Derivative of DH1 (Hanahan, 1983). High efficiency transformation strain for pUC-based plasmids. F^-,*recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17* (rk^-,mk^+), *supE44*, *relA1*, λ-, *mcrA*(+), *mcrB*(-), Δ(*argF-lacZYA*)U169, ϕ80d*lacZ*ΔM15.

XL1-Blue : *recA1*, *endA1*, *gyrA96*, *thi*, *hsdR17*, (rk^-,mk^+), *supE44*, *relA1*, λ-,(*lac*), {F', *proAB*, *lacI*^{q}*Z*ΔM15,Tn10(tet^R)} (Bullock *et al.*, 1987). A host for pUC-based plasmids.

2.4 VECTORS FOR CLONING IN E. COLI AND YEAST

2.4.1 Plasmid Vectors

pUC 9

The pUC series of plasmid vectors are based upon pBR322 (Bolivar et al., 1977) retaining both a high copy number, and the ampicillin resistance gene (β-lactamase) (Vieira and Messing, 1982). All of the pUC vectors contain a polylinker inserted into the β-galactosidase gene (lacZ). Non-recombinant plasmids are able to synthesise this enzyme which breaks down X-Gal to release a blue indolyl derivative. In recombinant plasmids the lacZ gene is interrupted by foreign DNA, thus the colonies remain colourless.

pTZ

pTZ is a 2.9 kb plasmid derived from pUC plamid with an insertion of the F1 origin of replication.

pBS

pBS, formerly pBluescribe, is a 3.2 kb phagemid derived from pUC19. The vector carries the colE1 origin, ampicillin resisitance, T3 and T7 promoters flanking the pUC19 polylinker and the lacZ promoter.

pYAC4NEO

pYAC4NEO (Cooke and Cross, 1988) is a yeast artificial cloning vector derived from pYAC4 (Burke et al., 1987).

2.4.2 Vectors based on the single-stranded DNA coliphage M13

M13, a single stranded (ss) DNA filamentous phage of E.coli infects cells via the F.pilus. Replications occurs via a double-stranded (ds) replicative form (RF) giving rise to ss progeny virions. M13 has been developed as a cloning vector for DNA sequencing (Messing, 1983; Norrander et al. 1983). Foreign DNA is cloned into the polylinker of the M13 RF. The two versions used, M13mp18 and M13mp19 contain the polylinker in different orientations. As for the pUC plasmid vectors, recombinant molecules can be identified by the disruption of the lacZ gene using the X-Gal assay.

2.5 MANIPULATION OF PLASMID DNA

2.5.1 Preparation of Competent DH5α

All glassware used was free from detergent. 10 colonies from a fresh overnight plate were innoculated into 100 ml Ψ broth, shaken at 37°C for 2-3 hours until OD_{560nm} = 0.45 - 0.55. The cells were placed on ice for 10 mins and all subsequent steps were performed at 4°C. After pelleting the cells at 2.5K for 10 mins the cells were gently resuspended in 33 ml cold TFB (100mM $RbCl_2$, 45 mM $MnCl_2$, 10mM $CaCl_2$, 3mM hexamine cobalt chloride, 10mM K-MES pH 6.3, filter-sterilised). After 10 mins they were repelleted at 2.5K for 10 mins and then resuspended in 8 ml TFB. 280µl ultrapure DMSO was then added to the cells, 5 mins later 280µl 2.25M DTT, 40mM KOAc was added and then 10 mins later another 280µl DMSO was added. After 5 mins 210µl aliquots of the cells were transferred to pre-cooled Eppendorf tubes.

2.5.2 Preparation of Frozen Cells for Electroporation

1 litre of L broth was innoculated with 0.1ml fresh overnight XL1-Blue culture and the cells grown at 37°C with vigorous shaking to an OD_{600nm} of 0.5 - 1. The cells were harvested by cooling for 15 - 30 mins on ice and then pelleting at 4,000 x g_{max} for 15 mins at 4°C. All subsequent steps were performed at 4°C. The cells were resuspended in 1 litre H_2O and repelleted as before. The cells were then resuspended in 500 ml H_2O and repelleted as before. The cells were resuspended in 100 ml H_2O and repelleted as before. Then the cells were resuspended in 20 ml 10% glycerol, repelleted as before and the cells resuspended to a final volume of 2 - 3ml in 10% glycerol. The cell concentration should be approximately 3 x 10^{10} cells/ml. 40µl aliquots were quick frozen and stored at -70°C.

2.5.3 Transformation of Competent DH5α

Plasmid DNA (<1µg and <10 µl) was added to an aliquot of competent DH5α and left on ice for 30 mins. The cells were heat-shocked at 42°C for 100 secs and replaced on ice. After addition of 800µl L broth containing 20mM glucose the cells were incubated at 37°C for 60 mins with shaking to allow for the expression of β-lactamase. Cells were spread on H agar containing ampicillin. 40µl IPTG was added to the cells and X-Gal (100µl 2% in dimethylformide/30 ml plate) was included in the agar if discrimination between recombinant and non-recombinant colonies was required.

2.5.4 Electroporation of XL1-Blue

An aliquot of the XL1-Blue cells prepared for electroporation was thawed at room temperature and then placed on ice. Plasmid DNA (<1µg in 1 - 2 µl TE or H_2O) was added and mixed well and the mixture left on ice for 1 min. The Gene Pulser apparatus was set to 25µF and 2.5 kV and the Pulse Controller to 200Ω. The mixture of cells and DNA was transferred to a cold, 0.2cm electroporation cuvette, the cuvette placed in a cold safety chamber slide and the slide pushed into the chamber until the cuvette was positioned between the contacts at the base of the chamber. The cells were pulsed once which should produce a pulse with a time constant of 4.5 to 5 msec (the field strength was 12.5 kV/cm). After removal of the cuvette from the chamber 1ml of SOC was immediately added and the cells quickly resuspended with a pastette. The cell suspension was then transferred to a 17 x 100 mm polypropylene tube and incubated for 60 mins at 37°C with shaking at 225 rpm. The cells were plated out onto selective medium as described in 2.5.3 for transformation of competent DH5α.

2.5.5 Large-scale Plasmid Preparation

Plasmid DNA was isolated by a modified version of the alkaline-lysis method of Birnboim and Doly (1979).

500ml broth containing ampicillin was innoculated with a single colony and the culture grown overnight. The cells were pelleted by centrifugation at 6K for 10 mins at 4°C and resuspended in 20ml solution I (25% glucose, 50mM Tris-HCl pH 8.0, 10mM EDTA pH 8.0) and lysozme was added to 0.5 mg/ml. After 10 mins at room temperature 60 ml fresh solution II (0.2M NaOH, 1%SDS) was added and mixed gently. After 10 mins on ice 45ml cold solution III (3M CH_3COOK pH 4.8 with

acetic acid) was added and the lysate left on ice for 30 mins. Cell debris was pelleted at 8K for 5 mins and the supernatant strained through muslin. DNA was precipitated by the addition of 0.6 vols of propan-2-ol and and collected by pelleting at 8K for 5 mins at 4°C. The pellet was washed with 70% ethanol and resuspended to 13ml with TE. 1.58g CsCl and 1.5ml of 10mg/ml ethidium bromide were added such that the refractive index of the resulting solution was 1.393. The plasmid DNA was banded by centrifugation in an 18ml polyallomer tube at 40K, 20°C overnight in a vertical rotor (which allows the CsCl gradient to reach equilibrium rapidly). Plasmid DNA is more dense than the bacterial chromosomal DNA, due to plasmid supercoiling constraining the amount of EtBr which can be intercalated into the molecule. The plasmid band was collected under UV light (300nm) with a syringe and the EtBr removed by a series of extractions with butan-2-ol until no colour remained in the aqueous phase. The DNA was then precipitated with 3 vols of 70% ethanol.

2.5.6 Small-Scale Plasmid Preparation

The solutions used were the same as for the large-scale method (2.5.5). 1.5 ml of a fresh overnight culture was transferred to an Eppendorf centrifuge tube, the cells pelleted by centrifugation, drained and resuspended in 200µl solution I. A few flakes of lysozyme were added and the cells left for 5 mins at room temperature. 400µl solution II was added and mixed gently. After 10 mins on ice 200µl solution III was added and the lysate left on ice for another 10 mins. Cell debris was precipiated by centrifugation and removed using a flat-ended toothpick. The DNA was precipitated with 480µl propan-2-ol

and left on dry ice until frozen. The DNA was pelleted by centrifugation for 10 mins, washed in 70% ethanol and the DNA resuspended in 50µl TE.

2.6 MANIPULATION OF M13 DNA

2.6.1 Preparation of JM101 cells competent for transformation by M13 DNA

JM101 cells were streaked onto glucose minimal medium plates which selects for cells carrying a plasmid encoding a gene involved in proline synthesis (the host cell has a deletion in this gene). The same plasmid is needed for synthesis of the F-pilus required for infection by M13.

1ml of an overnight culture of JM101 was used to inoculate 100ml 2 x TY and the cells grown to OD_{550nm} = 0.4. After harvesting the cells were resuspended in 50ml cold 50mM $CaCl_2$ and left on ice for 20 mins. The cells were then repelleted and resuspended in 10ml of the $CaCl_2$ solution, and stored at 4°C for up to 4 days.

2.6.2 Transformation of JM101 by M13 DNA

1ng ds recombinant M13 DNA was added to 300µl of competent cells. After 40 mins on ice the cells were heat-shocked at 42°C for 3 mins then returned to ice. 200µl of a fresh exponential culture of JM101 was added and the mixture plated in H top agar. For identification of recombinant phage-containing plaques 40µl 2% X-Gal and 40µl fresh 100mM IPTG were incorporated into the top agar.

2.6.3 Preparation of single-stranded M13 DNA

100ml 2 x TY were inoculated with 1ml of an overnight culture of JM101. An isolated M13 plaque was added to 1.5ml of these cells and grown at 37°C for 5 hours. The supernatant was respun to ensure no bacterial cells were remaining, then the phage particles precipitated by the addition of 200µl of 20% PEG 6000 in 2.5M NaCl. After 15 mins on ice the phage were collected by centrifugation and resuspended in 100µl TE. After phenol/ chloroform and chloroform extractions the ss M13 DNA was ethanol precipitated.

2.7 YEAST CELL CULTURE

2.7.1 Media and Additives

All media was sterilised by autoclaving.

YPD Broth and Agar

Per 500ml : 5g Bacto-yeast extract (Difco) and 10g Bacto-peptone (Difco) made up to 450ml with H_2O. 50ml of 20% glucose added after sterilising both solutions. YPD agar contains 10g agar in addition.

Yeast Selective Broth and Agar - Uracil

Per 500ml : 3g $(NH_4)_2SO_4$, 0.85g Bacto-Yeast Nitrogen Base w/o Amino Acids or $(NH_4)_2SO_4$ (Difco), 7g Cas-amino acids (Difco), 10g Bacto-agar and 400ml H_2O. Bacto-agar ommitted if broth required and 410ml H_2O added instead of 400ml. After autoclaving 22.9ml Adenine sulphate (1.2mg/ml), 30.5ml Tyrosine (1.8mg/ml) and 25ml 40% glucose (all sterile) added.

Yeast were grown at 30°C, with good aeration for liquid cultures. Yeast stocks were kept at -70°C in 20% glycerol.

2.7.2 Yeast Strains Used

AB1380 : MATα,ψ^+, ura3, trp1, ade2-1, can1-100, lys2-1, his5 (Burke et al., 1987).

YPH274 : a, ade2-101, ura3-52, trp1-Δ1, his3-Δ200, leu2-Δ1, lys2-1
α, ade2-101, ura3-52, trp1-Δ1, his3-Δ200, leu2-Δ1, lys2-1
(Heiter, unpublished).

2.7.3 Transformation of Yeast

The method used for transformation of yeast cells was that of Burgers and Percival (1987). 100ml of YPD was innoculated with a small loopful of yeast cells from a YPD plate and incubated overnight at 30°C shaking at 150 - 180 rpm. The cell concentration was checked in a haemocytometer and adjusted to 3 x 10^7/ml. This protocol is for 50 ml of cells at this concentration and all steps were performed at room temperature unless stated otherwise. The cells were pelleted at 600g for 5 mins, washed in 20ml H_2O, repelleted, washed in 20ml 1M Sorbitol and repelleted. The cells were then resuspended in 20ml SCE (1M Sorbitol, 0.1M Na Citrate, 10mM EDTA pH 5.8 with citric acid), 42µl β-mercaptoethanol and 1000u oxalyticase (30µl of a mg/ml solution) added and the cells incubated at 30°C with occasional inversion until spheroplasted (~15 min). After spheroplasting the cells were pelleted at 400g for 5 min and resuspended gently in 20ml 1M Sorbitol. The

cells were then repelleted, resuspended in 20ml STC (1M Sorbitol, 10mM Tris.HCl pH 7.5, 10mM $CaCl_2$), repelleted again and resuspended in 2ml STC. 0.1ml of the cells was added to DNA in a Falcon 2006 tube. For maximum transformation efficiency 1ng of plasmid and 5µg of carrier DNA in 10µl was used. After incubation for 10 mins 1ml PEG (20% PEG 8000, 10mM Tris.HCl pH 7.5, 10mM $CaCl_2$) was added and the incubation continued for another 10 mins. After pelleting the cells were resuspended gently in 150µl SOS (1M Sorbitol, 50% YPD broth, 6mM $CaCl_2$) and incubated for 30 - 40 mins at 30°C. 8ml of molten yeast selective agar was then added and the cells plated out onto selective plates. Once set the plates were wrapped and incubated at 30°C for 3 - 4 days.

2.8 EXTRACTION OF DNA FROM YEAST

2.8.1 Extraction of DNA for Cloning and Gel Electrophoresis

Solutions used are described in 2.7.3. 50ml of broth was innoculated with a small loopful of yeast cells and incubated overnight at 30°C shaking at 150 - 180 rpm. The cells were pelleted at 600g for 5 mins, washed in 20ml H_2O, repelleted, washed in 20ml 1M Sorbitol and repelleted. The cells were then resuspended in 10ml SCE, β-mercaptoethanol added to 100mM and zymolase added to 0.5mg/ml. The cells were incubated at 30°C with occasional inversion until spheroplasted to 90% (~15 min), pelleted at 400g for 5 min, washed in 1M Sorbitol, pelleted again and resuspended in 5ml 0.125M EDTA pH 8.0, 1% SDS, 200µl/ml Proteinase K and incubated for 2 hrs at 50°C. The DNA was extracted once with phenol, once with chloroform and precipitated with 0.6 vol propan-2-ol. The DNA was pelleted and resuspended in

500µl TE. RNAase was added to 50µg/ml and incubated for 30 mins at 37°C. To remove any residual RNA NaOAc pH 5.5 was added to 0.3M and 0.54 vol propan-2-ol was added dropwise with shaking between drops. The DNA was pelleted, washed in 70% ethanol and resuspended in 100µl TE.

2.8.2 Preparation of Chromosomes for Pulsed Field Gel Electrophoresis

The method used was that of Bellis *et al.* (1987). Cells from an overnight culture were harvested by centrifugation, washed twice in Solution 1 (50mM EDTA pH 8.0) and resuspended to a final concentration of 5 x 10^9 cells/ml in Solution 1. The cells were mixed with an equal volume of molten 1% low melting point agarose in Solution 1 (maintained at 55°C) and 100 µl aliquots were pipetted into a mould. When set the agarose plugs were transferred to Solution 2 (0.5M NaCl, 0.25M EDTA pH 8.0, 0.125M Tris.HCl pH 7.5, 0.5M β-mercaptoethanol) and incubated for 6 hrs at 37°C to obtain spheroplasts. The suspension was then made 1% for N-Lauroylsarcosine and Proteinase K was added (2mg/ml). After incubation for 24 hrs at 42°C the agarose plugs were stored at 4°C. Before use agarose plugs were washed 3 times in TE for 30 mins at 50°C, three times in TE for 30 mins at room temperature and 3 times in electrophoresis buffer for 30 mins at room temperature.

2.9 MANIPULATION OF DNA BY ENZYMES

2.9.1 Restriction endonuclease digestion

Digests were carried out according to the manufacturers' specifications using the A, B, L, M and H buffers provided by Boehringer Mannheim, except when a special buffer was required.

Genomic DNA digests were generally carried out overnight with 2 - 4 units of enzyme/µg DNA. If necessary spermidine at 5mM was added to digests. Reactions were stopped by heat inactivation or by extraction and precipitation of the DNA.

2.9.2 Dephosphorylation of DNA

To prevent unwanted ligation of certain DNA fragments, 5'-phosphate groups were removed by bacterial alkaline phosphatase (BAP). 1 unit of BAP was incubated with digested DNA at 65°C for 30 mins. To remove the BAP proteinase K was added to 100µg/ml and the incubation continued for 30 mins at 50°C. The DNA was extracted with phenol and then chloroform, precipiated, the pellet washed in 70% ethanol and the DNA resuspended in TE.

2.9.3 Ligation of DNA molecules

The enzyme ligase, isolated from E.coli infected with bacteriophage T4, catalyses the formation of a phosphodiester bond between 3'-OH and 5' phosphate groups of DNA (Weiss et al., 1968). This enzyme can therefore be used to join complementary cohesive termini of ds DNA molecules. Ligations were carried out in 50mM Tris-HCl pH7.4, 10mM $MgCl_2$, 1mM spermidine, 100µg/ml BSA, 1mM ATP, 10mM DTT. Reaction conditions depended on the type of cohesive ends being joined and followed the recommendations in the IBI catalogue (1986/1987).

2.9.4 Ligation of Linkers

Oligonucleotides were supplied after synthesis in an ammonium hydroxide solution. To prepare them for use the ammonium hydroxide solution was incubated overnight at 60°C and then transferred to -20°C for 10 mins. To precipitate the DNA 50µl of 2M NaOAc and 1.2ml of cold ethanol was added to 350µl of the solution and the mixture left for 60 mins at -20°C. The DNA was then pelleted by centrifugation and the pellet washed several times in 80% ethanol before resuspension of the DNA in TE. The DNA concentration was estimated by measuring the optical density (OD) at 260 nm (an OD of 1 at 260nm is equivalent to a DNA concentration of 20µg/ml).

Adaptors were inserted between DNA termini using the method described by Lathe et al. (1984).

2.9.5 Bal-31 Treatment of DNA

The exonuclease Bal-31 degrades linear duplex DNA from the termini, shortening both strands to yield progressively shorter duplexes without significant introduction of breaks away from the ends (Gray et al., 1975). Therefore if genomic DNA is treated with Bal-31 before digestion with restriction endonucleases the size of terminal restriction fragments would become shorter with increasing time of Bal-31 digestion. Bal-31 digestions were carried out as described in Maniatis et al. (1982).

2.9.6 Nested Deletions of Recombinant Plasmids

Blunt ends and those with 5' overhangs are susceptible to exonuclease III, whereas those with 3' overhangs are not. Therefore exonuclease III can be used to generate unidirectional deletions in ds

DNA if one end of the molecule has a 3' overhang and the other end has not (Henikoff, 1984). Treatment with exonuclease III results in progressive removal of nucleotides from one strand of the target DNA leaving a single-stranded region which can be removed using S1 nuclease. 'Nested' deletions are produced by removing aliquots at different times during exonuclease III digestion. The deleted molecules can be recircularised using ligase as described in 2.9.3 and transformed into XL1-Blue cells as described in 2.5.4. All nested deletions were done using the double-stranded nested deletion kit from Pharmacia.

2.10 SEPARATION OF DNA BY ELECTROPHORESIS

DNA fragments were separated by size by electrophoresis through agarose gels. Usually stop mix was added to DNA samples prior to loading. This stop mix acts as a visible marker and migrates in front of the DNA. After running gels were placed on a UV transilluminator and photographed using a Polaroid MP4 Land camera fitted with a red filter using Kodax TMAX Professional film 4052 with an exposure time of 15 secs. The film was developed in the automatic X-ray film processor.

2.10.1 Solutions and Buffers

50 x TAE

Per litre : 242g Tris base, 57.1ml glacial acetic acid, 100ml 0.5M EDTA (pH 8.0)

5 x TBE

Per litre : 54g Tris base, 27.5g boric acid, 20ml 0.5M EDTA (pH 8.0)

10 x stop mix

0.2M EDTA (pH 8.0), 15% ficoll and Orange G to a suitable colour

2.10.2 Conventional Agarose Gel Electrophoresis

Conventional horizontal agarose gel electrophoresis was used to separate DNA fragments of up to 50kb. Agarose concentrations of between 0.5% and 2% were used, depending on the size range of fragments to be resolved. Gels were made and run in 1 x TAE buffer containing ethidium bromide (EtBr) (0.5µg/ml). The gels were run at various voltages and for various times as required.

2.10.3 Pulsed-Field Gel Electrophoresis

Pulsed field gel electrophoresis was used to resolve DNA molecules which were greater than 50 kb. A 'waltzer' apparatus (Anand, 1986) was used. A 0.5cm thick 1% agarose gel was made and run in 0.25 x TBE at various 100 - 200V for 24 - 48 hours as required. The reorientation times of the gel were adjusted depending on the size of the fragments to be resolved. The buffer was recirculated through a cooling system in order to maintain it at a constant temperature of 10°C.

2.10.4 Polyacrylamide Gel Electrophoresis

Sequencing reactions were run on 5% polyacrylamide gels. The gels were cast and run in a Bio-rad sequencing apparatus which was assembled, and the gel poured, following the manufacturer's

recommendations. The glass plates were cleaned thoroughly before use and the bottom plate was siliconised. The glass plates used were 21 x 50cm and were separated by a wedge spacer which was 0.25mm thick at the top and 0.75mm thick at the bottom. The Sequagel gel casting solutions were used. For a 5% gel 20ml of Sequagel Concentrate was mixed with 70ml Sequagel Diluent and 10ml of Sequagel Buffer. 1ml of 25% (w/v) ammonium persulphate and 100µl of TEMED (NNN'N'-tetramethylethylethylenediamine) was added immediately before pouring the gel. After the gel had set (about 1 hour) it was pre-run in 1 x TBE for about 20 mins until the gel temperature had stabilised at 55°C. The gels were run at 1800V for the required time. After running the gels were fixed in 10% acetic acid, 10% ethanol for 30 mins to remove the urea before drying down and autoradiography.

2.10.5 Preparative Agarose Gel Electrophoresis

To isolate specific fragments of DNA for cloning or for radiolabelling DNA was separated on agarose gels were made and run as described in 2.10.2 except that low melting agarose was used. The gel slice containing the required DNA fragment was cut out under long-wave UV light and the DNA purified using the method of Burmeister and Lehrach (1989). The gel slice was melted at 65°C, EDTA pH 8.0 was added to 5mM and NaCl was added to 100mM. After 10 mins at 65°C the gel was equilibrated to 37°C and agarase added to 2U/100µl gel. After overnight incubation at 37°C to allow agarose digestion the DNA was phenol, then chloroform extracted, precipitated and resuspended in TE.

2.11 TRANSFER OF DNA TO NYLON MEMBRANES

2.11.1 Southern Transfer of DNA

DNA from agarose gels was transferred to reuseable nylon membranes (Hybond-N) using the method of Southern (1975).

UV illumination of the gel during photography introduces thymidine dimers into the DNA which increases the efficiency of transfer of large DNA molecules, however when transferring pulsed field gels the exposure time to UV was increased to 5 mins. After photography gels were denatured by shaking in denaturing solution (1.5M NaCl, 0.5M NaOH) for two 30 min periods and then neutralised by shaking in neutralising solution (3M NaCl 0.5M Tris.HCl pH 5.0) for a further two 30 min periods. The DNA was then capillary blotted onto nylon membrane with 20 x SSC (20 x SSC is 0.3M Tri-sodium citrate, 3M NaCl pH 7.0). The gel was placed onto Whatman 3MM paper pre-wet with 20 x SSC and a piece of Hybond-N pre-wet in 2 x SSC cut to the exact size of the gel was placed on top, care being taken to ensure that no air bubbles are trapped between the gel and the membrane. Two pieces of 3MM paper pre-wet in 2 x SSC were then placed on top followed by a stack of dry paper towels and finally a glass plate. After transfer for 4 hrs to 24 hrs the nylon membrane was washed in 2 x SSC and allowed to air dry. The DNA was covalently bound by exposure of the membrane to UV light for 5 mins with the DNA side opposed to the UV transilluminator.

Probe was removed from nylon membrane by washing in 0.4M NaOH for 30 mins at 45°C followed by neutralising in 0.2M Tris.HCl pH 7.5, 0.1 x SSC, 0.1% SDS for a further 30 mins.

2.11.2 Immobilisation of Bacterial Colonies

Hybond-N of the correct size was pre-wet by placing on a fresh agar plate and then carefully placed on the bacterial colonies. Orientation marks were made with a sterile needle and after 1 min the membrane was placed colony side up on a pool of denaturing solution and left for 7 mins before transferring to a pool of neutralising solution. After 3 mins it was transferred to a fresh pool of neutralising solution and left for a further 3 mins. The membrane was then washed in 2 x SSC, transferred to dry filter paper and allowed to air dry. The DNA was covalently bound to the membrane by exposure to UV light as described in 2.11.1.

2.11.3 Immobilisation of Yeast Colonies

The method used is based on that of Grunstein and Hogness (1975). Yeast colonies were grown on nylon filters placed on agar plates. Orientation marks were made using a sterile needle and the membrane was transferred colony side up onto 3MM paper soaked in SCE containing 100mM β-mercaptoethanol and 1mg/ml zymolase in a petri dish. After incubation for 3 hrs at 37°C the membrane was transferred to a pool of 0.5M NaOH and left for 7 mins. It was then transferred to a pool containing 0.5M Tris.HCl pH 7.5, 10 x SSC and left for 4 mins. This step was repeated and then the membrane was transferred to a pool of 2 x SSC for 2 mins. The membrane was allowed to air dry on dry filter paper after which the DNA was covalently bound to the membrane by exposure to UV light as described in 2.11.1.

2.12 RADIOLABELLING OF DNA

2.12.1 Random Priming

DNA probes were labelled by random priming using the method of Feinberg and Vogelstein (1983 and 1984). Random hexanucleotides are allowed to anneal to denatured probe DNA. They can then act as primers for synthesis of the complementary strand by Klenow enzyme from the 3'OH termini of the primers in the presence of ^{32}P-labelled dCTP and unlabelled dNTPs. Subsequent heat denaturation results in radiolabelled DNA ready for hybridization.

DNA for random priming was either in TE or in low gelling temperature agarose. The commercially available random priming kit from Boehringer Mannheim was used. 25ng DNA, boiled for 5 mins to denature the DNA, was added to a reaction mixture containing dATP, dGTP and dTTP (all at 25µM/l), random hexanucleotides and buffer. 2U of Klenow enzyme were added and 30µCi α-^{32}P dCTP (3000Ci/mmol, 10mCi/ml). The reaction volume was 20µl. After incubation at 37°C for 1-17 hrs the percentage incorporation of radioactive nucleotide into the DNA was estimated from the proportion of counts precipitated onto a Whatman GF/A filter by 10% TCA, which quantitatively precipitates oligonucleotides of >20 bases.

Proteins and unincorporated nucleotides were removed from radiolabelled probes to reduces background during hybridisation. This was done by passing the probe down a Sephadex G-50 (fine) column, DNA being excluded from the gel matrix (Maniatis et al., 1982).

2.12.2 End-labelling

The enzyme T4 polynucleotide kinase (PNK) can catalyzes the transfer of the γ- phosphate of ATP to a 5'OH terminus of DNA or RNA (Maniatis et al., 1982).

The reaction mix consisted of 20 - 30ng of oligonucleotide, 1 x PNK buffer (50mM Tris.HCl pH 7.6, 10mM $MgCl_2$, 5mM DTT, 0.1mM spermidine, 0.1mM EDTA pH 8.0), 100 µCi γ-^{32}P ATP and 20 units PNK in a total volume of 20µl. Incubation was at 37°C for 30 mins. The incorporation of counts was checked by TCA precipitation.

2.13 NUCLEIC ACID HYBRIDIZATION

2.13.1 Hybridization Solutions

Solution A

0.5M Na_2HPO_4 pH 7.2 with orthophosphoric acid, 7% SDS, 0.5% dried milk powder. Amended from the hybridization solution described in Church and Gilbert (1984). This solution was used for hybridizations using random primed DNA probes unless stated otherwise.

Solution B

0.5% SDS, 0.1% sodium pyrophosphate, 5 x SET/Denhardt's solution (20 x SET/Denhardt's is 0.4M Tris.HCl pH7.8, 3M NaCl, 20mM EDTA, 0.4% Ficoll, 0.4% polyvinylpyrrolidine (PVP) and 0.4% bovine serum albumin (BSA)). Sonicated salmon sperm DNA, denatured by boiling, was added at 100µg/ml as a competitor. This solution was used with certain random primed probes.

"Quick Hyb"

0.1% SDS, 0.1% sodium pyrophosphate, 0.05% BSA, 0.05% PVP, 0.05% Ficoll, 5 x SSC. This solution was used for hybridizations using radiolabelled DNA oligonucleotides.

2.13.2 Hybridization Protocols

Filters were hybridized in sealed plastic bags overnight in the appropriate solution containing radiolabelled probe. Filters were pre-hybridized for a minimun of 30 mins before the addition of probe. Pre-hybridization and hybridization was at 68°C for random primed DNA probes and at 50°C for oligonucleotide probes.

After hybridization filters were washed to remove non-specifically bound probe. When using random primed probes filters were washed 3 times at room temperature for 10 mins in 2 x SSC, 0.1% SDS followed by a stringent wash for 30 mins at 50°C - 65°C in 1 - 0.1 x SSC, 0.1% SDS depending on the degree of stringency required. When using oligonucleotide probes filters were washed 3 times at room temperature for 5 mins in 4 x SSC, 0.1% SDS followed by a stringent wash for 20 mins at 65°C in 4 x SSC, 0.1% SDS. After washing filters were placed in plastic bags ready for autoradiography.

2.14 AUTORADIOGRAPHY

Autoradiography was carried out using X-ray film (Fuji Medical or Kodak X-OMAT) in cassettes which contained intensifying screens except when exposing sequencing gels. Generally filters were exposed overnight at -70°C in the first instance and further exposures done for different lengths of time and/or at different temperatures as

required. Sequencing gels were exposed overnight at room temperature. The X-ray films were developed using a X1 X-OGRAPH automatic X-ray flim processor.

2.15 SEQUENCING

2.15.1 Sequencing Protocol

DNA was sequenced using the dideoxy chain termination reaction (Sanger et al., 1977) with the commercially available ^{T7}Sequencing kit (Pharmacia). DNA fragments to be sequenced were first cloned into M13 and ss template DNA made all as described in 2.6. The first step of the sequencing reaction involves annealing of a universal primer (17mer) to the ss M13 close to the cloning site followed by elongation from this primer by the enzyme T7 DNA polymerase and the second step involves chain termination. For each template four separate sequencing reactions were performed. All contained primer, template and all four dNTPs (one of them radiolabelled) but each included a different dideoxynucleotide (ddNTP). Incorporation of a ddNTP into the newly synthesised DNA strand results in chain termination as no 3'-OH group is available to form the next phosphodiester band. A series of DNA fragments is generated, all originating at the primer, and terminating at the sites of incorporation of the appropriate ddNTP. As one of the dNTPs is radiolabelled, the fragments can be resolved on polyacrylamide gels (Sanger and Coulson, 1978) and detected by autoradiography.

The annealing reaction contained 2µg of template M13 ssDNA, 0.8µM universal primer (17mer) and annealing buffer ($MgCl_2$, DTT) in a total volume of 14µl. This was mixed and centrifuged briefly before

incubation for 10 mins at 60°C followed by 10 mins incubation at room temperature. 3µl of labelling mix (dGTP, dCTP and dTTP in solution), 10µCi α-^{35}S dATPαS (>400Ci/mM) in 1µl and 3u T7 DNA polymerase in 2µl enzyme dilution buffer, were added and the reaction left for a further 5 mins. Four tubes labelled G, A, T and C containing 2.5µl of the appropriate termination mix was pre-incubated for 1 min before addition of 4.5µl of the labelling reaction to each tube. The components were mixed and incubated for 5 mins at 37°C. The reactions were stopped by the addition of 5µl stop solution (deionised formamide solution containing EDTA, xylene cyanol and bromophenol blue). The reactions were kept on ice until loading or they could be stored at -20°C.

Samples were heated at 80°C immediately prior to loading. 1.5µl was loaded into a freshly washed well using a micro-pipette. The gels were run and fixed as described in 2.10.4. The dried-down gels were autoradiographed as described in 2.14.

2.15.2 Sequence Analysis

All sequence analysis was carried out using the Amersham Staden-Plus suite of programs on a DCS IBM AT clone. Homology searches were carried out using the University of Wisconsin Genetics Computer Group's software package on the Daresbury computer.

CHAPTER 3

ORGANISATION OF THE PSEUDOAUTOSOMAL REGION TELOMERE

3 ORGANISATION OF THE PSEUDOAUTOSOMAL REGION TELOMERE

3.1 Introduction

As outlined in the Introduction the first DNA probe to be shown to be closely linked to a human telomere was 29C1. This probe defines the locus DXYS14 which is located within 20 kb of the telomere of the pseudoautosomal region of the X and Y chromosomes. In this chapter I will describe experiments designed to better define the terminal organisation of the pseudoautosomal region of the X and Y chromosomes from one individual. In addition I will describe experiments using 29C1 and linked probes to examine whether the human pseudoautosomal telomere undergoes any gross changes when in a mouse/human hybrid cell background.

In the human male lymphoblastoid cell line PES there are four blocks of homology to 29C1, two on each of the sex chromosomes within the terminal 20 kb. By screening a HindIII insert library, made from PES with 29C1 Dr Cooke isolated a clone, A35, which contains a block of homology to 29C1. By restriction mapping and Southern blotting this was shown to be the terminal block of homology on one of the sex chromosomes and the distal HindIII site was approximately 5 kb from the terminus. Although we knew that the clone A35 was from the distal part of the pseudoautosomal region we did not know which of the sex chromosomes it was from. It would be useful to be able to distinguish between the telomeres of the two sex chromosomes, especially for analysis and mapping. One way of doing this is to construct a mouse/human hybrid cell line which contains the X chromosome of PES but not the Y chromosome. By probing Southern blots of DNA made from

hybrids which have an intact X chromosome and have lost the Y chromosome with 29C1 the part of the restriction pattern which belongs to the X chromosome will be revealed.

Any mouse/human hybrids generated could then be used to investigate whether any gross changes or rearrangements occur at the telomere of the human X chromosome in the hybrid. These cell fusion mouse/human hybrids could be compared to hybrids generated by chromosome mediated gene transfer (CMGT) to see if the method used to create the hybrid has any effect on telomere structure.

3.2 Fusion of PES and X63/NS1

When mouse and human cells are fused, most of the human chromosomes are lost and only a few are relatively stably maintained (Ruddle and Creagan, 1975). By fusing an $Hprt^-$ mouse cell line, X63/NS1, with PES, which is $Hprt^+$, and selecting for cell survival in medium containing HAT and ouabain the only cells which can survive are mouse/human hybrids which contain the human X chromosome. The mouse cells are selected against because they cannot survive in HAT medium and the human cells are selected against because there is a naturally occurring difference between mouse and human cells with respect to resistance to ouabain. Concentrations of ouabain which are sufficient to kill human cells do not kill mouse cells.

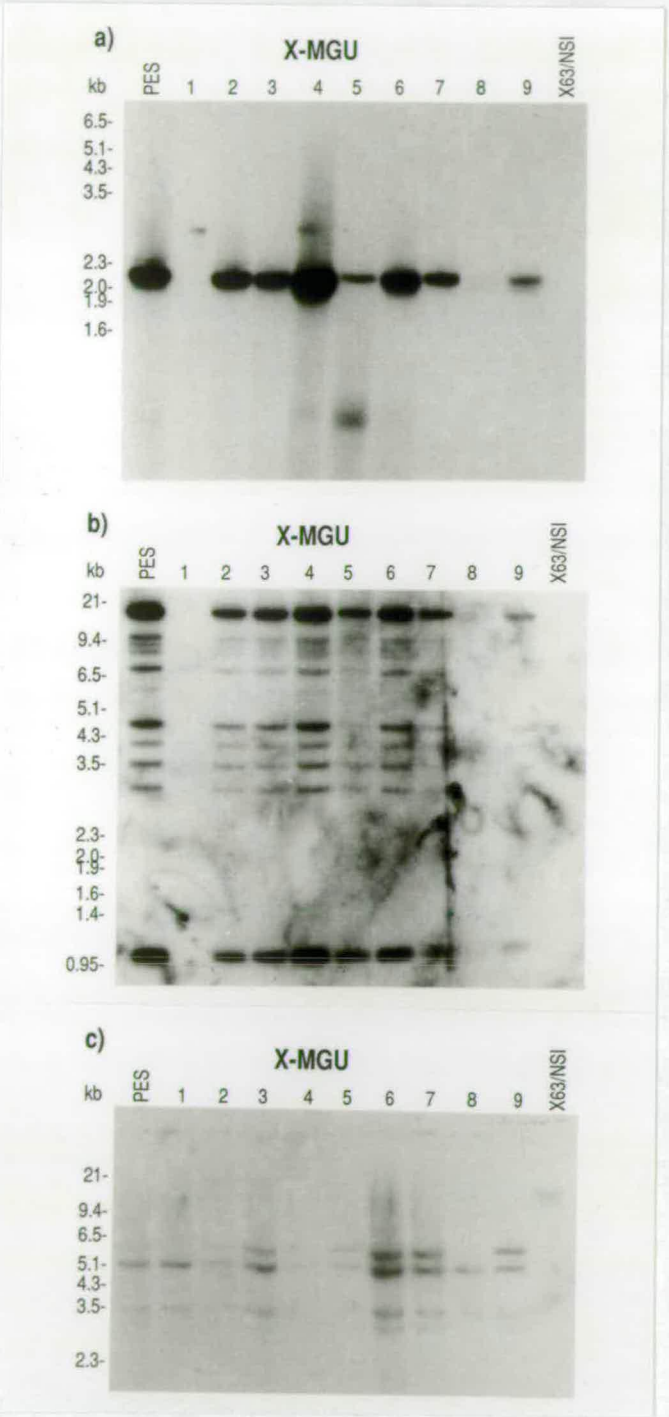
$5x10^6$ X63/NS1 cells were fused with 10^7 PES cells as described in Materials and Methods and they were initially put onto HAT selection only. The next day ouabain was added. Nine independent clones, X-MGU1 to 9 were obtained which were resistant to HAT and ouabain.

3.3 DNA Analysis of Clones X-MGU1 to 9

To determine if the human Y chromosome had been lost from the hybrids a Southern blot was prepared of HaeIII digested DNA from X-MGU1 to 9, PES and X63/NS1. This was hybridized with pHY2.1, a probe which detects a major satellite repeat present on the long arm of the Y chromosome and also cross-hybridizes slightly to the X chromosome (Cooke et al., 1983). A hybridizing band of 2.1 kb is characteristic of the Y chromosome. As can be seen from Figure 3.1a X-MGU 2,3,4,5,6 and 7 all have the Y chromosome pattern of hybridization whereas X-MGU1 and 8 do not. X-MGU9 has a faint Y specific band. To confirm this result another filter was prepared with HindIII digests of X-MGU1 to 9, PES and X63/NS1 and this was probed with the insert from Y-367. Y-367 is a 0.9 kb probe which detects Y-specific repeats, one 20 kb HindIII fragment derived from the proximal region of Yq and several other HindIII fragments from Yp (Müller et al., 1989). As can be seen from Figure 3.1b only X-MGU1 and 8 do not hybridize at all to Y-367, X-MGU9 gives a faint signal and the other hybrids hybridize as strongly as PES to Y-367. From these results it appears that only X-MGU1 and 8 have completely lost the Y chromosome and the other hybrids have not, although X-MGU9 is probably a mixed population of cells, most of which have lost the Y chromosome and some of which have not. The result of probing a filter of EcoR1 digests of PES, X63/NS1 and X-MGU1 to 9 with 29C1 can be seen in Figure 3.1c. As expected two hybridizing bands are seen with X-MGU1 and 8 and four hybridizing bands are seen with PES, X-MGU2,3,4,5,6 and with X-MGU 7 and 9 there are two strong and two weakly hybridizing bands. Presumably the two

Figure 3.1 Analysis of X-MGU Clones

10 µg of PES, X-MGU1 to 9 and X63/NS1 DNAs were digested with HaeIII in a), HindIII in b) and EcoRI in c). The digests were electrophoresed on three separate agarose gels and transferred to Hybond-N. The filters were hybridized with these ^{32}P-labelled probes : a) pHY2.1, b) Y-367 and c) 29C1. The hybridization buffer used for b) was solution B as described in Section 2.13.1. Washing conditions were 1 x SSC 0.1% SDS at 65°C for a) and 0.1 x SSC 0.1% SDS at 65°C for b) and c). Filters were exposed to XAR-5 film at -70°C for the appropriate time.



bands seen with X-MGU1 and 8 are derived from the X chromosome pseudoautosomal region and the extra two bands seen in PES and the other X-MGU hybrids derive from the Y chromosome.

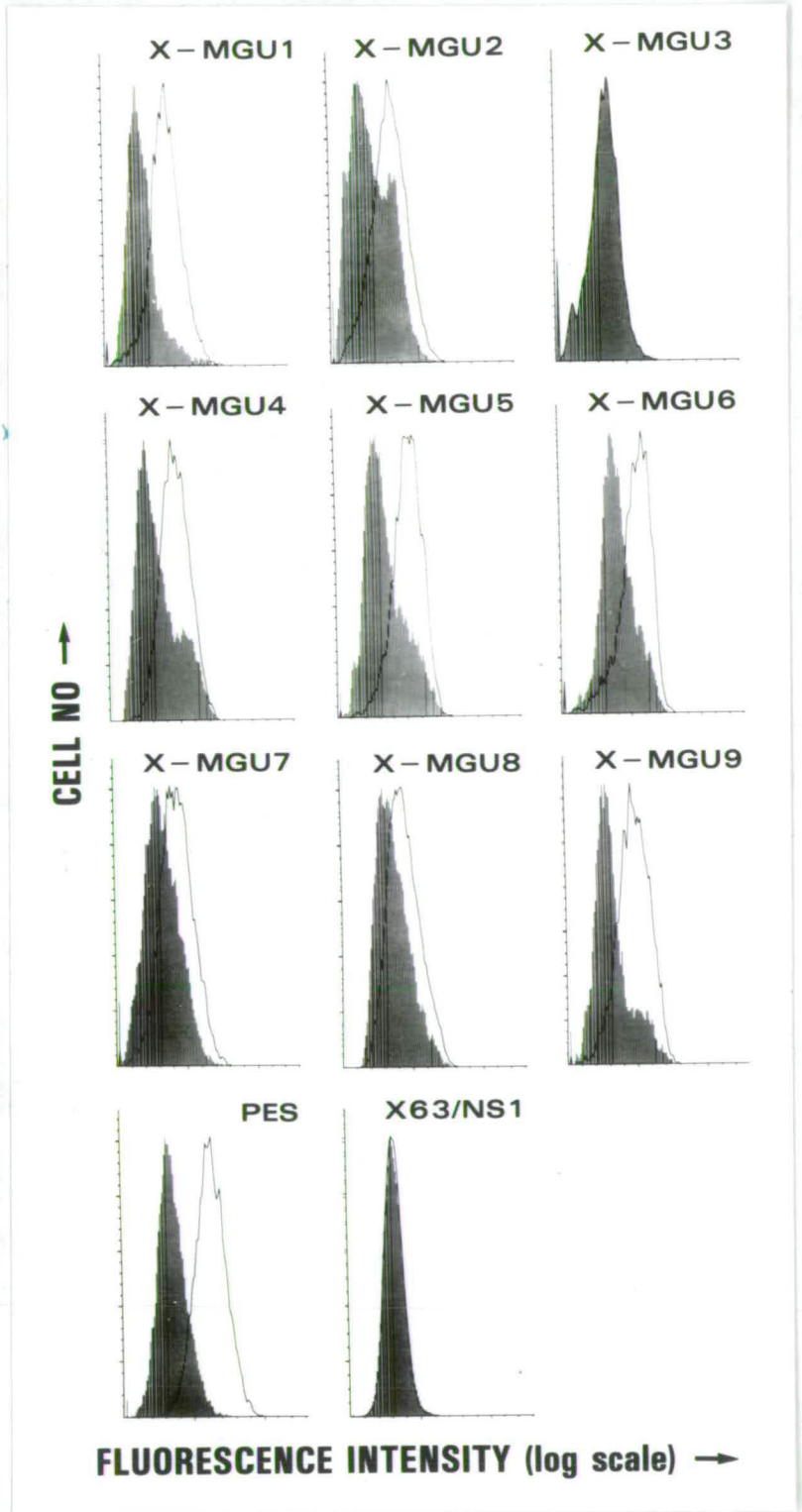
3.4 FACS Sorting the X-MGU Hybrids

Often the human chromosomes retained in mouse-human hybrids are deleted and rearranged and sometimes only fragments of human chromosomes are present which are integrated into mouse chromosomes. The region of interest was the terminal part of the pseudoautosomal region including the Xp telomere whereas the gene selected for by HAT, the HPRT gene, is located in the distal part of Xq. It was possible that the human X chromosomes present in the X-MGU hybrids might have rearranged in such a way as to retain a functional HPRT gene but have an altered Xp telomeric region. Therefore as one test for the integrity of the pseudoautosomal region the cells were FACS sorted for the presence of the cell-surface antigen 12E7. This gene is encoded for by the MIC2 gene which is located in the very proximal part of the pseudoautosomal region (Darling et al., 1986).

Figure 3.2 shows the results of FACS sorting the X-MGU hybrids, PES and X63/NS1. PES and all the X-MGU hybrids except X-MGU3 react with the 12E7 antibody and X63/NS1 is negative. The percentage positive counts for all the X-MGU hybrids is reduced in comparison to PES, perhaps indicating loss of the Y chromosome in a proportion of the cells or that the MIC2 is expressed to a lower level in a mouse/human hybrid than it is in a human lymphoblastoid cell line. The negative result with X-MGU3 is surprising as from blotting data this hybrid appears to contain intact human X and Y chromosomes. This

Figure 3.2 FACS Sort Analysis of PES, X-MGU Clones and X63/NS1

FACS sort profiles of the cell-lines PES, X-MGU 1 to 9 and X63/NS1. The vertical axis shows cell number and the horizontal axis shows relative fluorescence intensity (log scale). The shaded peak is the fluorescence profile with the negative control antibody and the open peak is the fluorescence profile for the MIC2 cell surface marker.



result needs to be confirmed, but if it is correct one explanation would be that the MIC2 genes have undergone some sort of rearrangement or modification in this hybrid so that they are no longer expressed in such a way that they can be detected by 12E7 antibody.

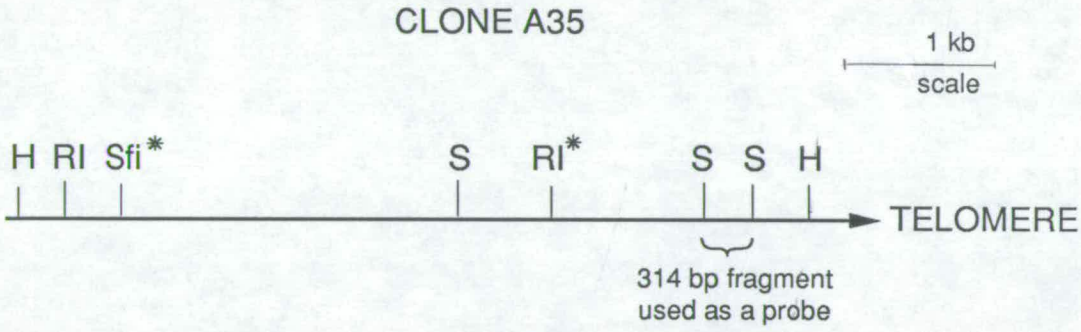
3.5 Probes for Defining the Map

The probe 29C1 detects two blocks of homology on both the X and the Y chromosome in PES. If only 29C1 is used for mapping purposes the only way to distinguish terminal from pro-terminal hybridizing fragments is by determining which fragment is lost first when intact DNA is treated with the exonuclease Bal-31 for different times prior to restriction digestion, Southern blotting and probing with 29C1. To provide a more straightforward alternative, fragments from clone A35 were sub-cloned and examined in order to isolate single-copy probes that detect only terminal restriction fragments.

A restriction map of clone A35 is shown in Figure 3.3. The distal part of this clone was examined for single-copy sequences. An EcoR1 site is located between the region containing homology to 29C1 and the HindIII site closest to the telomere. This 1.8 kb EcoR1 to HindIII fragment is repetitive, as judged by hybridization with total human DNA (data not shown), so is not an ideal probe. This EcoR1 to HindIII fragment was sub-cloned into the vector pTZ. Two SstI sites were mapped between the EcoR1 site and HindIII site, by end-labelling and restriction mapping A35, which split this fragment into three parts of 1.11 kb, 314 bp and 376 bp. The sub-clone was cut with HindIII, EcoR1 and SstI, electrophoresed on an agarose gel and transferred to Hybond-N. This filter was probed with ^{32}P-labelled total human DNA and washed

Figure 3.3 Restriction Map of Clone A35

Restriction map of clone A35 as deduced from double-digestion, end-labelling and partial digestion analysis.



KEY TO ENZYMES
B = Bam HI
H = Hind III
RI = EcoRI
S = SstI
Sfi = SfiI

A* above a site indicates that it is the final site before the telomere for this restriction enzyme.

at high stringency. The 314 bp SstI fragment was the only fragment which did not hybridize (data not shown) and so was the only likely candidate for use as a single-copy probe. This fragment was purified and the results of hybridizing it and 29C1 to filters with HindIII, SfiI and HindIII/SfiI digests of PES are shown in Figure 3.4. With SfiI, the most distal site for which is present between the two blocks of homology to 29C1, 29C1 detects two discrete fragments and a smear whereas the 314 bp SstI fragment detects only a smear. Also whereas 29C1 detects four HindIII fragments in PES the 314 bp SstI fragment detects only the two distal fragments. This fragment can therefore be used to differentiate between terminal and pro-terminal restriction fragments in the distal part of the pseudoautosomal region.

3.6 Restriction Map Terminal Part of the Pseudoautosomal Region of PES

X-MGU1 was selected for further study because from blotting and FACS data this hybrid appears to have lost the human Y chromosome and to have retained an intact X chromosome.

Southern blots were prepared of various digests of PES and X-MGU1. The results of probing these with 29C1 and/or the 314 bp SstI fragment derived from A35 are shown in Figure 3.5. Several conclusions can be made from examining these results. In HindIII digests both 29C1 and the 314 bp SstI fragment detect a band of 5.7 kb in PES which is absent in X-MGU1. Therefore the clone A35, which contains a 5.7 kb HindIII insert, is derived from the distal region of the Y chromosome in PES. In EcoR1 digests of PES the 314 bp SstI fragment

Figure 3.4 Comparison of 29C1 and the SstI Probe from A35

10µg of PES DNA was digested with HindIII, HindIII/SfiI and SfiI, electrophoresed on a 0.8% agarose gel and transferred to Hybond-N. The filter was first probed with ^{32}P-labelled 29C1, stripped and then probed with ^{32}P-labelled 314 bp SstI fragment from A35. Washing and autoradiography conditions were as described in Figure 3.1c.
H = HindIII and S = SfiI.

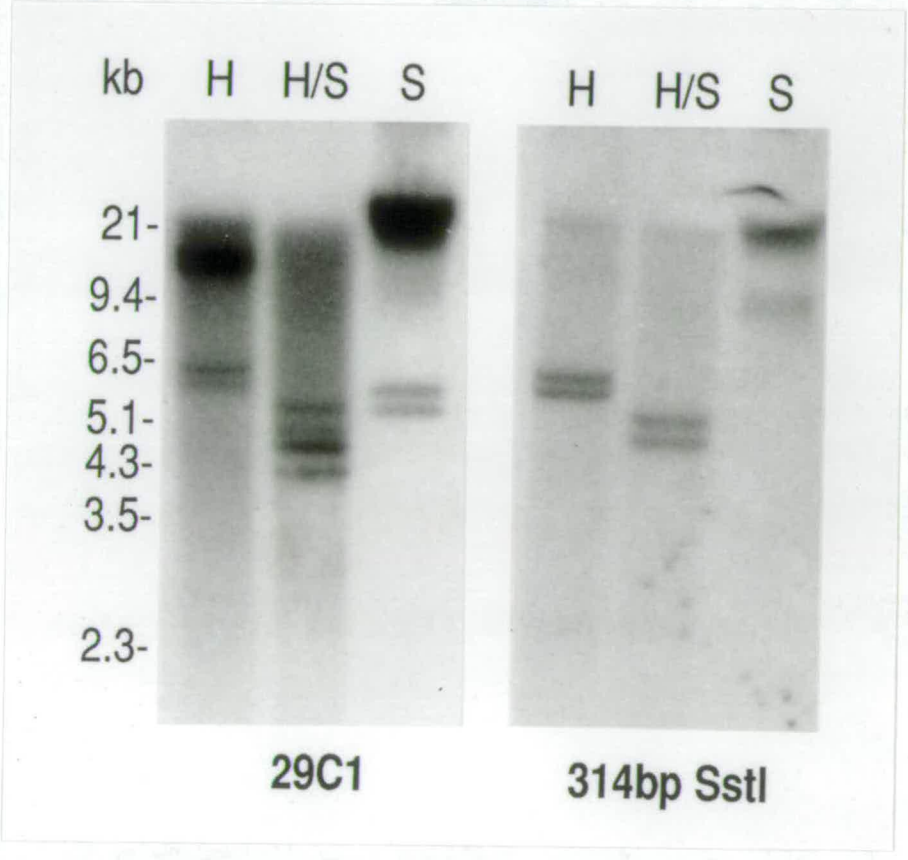
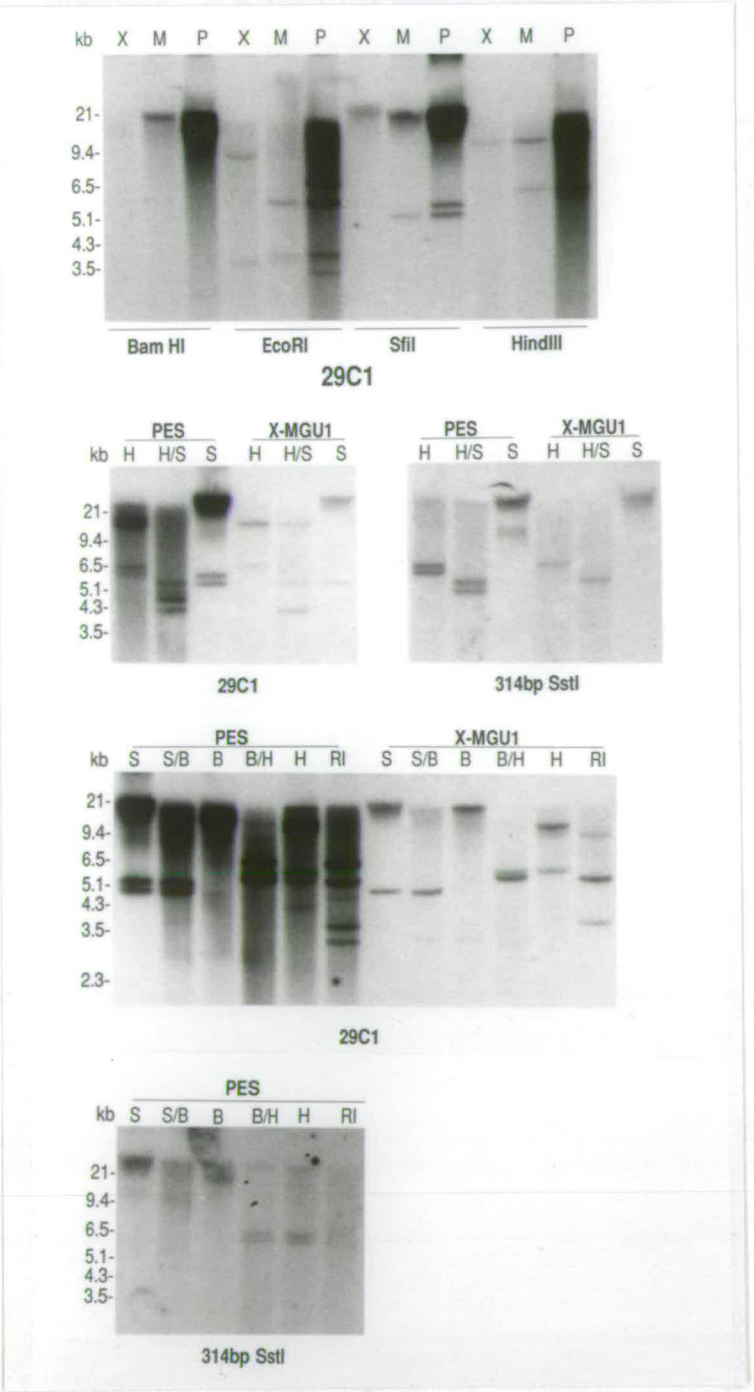


Figure 3.5 Hybridization Analysis of PES and X-MGU1

DNAs from the cell-lines PES, X-MGU1 and X63/NS1 were digested with restriction enzymes as shown, electrophoresed on 0.8% agarose gels and transferred to Hybond-N. The filters were hybridized with 29C1 and/or the 314 bp SstI fragment from A35 as indicated. Conditions used as described in Figure 3.4. Key to restriction enzymes B = BamHI, H = HindIII, RI = EcoRI and S = SfiI. Key to cell-lines M = X-MGU1, P = PES and X = X63/NS1.



detects a smear whereas 29C1 detects two discrete bands. This indicates that the 314bp SstI probe is located on the telomeric side of the final EcoR1 site before the telomere.

Restriction maps of the X and Y pseudoautosomal telomeres of PES derived from the results of Figure 3.5 are shown in Figure 3.6. As can be seen the maps are very similar with two blocks of homology to 29C1 located quite close together on both chromosomes. These blocks are probably part of a larger region of homology generated by duplication (Inglehearn, pers. comm.), although this duplicated region cannot extend far on the telomeric side of the 29C1 homology as the 314 bp SstI fragment is not duplicated.

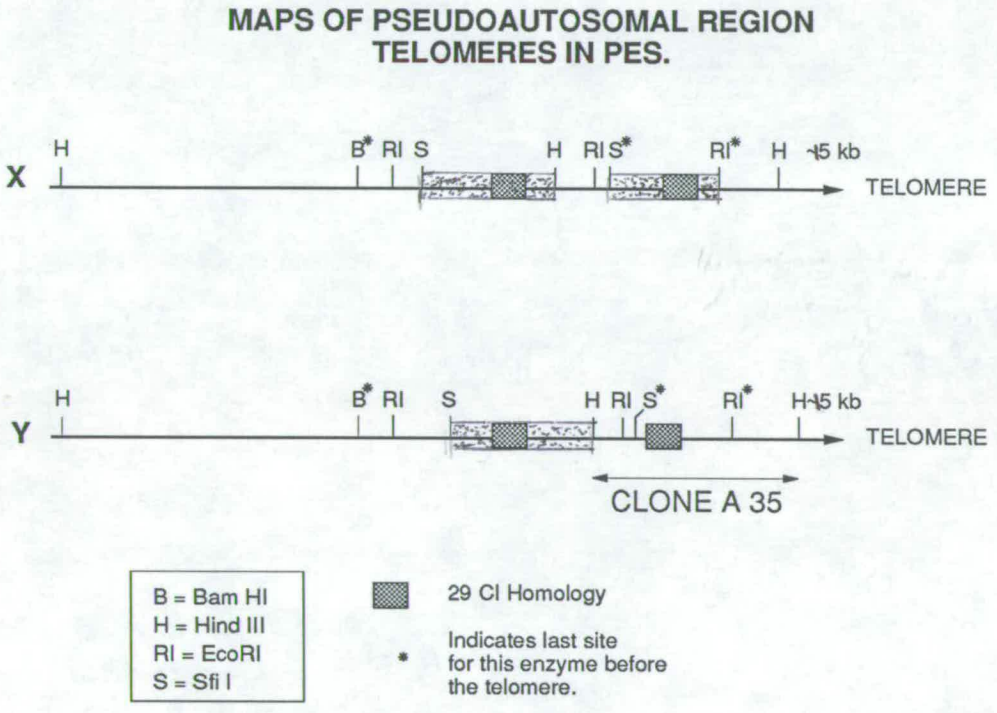
3.7 Changes in Human Telomere On a Mouse Background

When linear chromosomes with Tetrahymena telomeres are introduced into yeast cells the terminal array of Tetrahymena repeats is shortened and then extended by yeast telomeric repeats (Shampay et al. 1984). The X-MGU hybrids and CMGT generated hybrids (see below) were used to investigate whether changes of a similar sort occur at the X pseudoautosomal telomere in mouse/human hybrids as a consequence of being on a mouse background.

From the results presented in Figure 3.5 it can be seen that no differences can be detected by the probes 29C1 and the 314 bp SstI fragment from A35 between the human X pseudoautosomal telomere in PES and any of the X-MGU hybrids.

In addition to cell fusion human chromosomes can be introduced into a mouse cell background by CMGT. In this method isolated human chromosomes arrested at mitosis are co-precipitated with $CaPO_4$ and

Figure 3.6


MAPS OF PSEUDOAUTOSOMAL REGION
TELOMERES IN PES.
X
H
B* RI S
H
RI S*
RI*
H
~15 kb
TELOMERE
Y
H
B* RI
S
H RI S*
RI*
H
~15 kb
TELOMERE
CLONE A 35
B = Bam HI
H = Hind III
RI = EcoRI
S = Sfi I
29 Cl Homology
*
Indicates last site
for this enzyme before
the telomere.

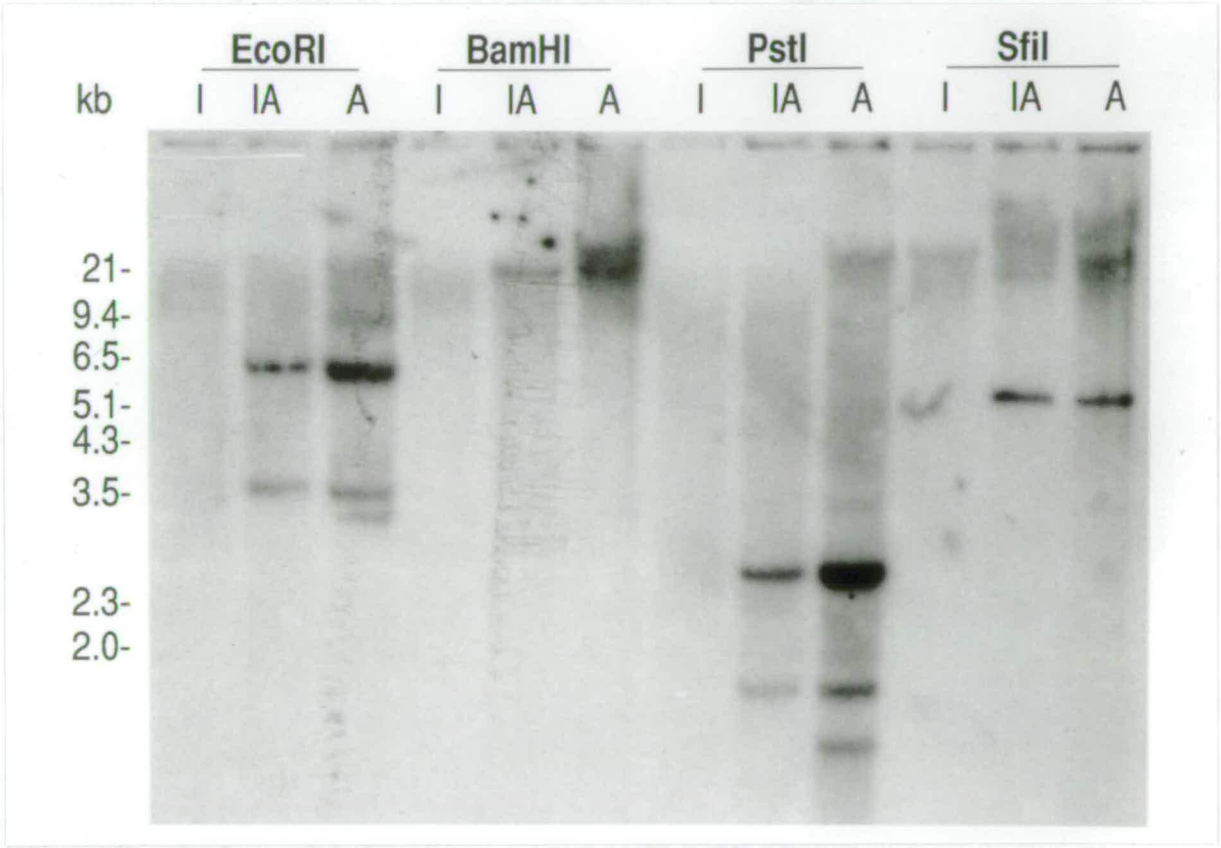
applied to a monolayer of mouse cells. Fragments of the human chromosomes transform the mouse cells and those fragments carrying selectable markers can be retained (reviewed in Porteous, 1987). In the resulting transgenomes the human complement is often rearranged and frequently integrates into mouse chromosomes. However markers adjacent to the selectable marker can be retained in an unrearranged manner and sometimes almost intact chromosomes are stably maintained.

Dr Porteous provided DNA samples of a mouse/human hybrid cell line, IR/AKLO, which contains a single human X chromosome introduced by transfection into a mouse cell line and the parental cell-lines used to generate this hybrid which were the $Hprt^-$ mouse cell-line IR-3D and the $Hprt^+$ human cell-line AKLO. The human X chromosome is retained in the hybrid by selection on HAT medium. The human X chromosome in this hybrid has lost about 1/3 of its material, probably by interstital deletions as judged by probing blots with IR repeats, and no mouse DNA can be detected in the chromosome by in situ hybridisation (Porteous, pers. comm.).

Figure 3.7 shows the results of probing a Southern blot of various restriction digests of this hybrid cell-line and its mouse and human parental cell-lines with 29C1. There are no detectable differences between the pseudoautosomal telomere of the X chromosome in the human parental cell line, AKLO, and that in the mouse/human hybrid cell line, IR/AKLO. With those enzymes whose most telomeric site is on the proximal side of 29C1 the terminal restriction fragments detected in AKLO and IR/AKLO cover the same size range. However in all cases the fragments detected are large so it would be difficult to see any small size changes. It is also possible that any mouse DNA added contains a site for the restriction enzymes used.

Figure 3.7 Hybridization Analysis of a CMGT Mouse/Human Hybrid

DNAs from the cell-lines IR-3D, IR/AKLO and AKLO were digested by restriction enzymes as shown below, electrophoresed on a 0.8% agarose gel and transferred to Hybond-N. The filter was hybridized with 29C1 as described in Figure 3.4. Key to cell-lines I = IR-3D, IA = IR/AKLO, A = AKLO.



Telomere length in cell-lines is altered in comparison to that in tissue DNA . The 29C1 hybridization pattern of BamHI digests of PES compared to those of blood and sperm from the individual from whom PES was made show that the terminal restriction fragments in the lymphoblastoid cell line are shorter and more homogeneous than those in blood and sperm (Cooke and Smith, 1986). It may be that telomeres in cell-lines reach a base-line length which is species-specific but is independent of the cellular origin of the cell-line.

3.8 Discussion

The deduced restriction maps of the distal portion of the X and Y pseudoautosomal regions of PES show that although they are very similar in the overall organisation of restriction sites there are differences in the sizes of restriction fragments. This would be expected as the locus DXYS14 recombines with sex at a frequency of 50% in male meiosis so that any major changes occurring in this highly polymorphic region would be likely to spread throughout the population.

The locus DXYS14 is part of a tandem duplication which is present both on the X and Y chromosomes of PES. The distal part of the clone A35 is not part of this duplication as it only detects terminal restriction fragments. The orientation of restriction fragments with respect to the telomere appears to the same on both chromosomes so the duplication does not involve an inversion.

Clone A35 consists largely of very repetitive sequences, shown when restriction digests of A35 were probed by total human DNA and most fragments were detected (data not shown). Whether the repetitive

nature of sub-telomeric sequence is the same on other chromosomes will be known when telomere-adjacent sequences from other telomeres have been isolated and characterised (see Chapter 7).

When mouse and human cells are fused the mouse chromosomes are retained and the human chromosomes tend to be lost unless they are selected for. From the results described in this section it seems unlikely that the reason for the observed instability of human chromosomes on a mouse cell background is due to a failure of the human telomere to function in the mouse cell. The human X pseudoautosomal telomere appears to be stable and has not undergone gross rearrangements when introduced onto a mouse cell background by the two methods of chromosome transfer examined here. For these reasons it is probably functional.

CHAPTER 4

CROSS-SPECIES CONSERVATION OF TELOMERIC SEQUENCE REPEATS

4 CROSS-SPECIES CONSERVATION OF TELOMERIC SEQUENCE REPEATS

4.1 Introduction

As discussed in Sections 1.12 and 1.13, human telomeres appear to share many characteristics with those of lower eukaryotes in that terminal restriction fragments are heterogeneous and the terminus appears to be a 3' overhang (Cooke et al., 1985; Cooke and Smith, 1986; Cooke, 1987). These results strongly suggest that arrays of repeats are present at the ends of human chromosomes. The sequence composition of these repeats will be investigated in this chapter.

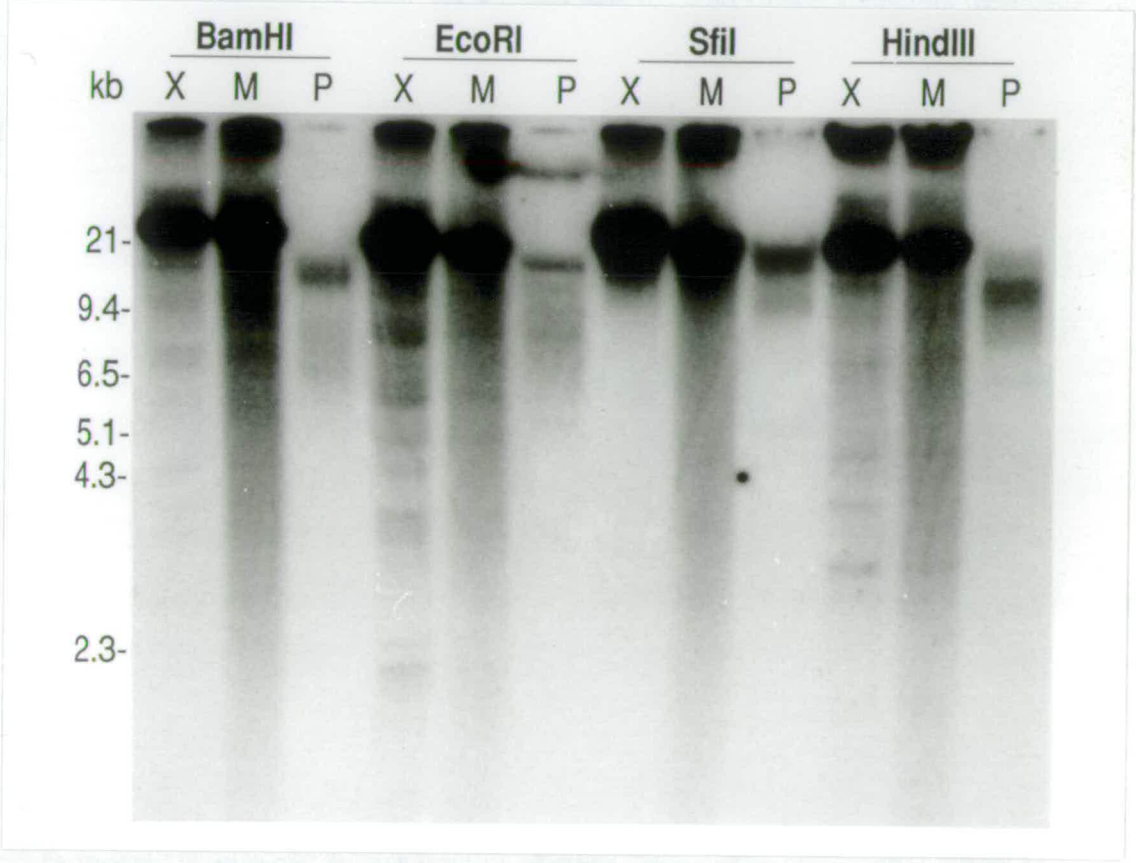
In the course of other experiments Dr Allshire hybridized a filter of mouse DNA digests with pSPT16 (Sugawara and Szostak, 1986), a probe which contains arrays of both the Tetrahymena telomeric repeat and the S. pombe telomeric repeat. pSPT16 detected heterogeneous bands in all restriction digests. As heterogeneity of terminal restriction fragments is a feature of telomeres it seemed possible that at least some part of this probe was cross-hybridizing to the mouse telomeres. In this chapter I will describe experiments which confirmed this result and extended the analysis to human telomeres.

4.2 pSPT16 Cross-hybridizes to Human Telomeres

To establish if pSPT16 cross-hybridizes to human DNA in addition to mouse DNA pSPT16 was hybridized to a filter of various digests of a mouse cell-line, X63/NS1, a human cell-line, PES, and a mouse/human hybrid cell-line, X-MGU1 derived from these two cell-lines (all described in Chapter 3). Figure 4.1 shows the results of this

Figure 4.1 Cross-hybridization of pSPT16 to Human DNA

X63/NS1, X-MGU1 and PES DNAs were digested with the restriction enzymes as shown, 10 µg samples were electrophoresed on a 0.8% agarose gel and transferred to Hybond-N. The filter was hybridized with ^{32}P-labelled BamHI/HindIII fragment from pSPT16. The filter was washed at 0.1 x SSC 0.1% SDS at 65°C and exposed to XAR-5 film.
Key to cell-lines, X = X63/NS1, M = X-MGU1, P = PES.



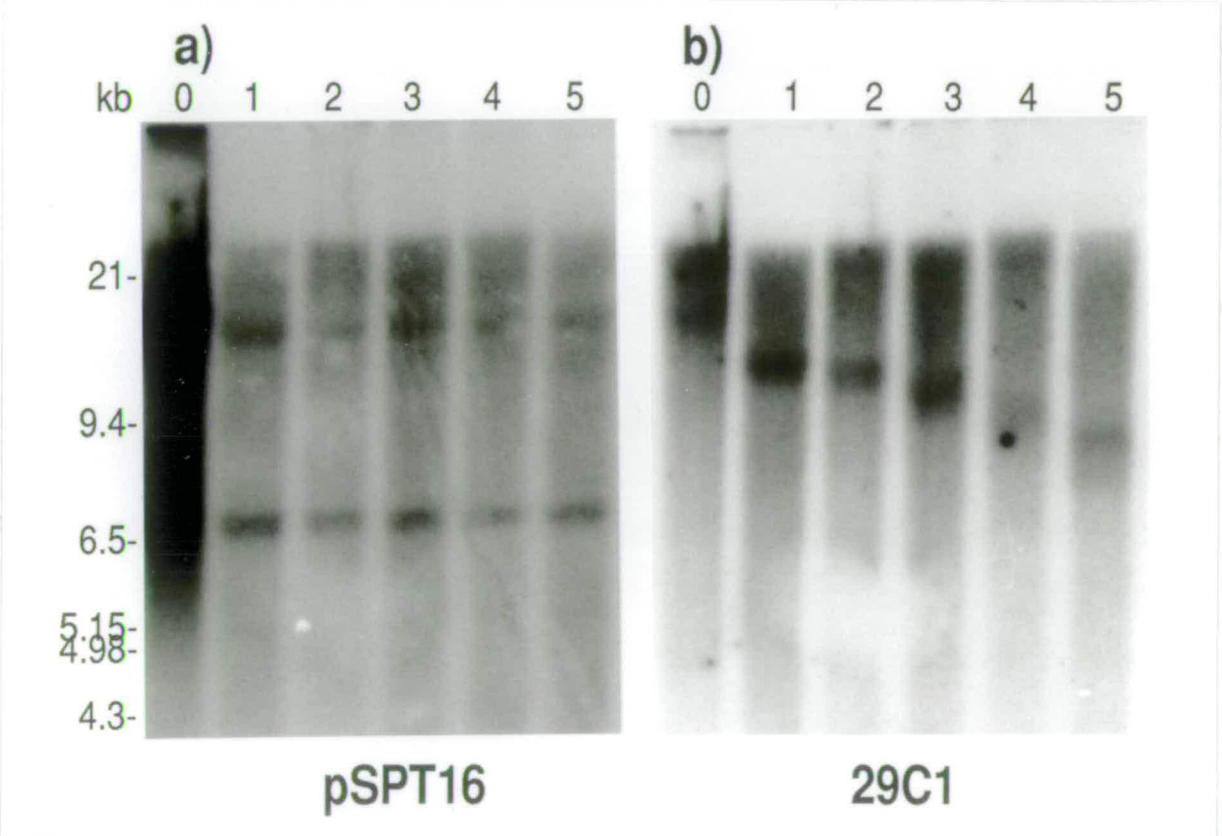
experiment. A heterogeneous smear of hybridization, in addition to discrete bands, is detected in all the cell-lines and in all the digests of those cell-lines. Therefore pSPT16 cross-hybridizes to human DNA as well as mouse DNA in a pattern consistent with it cross-hybridizing to telomeric sequences.

The next objective was to determine whether any of the cross-hybridizing (heterogeneous or discrete) human fragments were indeed telomeric. Terminal restriction fragments from human DNA are sensitive to the exonuclease Bal-31 as has been demonstrated with the telomere-adjacent probe 29C1 (Cooke et al., 1985). Figure 4.2a shows the result of hybridizing pSPT16 to PES DNA that had been treated for increasing lengths of time with Bal-31 prior to digestion with the restriction endonuclease BamHI. The heterogeneous smear of hybridization detected by pSPT16 is not apparent after one hour of Bal-31 digestion. In contrast the discrete hybridizing bands are largely unchanged even after five hours of Bal-31 digestion. Figure 4.2b shows the result of reprobing the filter with 29C1. Hybridizing fragments become smaller with increasing time of Bal-31 digestion demonstrating that the enzyme remains active even after several hours of incubation. That the heterogeneous smear of hybridization to pSPT16 is sensitive to Bal-31 must mean that these cross-hybridizing sequences are located at, or close to telomeres. However the discrete bands are not immediately telomeric. They may still be sub-telomeric but are not degraded by Bal-31 in this experiment.

The above results demonstrate that pSPT16 cross-hybridizes to human telomeres. This conclusion was supported by in situ experiments carried out by Dr Allshire and Dr Gosden, (Allshire et al., 1988, see

Figure 4.2 pSPT16 Detects Bal-31 Sensitive Restriction Fragments

PES DNA was digested with Bal-31 for the times indicated (hours). The samples were then digested with the restriction enzyme BamHI, separated by electrophoresis on a 0.7% agarose gel and transferred to Hybond-N. The filter was first hybridized with pSPT16 shown in a), stripped and then hybridized with 29C1 shown in b). Washing conditions were as described in Figure 4.1.



Appendix). These showed that when pSPT16 was hybridized to metaphase spreads of human chromosomes the majority of the grains were located over telomeric regions of most chromosomes.

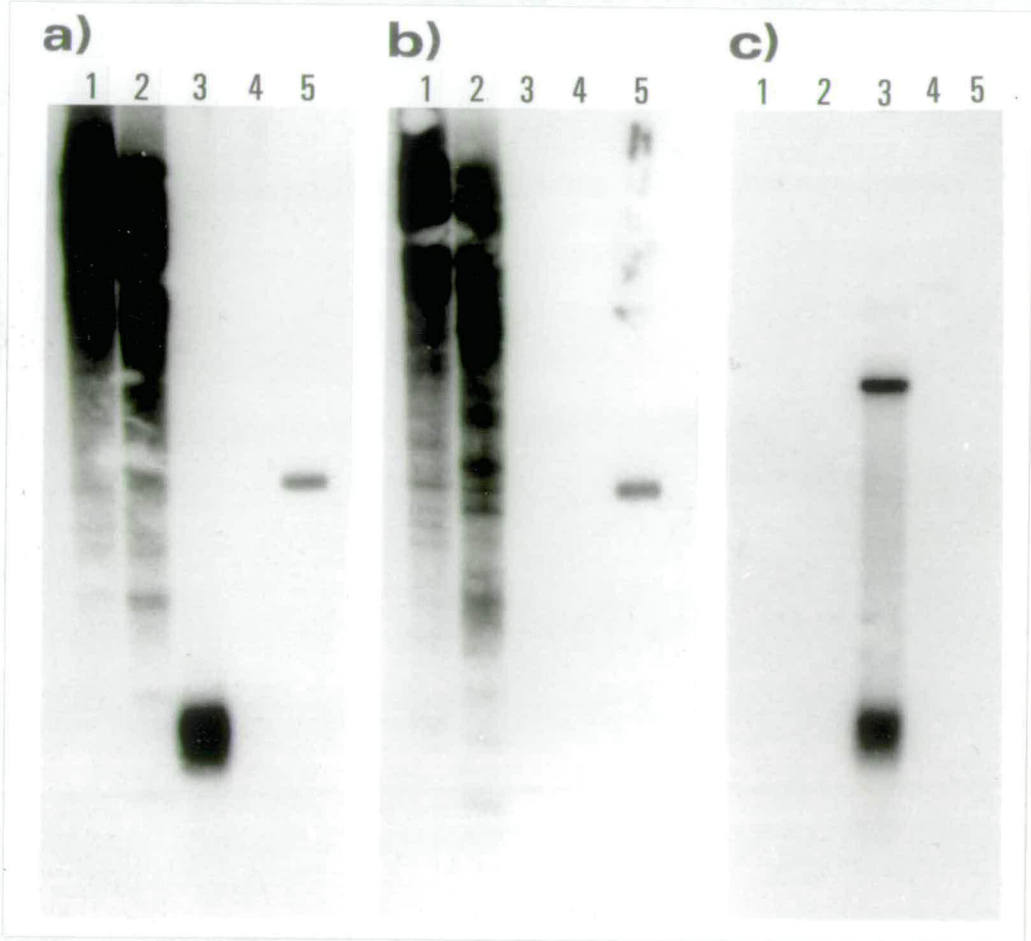
4.3 Tetrahymena Telomeric Repeat Cross-Hybridizes to Human Telomeres

The pSPT16 probe consists of ~100 bp of the Tetrahymena telomeric repeat and ~300bp of the S. pombe telomeric repeat (Sugawara and Szostak, 1986). It was important to establish if the cross-hybridization to human telomeres observed with pSPT16 is due to the Tetrahymena or the S.pombe repeat. Figure 4.3 shows the results of hybridizing a filter of human, mouse, S.cerevisiae, S. pombe and Tetrahymena DNA digests with pSPT16, a Tetrahymena telomeric probe and a S. pombe telomeric probe. The filter was prepared by Dr Allshire who also carried out the S. pombe hybridization. The S. pombe, S. cerevisiae and Tetrahymena telomeric sequences did not cross-hybridize to each other. This implies that the cross-hybridization seen with pSPT16 to human telomeres reflects sequence identity rather than an overall G-richness of the probe. The Tetrahymena telomeric repeat probe cross-hybridizes to human and mouse telomeres but the S. pombe telomeric repeat probe does not. From this result it is clear that it is the simple repeat Tetrahymena telomeric repeat TTGGGG which is responsible for the cross-hybridization of pSPT16 to human telomeres.

In order to estimate the position and extent of the telomeric sequences cross-hybridizing to the Tetrahymena telomeric repeat PES DNA was treated for increasing amounts of time with Bal-31 prior to digestion by the restriction endonuclease BamHI. A filter of these digests was first hybridized with the pYAC4 fragment containing the

Figure 4.3 Cross-Hybridization Analysis of Telomeric Sequences

DNA samples as described below were elecrophoresed on three separate 0.8% agarose gels and transferred to Hybond-N. Lane 1, 10 µg human placental DNA; lane 2, 10 µg mouse C_{127} cells; lane 3, 300 ng *S. pombe* ED628; lane 4, 300 ng *S. cerevisiae* DC5; lane 5 ~80 pg purified telomeric 3.5kb BglII fragment from *T. thermophila* rDNA. DNAs in 1-4 were digested with EcoRI. The filters were probed with the following ^{32}P-labelled probes; a) pSPT16 (fragment described in Fig 4.1), b) BamHI/XhoI fragment from pYAC4 (Burke *et al.*, 1987) containing 700 bp *Tetrahymena* telomeric repeat, c) SacI/HindIII fragment from pNSU68 containing 200 bp *S. pombe* telomeric repeat. Filters in a) and b) were washed as described in Fig 4.1, filter in c) was washed at 1 x SSC, 0.1% SDS at 68°C.



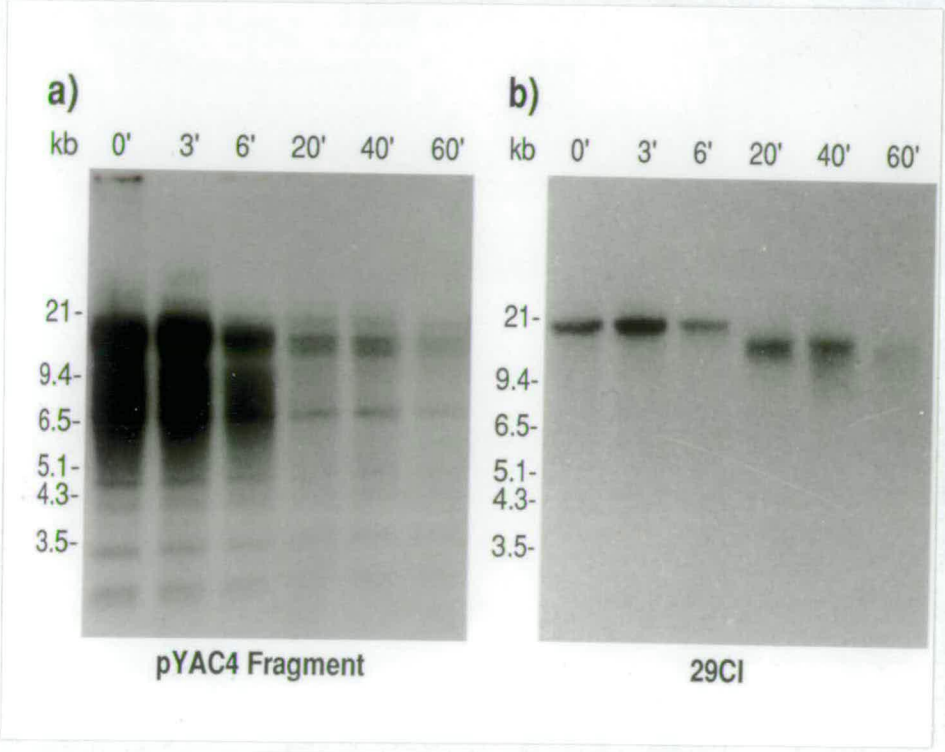
TTGGGG repeat described in Figure 4.3, stripped and then rehybridized with 29C1. As 29C1 detects sequences located within the terminal 20 kb of the pseudoautosomal telomere of PES (see Figure 3.6) an estimate of the rate of Bal-31 digestion can be made from the results of this hybridization. The results of these hybridizations are shown in Figure 4.4a and b. In this experiment the rate of removal of nucleotides was ~10 kb per hour. As the smear detected by Tetrahymena telomeric probe has disappeared completely by 20 mins, and had been reduced by 6 mins, the sequences detected may be within 2.5 - 3.0 kb from the terminus, although this is a gross estimate as the rate of Bal-31 digestion is not constant and depends on base composition.

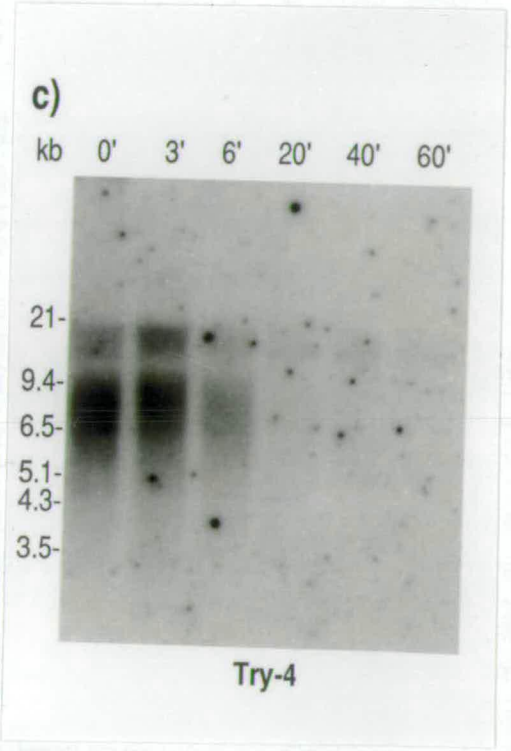
4.4 Trypanosoma Telomeric Repeat Cross-Hybridizes to Human Telomeres

Other workers have shown that other repeat sequences cross-hybridize to human telomeres. A human repetitive sequence was shown to cross-hybridize strongly to most if not all human telomeres (Moyzis et al., 1988). The sequence of this repeat is TTAGGG which is identical to the telomeric repeat of Trypanosoma. In addition the Arabidopsis telomeric repeat TTTAGGG (Richards and Ausubel, 1988) was shown to cross-hybridize to human telomeres. As the cross-hybridization of the Trypanosoma telomeric repeat appeared to be as strong, if not stronger, than the cross-hybridization of the Tetrahymena telomeric repeat to human telomeres an oligonucleotide, $(TTAGGG)_4$ (a gift from Dr Greider), which contains four copies of the G-rich strand telomeric repeat of Trypanosoma was used as a probe in subsequent experiments. It will be referred to as Try-4 in the rest of this thesis. The hybridization and washing conditions used with this oligonucleotide were determined empirically.

Figure 4.4 Extent of Telomeric Sequence Cross-Hybridization

PES DNA was digested with Bal-31 for the times indicated (mins). The samples were then digested with the restriction enzyme BamHI, 10µg aliquots separated by electrophoresis on a 0.8% agarose gel and then transferred to Hybond-N. The filter was hybridized with a) ^{32}P-labelled fragment from pYAC4 containing the Tetrahymena telomeric repeat (described in Figure 4.3), stripped and then hybridized with b) ^{32}P-labelled 29C1. The conditions used are described in Figure 4.1. The filter was stripped again and hybridized with c) ^{32}P-labelled oligonucleotide Try-4. The hybridization conditions used are described in Materials and Methods. The filter was washed at 4 x SSC 0.1% SDS at 65°C.





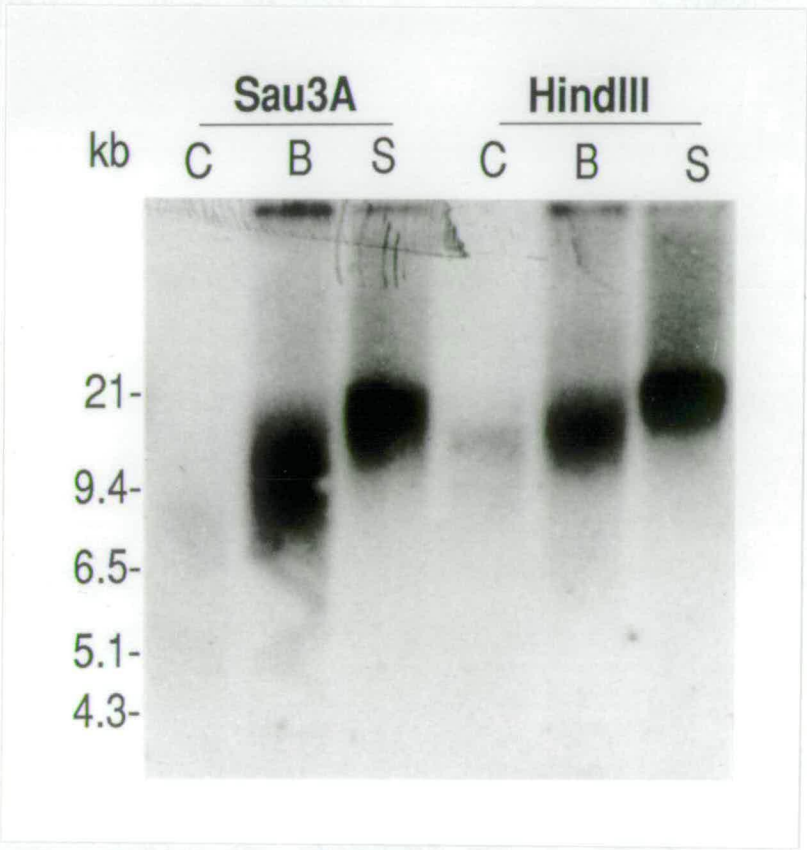
The result of hybridizing Try-4 to the filter of the PES Bal-31 series used in Figure 4.4a and b is shown in Figure 4.4c. The TTAGGG sequence hybridizes very strongly to the Bal-31 sensitive heterogeneous telomeric smear and no discrete hybridizing fragments are apparant. It seems likely that the majority of the sequences cross-hybridizing to Try-4 are located at the telomeres.

The Try-4 probe was used to investigate the differences in telomere length that had been previously observed between pseudoautosomal telomeres from blood, sperm and lymphoblastoid cell lines (Cooke and Smith, 1986). Figure 4.5 shows the result of hybridizing Try-4 to a filter of various digests of the cell-line PES and blood and sperm DNAs from the individual from whom PES was made. The PES telomeric smear is less intense, more homogeneous and smaller than that of either blood or sperm. In addition the sperm DNA telomeric smear is larger than the blood DNA telomeric smear. The sperm telomeric smear also appears to be more homogeneous but this could just be due to the fact that larger DNA fragments are not well resolved on conventional agarose gels. This result suggests that the differences between blood and sperm telomeres previously observed are not confined to the pseudoautosomal telomere but apply to all human telomeres.

Figure 4.6a and b shows the result of hybridizing Try-4 to a filter of various digests of blood and sperm DNAs from the same individual. Figure 4.6a is a short exposure which shows that the TTAGGG sequence hybridizes to large heterogeneous fragments in all digests and that the size of these fragments is larger in the sperm DNA digests than the equivalent blood DNA digests. Terminal restriction fragments derived from blood DNA are between 5 and 12 kb

Figure 4.5 Hybridization of Try-4 to PES, Blood and Sperm DNA Digests

PES DNA, and blood and sperm DNAs from the individual from whom PES was made, were digested with restriction enzymes as shown. The digests were separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The filter was hybridized with Try-4 as described in Figure 4.4c. Key to lanes C = PES cell-line, B = blood and S = sperm.



and those derived from sperm DNA are between 12 and 21 kb. The size of the terminal restriction fragments appears similar in all digests irrespective of whether a restriction enzyme with a 6 bp or a 4 bp recognition sequence is used. An extensive region at telomeres appears "barren" for restriction endeonuclease sites. The simplest explanation for this is that telomeric and immediately sub-telomeric regions consist of simple repeat sequences that do not contain restriction endonuclease recognition sites. Figure 4.6b is a longer exposure of 4.6a. In this exposure discrete hybridizing bands can be seen in addition to the heterogeneous smear in all digests. This indicates that arrays of TTAGGG repeats are not confined to telomeres but are present at internal locations as well. The exact chromosomal position of these internal arrays is unknown at present (see below) but as they only become apparent after long exposure times they are likely to be short and perhaps they may diverge to some extent from the TTAGGG sequence. Figure 4.6c shows the result of hybridizing the filter to an oligonucleotide $(TTGGGG)_4$, named Tet-4 (synthesised by John Inglis), which is equivalent to four repeats of the G-rich strand of the Tetrahymena telomeric repeat. Heterogeneous and discrete restriction fragments are detected. If the results shown in Figure 4.6b and c are compared it is apparent that both Tet-4 and Try-4 hybridize to high molecular weight, heterogeneous restriction fragments but the discrete restriction fragments hybridizing to the TTAGGG sequence are different from those hybridizing to the TTGGGG sequence. In order to determine the extent of cross-hybridization between Try-4 and Tet-4 the two oligonucleotides were hybridized to restriction digests of plasmids containing cloned telomeric sequences of Tetrahymena, Trypanosoma, and Plasmodium. As can be seen in Figure

Figure 4.6 Hybridization Analysis of Blood and Sperm DNA Digests

Blood and sperm DNAs from the the same individual were digested with the restriction enzymes as shown. 10µg samples of the digests were separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The filter was probed with ^{32}P-labelled Try-4. A short exposure is shown in a) and a long exposure in b). The filter was stripped and reprobed with ^{32}P-labelled Tet-4 shown in c). The conditions used are described in Figure 4.4.
Key to DNAs S = sperm, B = blood.

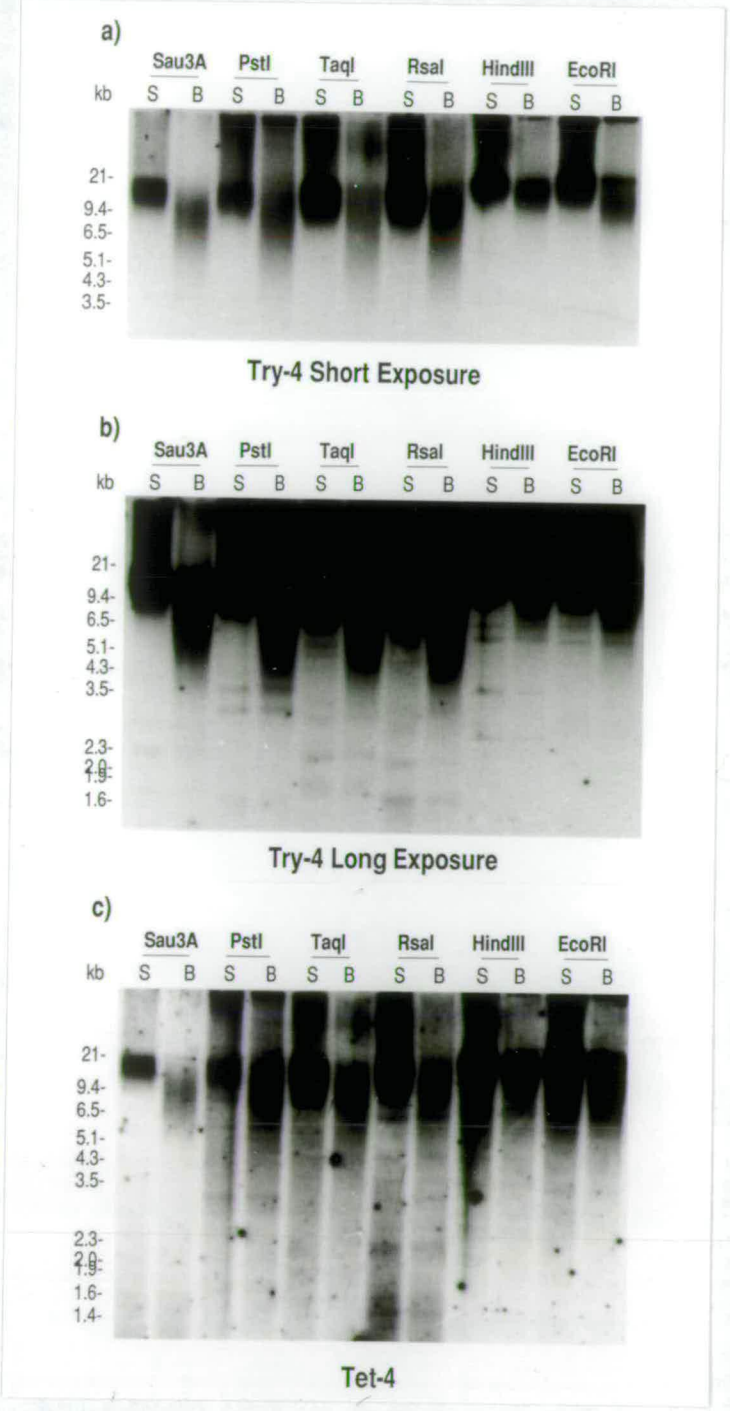
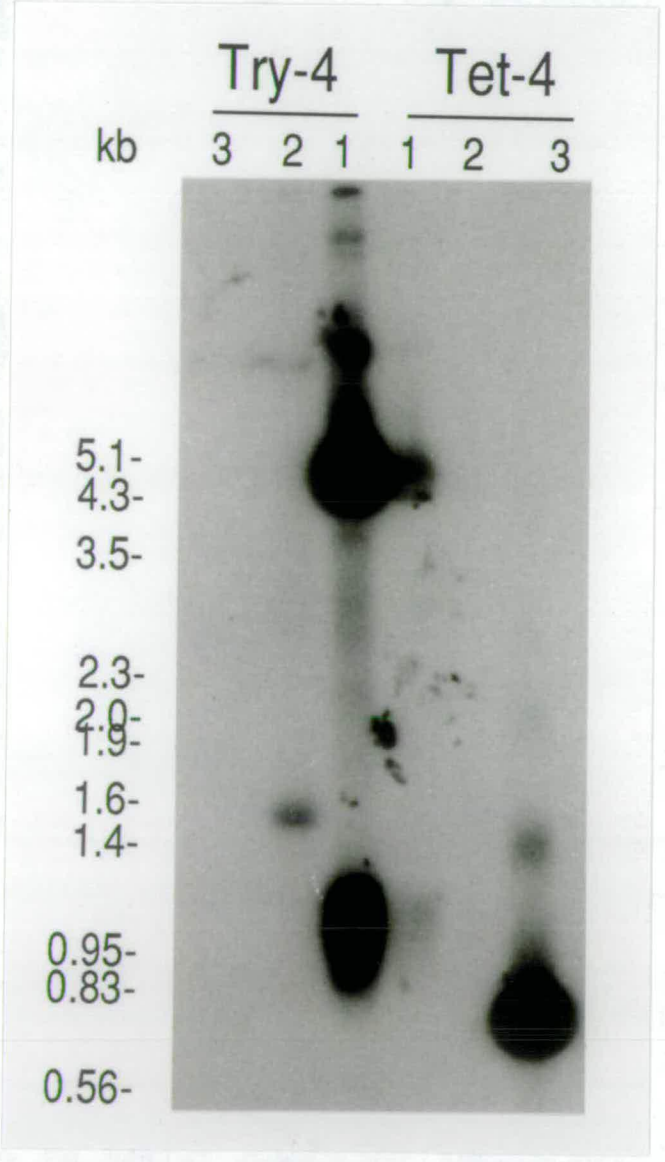


Figure 4.7 Cross-Hybridization of Try-4 and Tet-4

DNA samples as described below were elecrophoresed on a 0.8% agarose gel and transferred to Hybond-N. Lane 1, 500 ng partial BamHI/EcoR1 digest of clone Dirt3 (Forney *et al.*, 1988); lane 2, 300 ng 1.4 kb fragment containing *Plasmodium* telomeric repeat (gift of Dr Allshire); lane 3, 50 ng purified BamHI/EcoR1 fragment from pYAC4 containing 700 bp *Tetrahymena* telomeric repeat. Dirt3 contains a 1 kb BamHI/EcoR1 insert which includes ~58 tandem repeats of the *Didymium* telomeric sequence which is identical to that of *Trypanosoma*. The two halves of the filter were probed with Try-4 or Tet-4 as indicated. The conditions used are described in Figure 4.4.



4.7 Try-4 detects the Trypanosoma telomeric sequence strongly, as would be expected, the Plasmodium clone weakly and it does not detect the Tetrahymena clone. Tet-4 detects the Tetrahymena clone strongly, the Trypanosoma clone very weakly and it does not detect the Plasmodium clone. Clearly the two sequences do not cross-hybridize to any great extent under the conditions used so therefore arrays of both the TTAGGG sequence and the TTGGGG sequence are likely to be found at telomeres. It might be expected that the TTAGGG sequence would be more predominant at telomeres than the TTGGGG sequence as the TTAGGG sequence cross-hybridizes to human telomeres more intensely.

Figure 4.8 shows the result of hybridizing Try-4 to a human Bal-31 series which had subsequently been cut by the restriction enzyme Sau3A. A short and a long exposure of the same filter are shown. The heterogeneous hybridizing fragments are sensitive to Bal-31 digestion but the discrete hybridizing fragments are unchanged even after extensive Bal-31 digestion. Therefore internal fragments which cross-hybridize to Try-4 are not immediately sub-telomeric.

4.5 Trypanosoma Telomeric Repeat Cross-Hybridizes to Other Species

Allshire et al. (1988) showed that the the Tetrahymena telomeric repeat cross-hybridized to heterogeneous restriction fragments in many eukaryotic species. This result suggested that the Tetrahymena telomeric repeat may be cross-hybridizing to the telomeres of these species as well as those of human. Figure 4.9 shows the result of hybridizing the Trypanosoma oligonucleotide Try-4 to Sau3A restriction digests of DNA from several different species. Large heterogeneous restriction fragments are detected in all species, consistent with the

Figure 4.8 Hybridization of Try-4 to a Human DNA Bal-31 Series

Human sperm DNA was digested with Bal-31 for the times indicated (mins). The samples were then digested with the restriction enzyme Sau3A, 10µg aliquots separated by electrophoresis on a 1% agarose gel and then transferred to Hybond-N. The filter was hybridized with ^{32}P-labelled Try-4. The hybridization conditions used are described in Materials and Methods. The filter was washed at 4 x SSC 0.1% SDS at 65°C. An overnight and a 6 day exposure are shown.

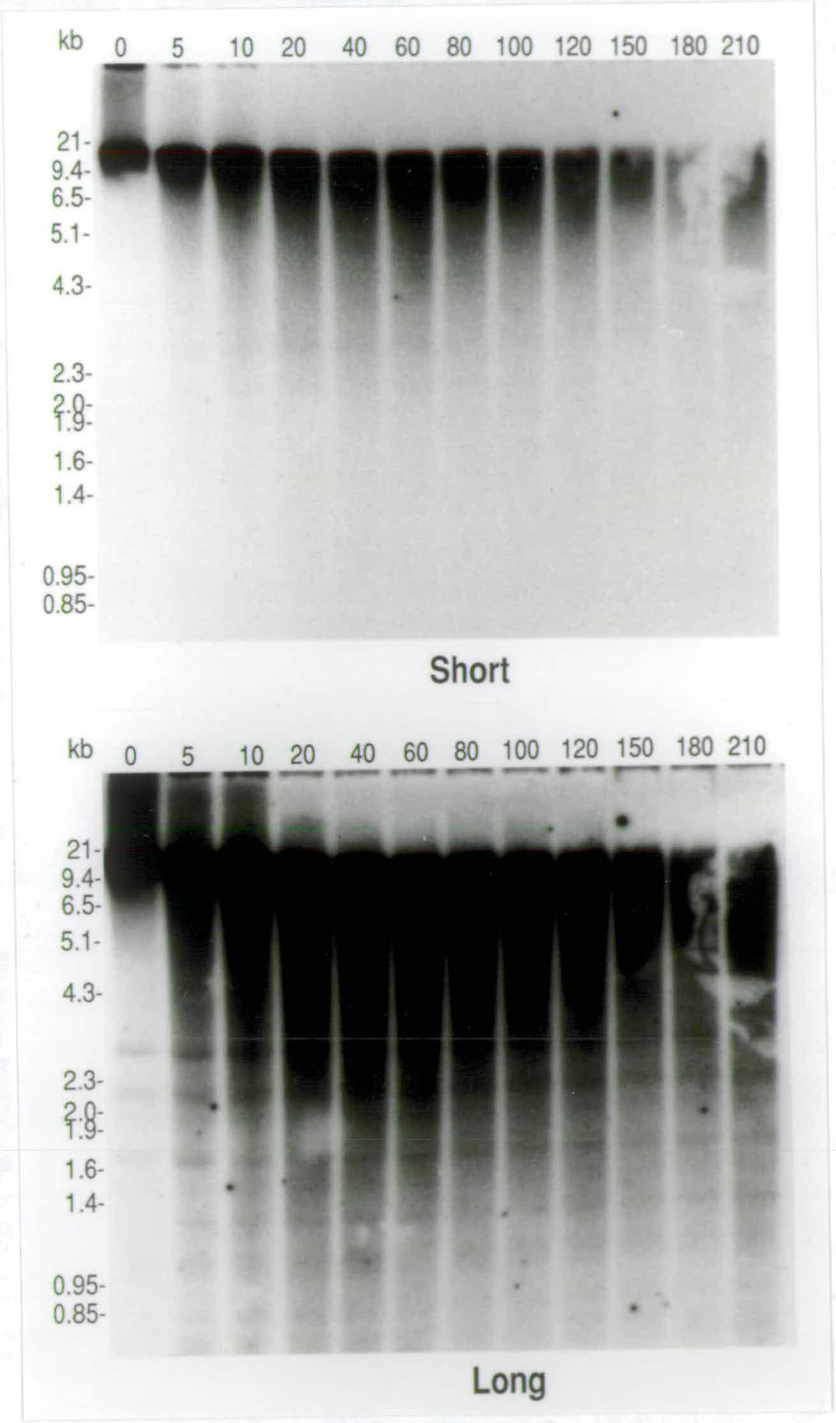
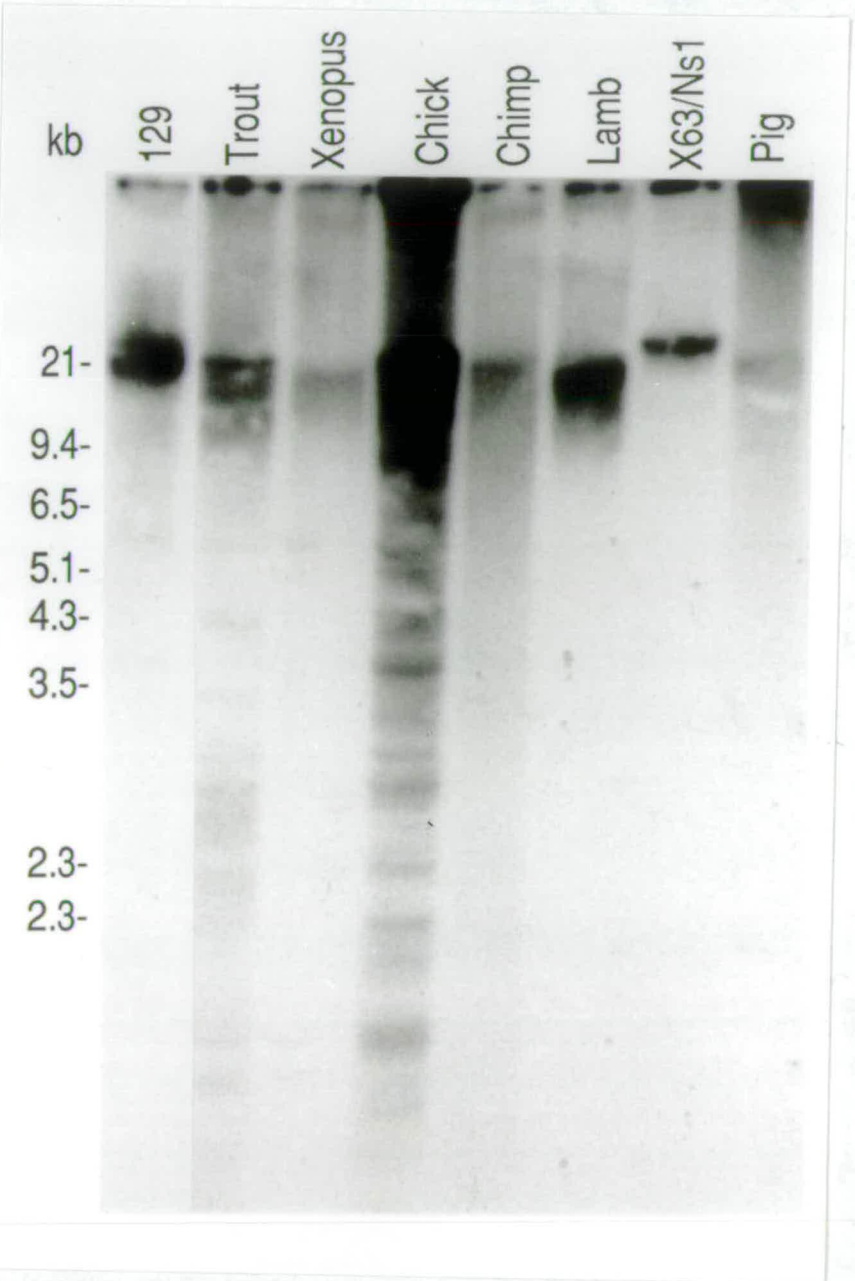


Figure 4.9 Cross-Hybridization of Try-4 to Several Species

DNAs from the species shown were digested with the restriction enzyme Sau3A, separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The filter was hybridized with ^{32}P-labelled Try-4 using the conditions described in Figure 4.4. 129 is a mouse strain, DNA extracted from male liver. X63/NS1 is a mouse cell-line.



interpretation that arrays of TTAGGG repeats are also found at the telomeres of these species. In the mouse the Sau3A heterogeneous restriction fragments detected are ~23 kb, implying that terminal repeat arrays on mouse chromosomes are much longer than those of human chromosomes(also see Figure 4.1). The result of hybridizing Try-4 to several mouse DNA digests is shown in Figure 4.10. A high molecular weight heterogeneous smear is detected with all enzymes, even when a restriction enzyme such as MnlI, which has a non-palindromic recognition sequence, is used. This result supports the conclusion that the arrays of sequences which cross-hybridize to TTAGGG found at the ends of mouse chromosomes are considerably longer than those found at the ends of human chromosomes.

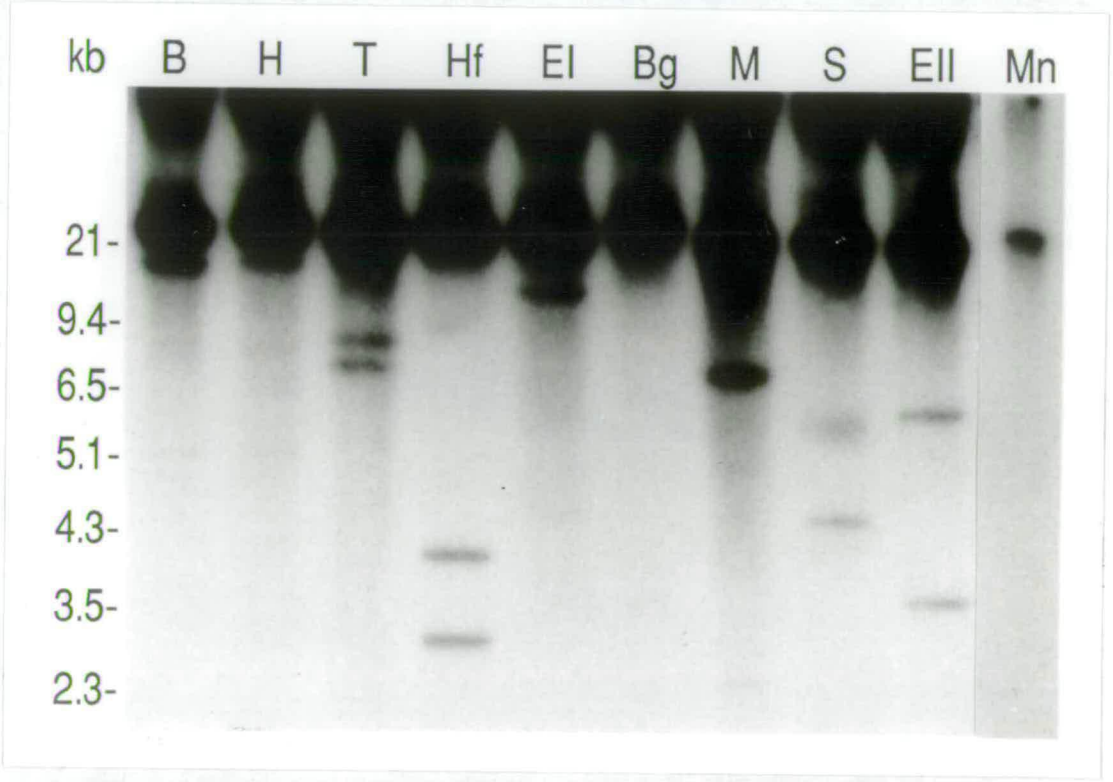
4.6 Discussion

The cross-hybridization experiments described here, and those performed by other workers, confirm the idea that arrays of simple repeats are found at the ends of human chromosomes which was suggested by the experiments done with 29C1. In addition they show that human telomeric repeats must be very similar, if not identical, to those of Trypanosoma and Tetrahymena. The results also show that the telomeres of other eukaryotic species are highly likely to consists of arrays of these repeats.

Although the sequence composition of the telomeres appears to be similar there are differences in the length of the telomeric array. In humans there are differences between blood and sperm telomeres in the length of terminal restriction fragments. Telomeres in sperm are ~5 kb longer than those of blood from the same individual. This difference

Figure 4.10 Hybridization of Try-4 to Mouse DNA Digests

DNA from the mouse strain 101 was digested with the restriction enzymes shown, separated by electrophoresis on 0.8% agarose gels and transferred to Hybond-N. The filters were hybridized with ^{32}P-labelled Try-4 using the conditions described in Figure 4.4. Key to restriction enzymes B = BamHI, H = HindIII, T = TaqI, Hf = HinfI, EI = EcoR1, Bg = BglII, M = MspI, S = Sau3A, EII = Eco RII and Mn = Mn1I.



is probably accounted for by a longer terminal repeat array at the end of sperm telomeres judging by the more intense hybridization of the cross-hybridizing probes to sperm DNA terminal restriction fragments. These extra terminal repeats may have an important germ-line function, or it may be that the mechanism for maintaining the telomeric repeats is less efficient in somatic as opposed to germ-line cells. There are also differences between species as demonstrated by the differences between the size of human terminal restriction fragments compared to those of mouse. As there are differences between human blood and sperm telomeres, and as mouse liver DNA was used, the possibility that the differences seen between species are just tissue-specfic cannot be completely excluded. However as the size of terminal fragments from mouse are so large in comparison to those of human blood DNA it is likely that the differences seen do reflect genuine species differences. In addition terminal restriction fragments of the mouse cell-line X63/NS1 are larger than those in the human cell-line PES (see Figure 4.1).

Both the Tetrahymena and the Trypanosoma telomeric repeat cross-hybridize to human telomeres. Recently Allshire et al. (1989) have demonstrated by cross-hybridization experiments that both these repeat sequences, and other similar sequences, are found at the ends of human chromosomes. From the Bal-31 experiments shown in Figure 4.4 it is clear that these sequences are probably confined to the most terminal 3 kb, although this is a gross estimate. However the "barren" region is at least 5 kb with all restriction enzymes tested so far which strongly suggests that other simple repeats will be found at human telomeres.

Although these cross-hybridization experiments have been very imformative, they cannot define exactly what sequences are found at the ends of human chromosomes and how they are arranged. The only way this can be done is by cloning and sequencing of human telomeres. The next chapter describes experiments designed to clone human telomeres.

CHAPTER 5

CLONING OF HUMAN TELOMERES BY COMPLEMENTATION IN YEAST

5 CLONING OF HUMAN TELOMERES BY COMPLEMENTATION IN YEAST

5.1 Introduction

The experiments in Chapter 4 show that arrays of simple repeats, whose sequence is very similar to that of the telomeric repeat of Trypanosoma, are likely to be found at the ends of human chromosomes. It is important to isolate human telomeres in order to determine precisely their structure and sequence, and to generate sub-telomeric sequences which could be used to define the end of the genetic and physical maps of human chromosomes.

I carried out a series of experiments designed to clone human telomeres by attachment of linkers or selectable markers to the ends of human chromosomes but these proved unsuccessful. As the human haploid genome size is $3x10^9$ bp organised into 23 chromosomes there is one chromosome end per $6.5x10^7$ bp. Most methods of DNA extraction cause reduction in size of the DNA resulting in large numbers of free ends in addition to the telomeres. Any cloning strategy designed to select for telomeres will inevitably clone these free ends as well, making the detection of telomeric clones harder. Therefore large numbers of clones would have to be generated in order to successfully clone a genuine telomere as opposed to a random end generated during DNA extraction.

A different strategy for isolating human telomeres based on selection for telomere function was devised. Telomeres of organisms such as Tetrahymena and Oxytricha can provide telomere function in the yeast S. cerevisiae (Shampay et al., 1984; Pluta et al., 1984). It seemed likely that human telomeres, which appear to have similar

sequences, would be able to provide telomere function in yeast as well. In addition, as the Trypanosoma telomeric repeat cross-hybridizes to human telomeres, but not to the Tetrahymena telomeric repeat (see Figure 4.7) or yeast DNA (not shown), this sequence could be used as a probe to identify clones containing human telomeric sequences.

5.2 Telomere Cloning Vector Construction

The yeast artificial chromosome (YAC) vector pYAC4NEO (Cooke and Cross, 1988) was modified to give a specialised telomere cloning vector. This vector, pYAC4NEONOT, when linearised requires the addition of a second telomeric sequence to be maintained as a linear minichromosome in yeast. Figure 5.1 illustrates how pYAC4NEONOT was derived from pYAC4NEO. pYAC4NEO was digested partially with XhoI, run on a 1% low gelling temperature (LGT) gel and the 12.5 kb fragment cleaved only once was cut out and purified. This fragment was then digested to completion with BamHI, run on a 1% LGT gel and the 10.1 kb fragment cut out and purified. The adaptor sequence (shown in Figure 5.1, oligonucleotides synthesised by John Inglis) was ligated to the pYAC4NEO fragment as in Lathe et al., (1984). This adaptor destroys the original BamHI site and recreates the XhoI site. The HIS3 gene and one of the Tetrahymena telomeric repeats $(TTGGGG)_n$ in pYAC4NEO are deleted. A NotI restriction site is introduced adjacent to the remaining Tetrahymena telomeric repeat and a BamHI site between the NotI and XhoI sites. The ligation was transformed into competent DH5α cells and transformants selected on Amp plates. DNA mini-preps were prepared from Amp resistant colonies and tested for the presence of a

Figure 5.1 Construction of the Vector pYAC4NEONOT

The YAC vector pYAC4NEO was modified as shown below to give pYAC4NEONOT. The cloning steps are described in Section 5.2.

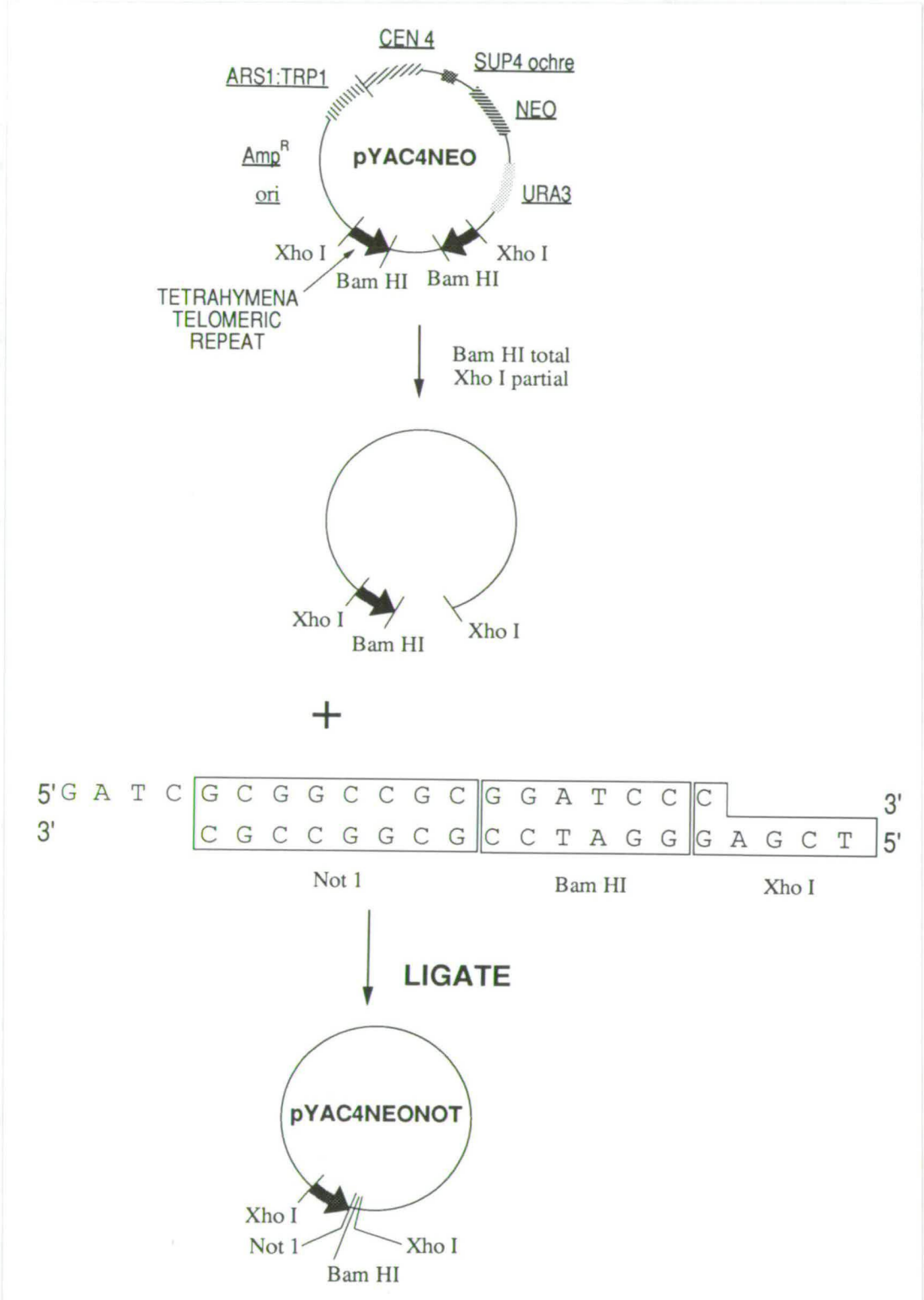
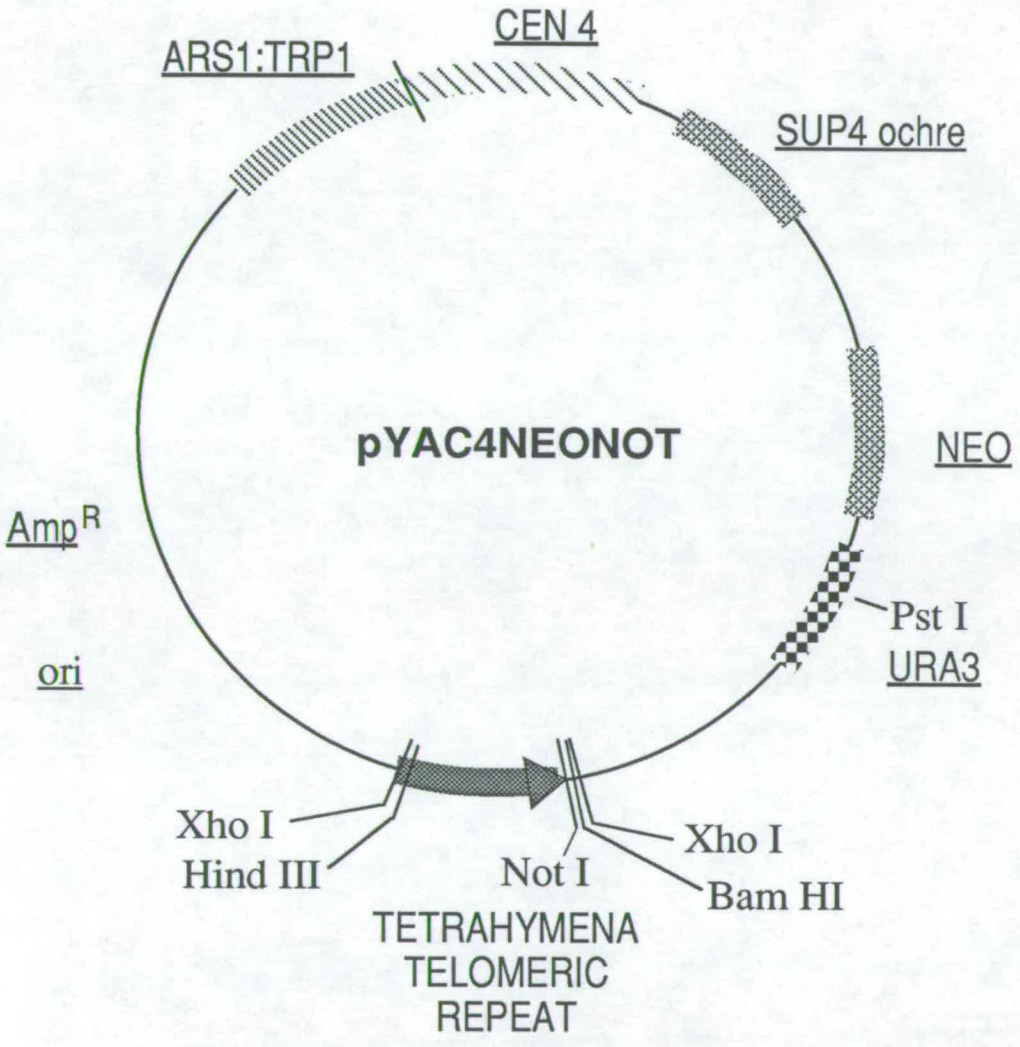


Figure 5.2 Diagram of pYAC4NEONOT

Diagram of pYAC4NEONOT showing the genes and important restriction sites.



NotI restriction site. A plasmid with a NotI site was identified in this way and named pYAC4NEONOT. The structure of pYAC4NEONOT was confirmed by restriction site mapping and is shown in Figure 5.2.

5.3 Human Telomere Cloning Strategy

The cloning scheme used is shown in Figure 5.3. Total human placental DNA was digested with Sau3A, electrophoresed on a 1% LGT agarose gel and the 5 - 15 kb fraction purified (this step was done by Dr Allshire). This size-selection enriches for telomeric sequences. Figure 4.6a shows that the Sau3A fragments detected by Try-4, a Trypanosoma telomere oligonucleotide $(TTAGGG)_4$, are between 5 and 15 kb. As Sau3A has a 4 bp recognition sequence, which is frequent in human DNA, most of the fragments generated by it are small. Therefore this size-selection should enrich greatly for telomeric sequences.

pYAC4NEONOT was digested with NotI and BamHI and then dephosphorylated with bacterial alkaline phosphatase (BAP). Test ligations were done to check that the vector could no longer ligate to itself but could ligate to BamHI digested pUC9 DNA.

The size-selected Sau3A digested human DNA was ligated to the prepared vector. In addition control ligations containing no DNA and only vector DNA were set up. These ligations and control DNAs were used to transform spheroplasts of the Ura^- haploid yeast strain AB1380 (Burke et al., 1987). The results of this transformation are shown in Table 5.1. The 240 Ura^+ colonies resulting from the transformation with the vector/human ligation, and 3 colonies from the pYAC4NEONOT

Figure 5.3 Construction of Sau3A Human Telomere Library

Diagram showing the scheme used to clone human telomeres by selection for function in yeast. Clones containing human telomeric sequences were identified by hybridization to Try-4.

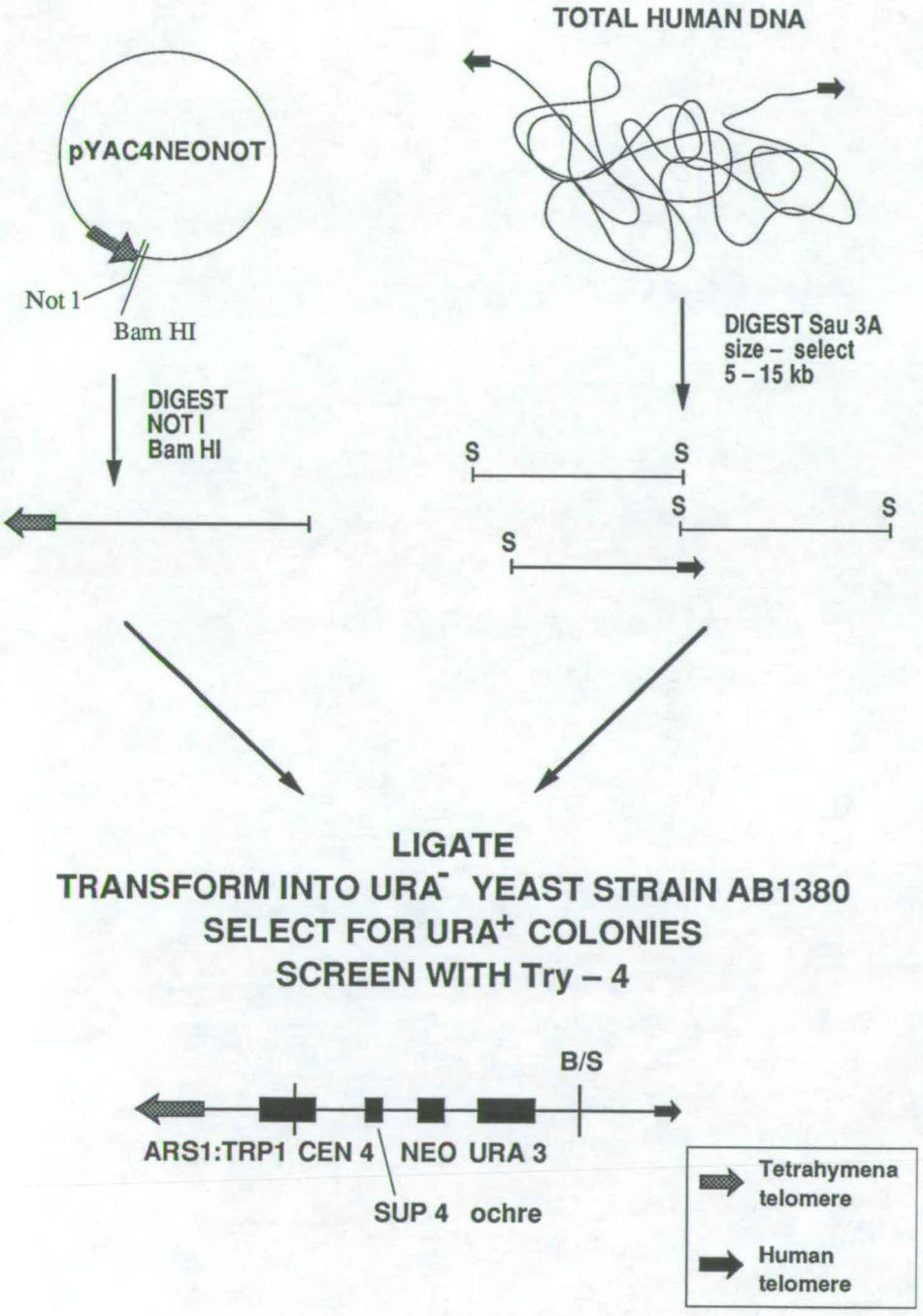


Table 5.1 Result of Transformation Experiment

Table showing the numbers of Ura^+ colonies obtained on transformation of AB1380 spheroplast with various DNAs.

DNA	Amount (ng)	Number of Ura^+ Colonies
YAC Plasmid DNA		
pYAC4	1	320
pYAC4NEO	1	120
pYAC4NEONOT	1	504
Linear YAC DNA		
pYAC4NEONOT (NotI/BamHI)	10	14
pYAC4NEONOT (NotI/BamHI BAP)	10	26
Ligations		
pYAC4NEONOT (NotI/BamH1 BAP)	10	22
pYAC4NEONOT/human	200	240
Blank	-	0
Cells only	-	0

transformation, were screened with Try-4. Hybridizing colonies were picked and rescreened with Try-4 and two strongly hybridizing positive colonies, Hutel 1 and Hutel 22, were found.

5.4 Analysis of Hutel 1 and Hutel 22

Hutel 1 and Hutel 22 were first analysed by pulsed field gel (PFG) electrophoresis. Figure 5.4 shows the results of this analysis. Figure 5.4a shows the analysis for Hutel 1. The ethidium bromide stained gel of AB1380 and Hutel 1 chromosomes showed that in Hutel 1 the 1.5 megabase (Mb) chromosome 4 had been replaced by two new chromosomes of 1.0 and 0.5 Mb. Both novel chromosomes hybridized with a CEN4 probe and with pBR322. The 1 Mb chromosome hybridized to Try-4 and to URA3. The 0.5 Mb chromosome hybridized with the ori region of pBR322 and the Tetrahymena telomere oligonucleotide Tet-4, $(TTGGGG)_4$. The 1 Mb chromosome is detected by Tet-4 as this oligonucleotide cross-hybridizes slightly to the TTAGGG sequence. Homologous recombination between the CEN4 region of pYAC4NEONOT and the original chromosome 4 in AB1380 would give rise to two chromosomes of the size and organisation observed as shown in Figure 5.5. Figure 5.4b is the PFG analysis of Hutel 22 which showed the presence of a 12 kb linear artificial chromosome which hybridized to pBR322, Try-4 and Tet-4.

Short linear chromosomes are mitotically unstable in yeast, whereas chromosomes of 55 kb and above behave much like natural chromosomes (Murray and Szostak, 1983). The presence of the ochre suppressor gene SUP4, which is an allele of the tryosine transfer RNA gene, in pYAC4NEONOT enabled the mitotic stability of the artificial chromosomes in the two Hutel clones to be easily tested. AB1380 carries the ade2-ochre mutation so cells in which the suppressor is

Figure 5.4 Pulsed Field Gel Analysis of Hutel 1 and Hutel 22

Yeast DNAs were prepared in agarose plugs as described in Materials and Methods. The pulsed field gel conditions used were 7 V cm^{-1} at 10°C and 90" between 120° reorientations. The gels were stained with ethidium bromide, photographed and transferred to Hybond-N. The filters were hybridized with the ^{32}P-labelled probes as listed below using the conditions described in Materials and Methods. Random-primed hybridizations were washed at 0.1xSSC, 0.1% SDS at 65°C. Oligonucleotide hybridizations were washed at 4xSSC, 0.1% SDS at 65°C.

a) Hutel 1 and AB1380 analysed by pulsed field gel electrophoresis. A,M and P, yeast marker chromosomes; B,D,F,H,J and N, AB1380; C,E,G,I,K, L and O, Hutel 1. A-C ethidium bromide stained gel. The hybridization probes used are shown. pBRori was a (PvuII-PstI) fragment. The two new chromosomes are arrowed.

b) Hutel 22 and AB1380 analysed by pulsed field gel electrophoresis. A,D and I yeast marker chromosomes; B,E and H AB1380; C,F and G Hutel 22. The hybridization probes used are shown.

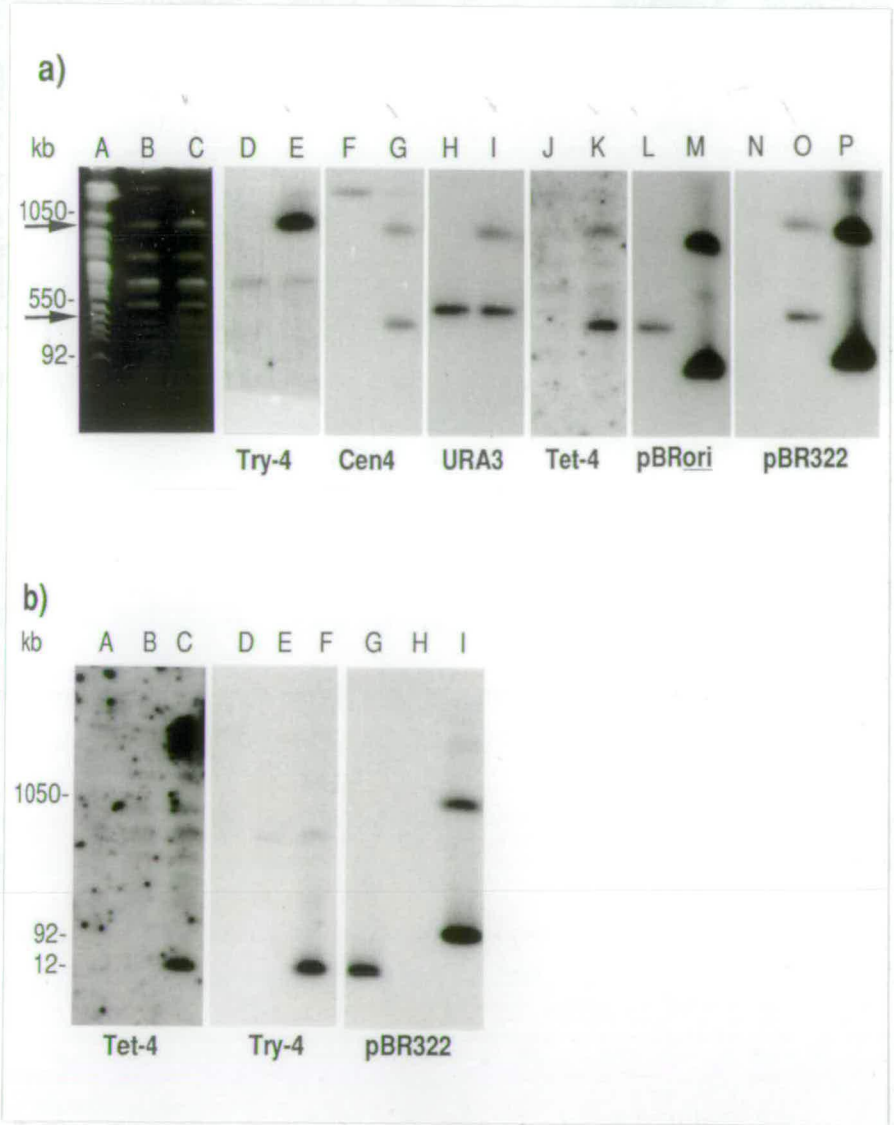
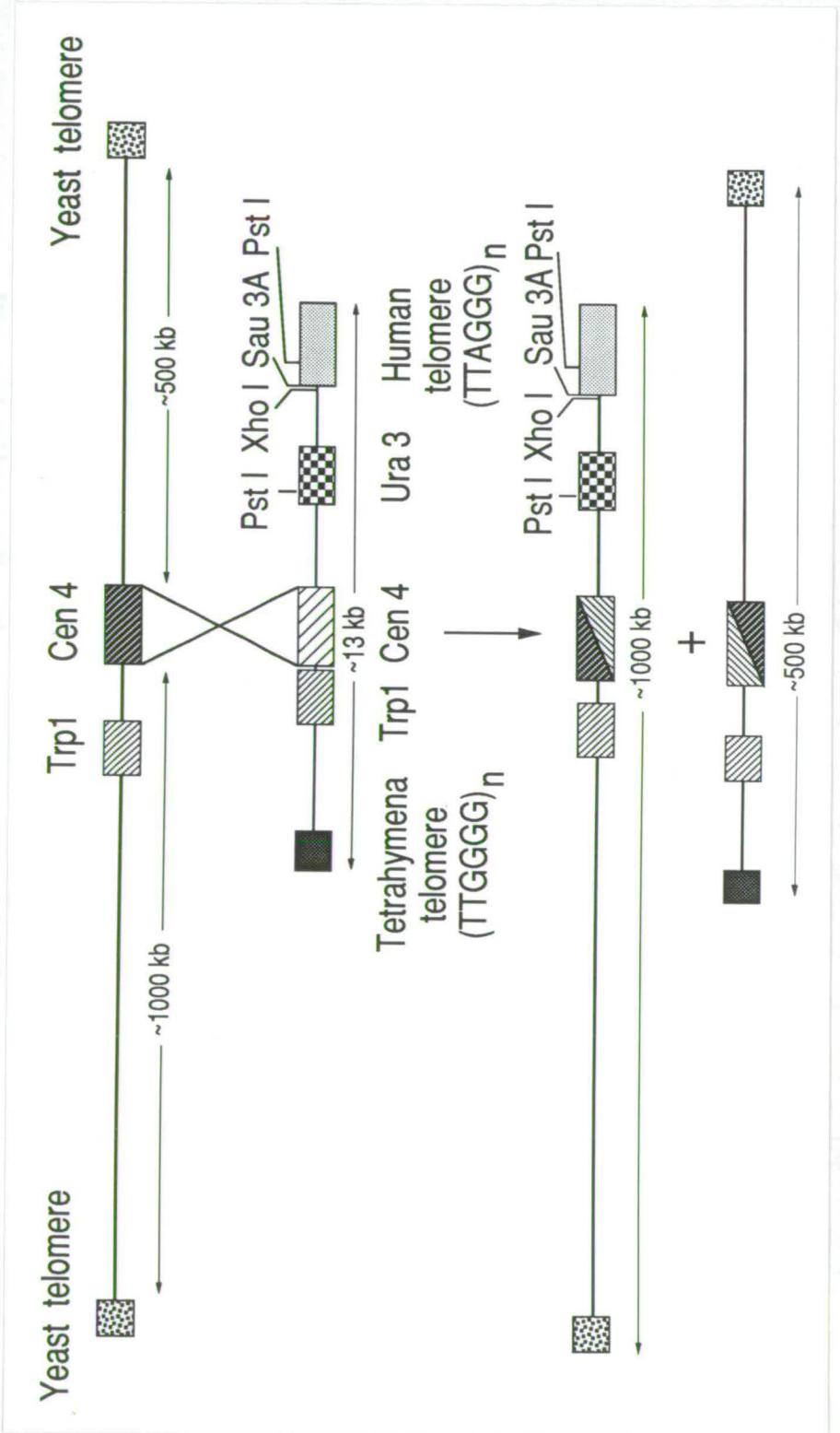


Figure 5.5 Proposed Recombination Event in Hutel 1

Recombination proposed to have occurred in Hutel 1 between the endogenous 1.5 Mb chromosome 4 and the transforming molecule. Two new chromosomes of 1 Mb and 0.5 Mb are generated. The restriction sites relevant to the chromosome construction and subcloning are shown. Figure is not to scale.



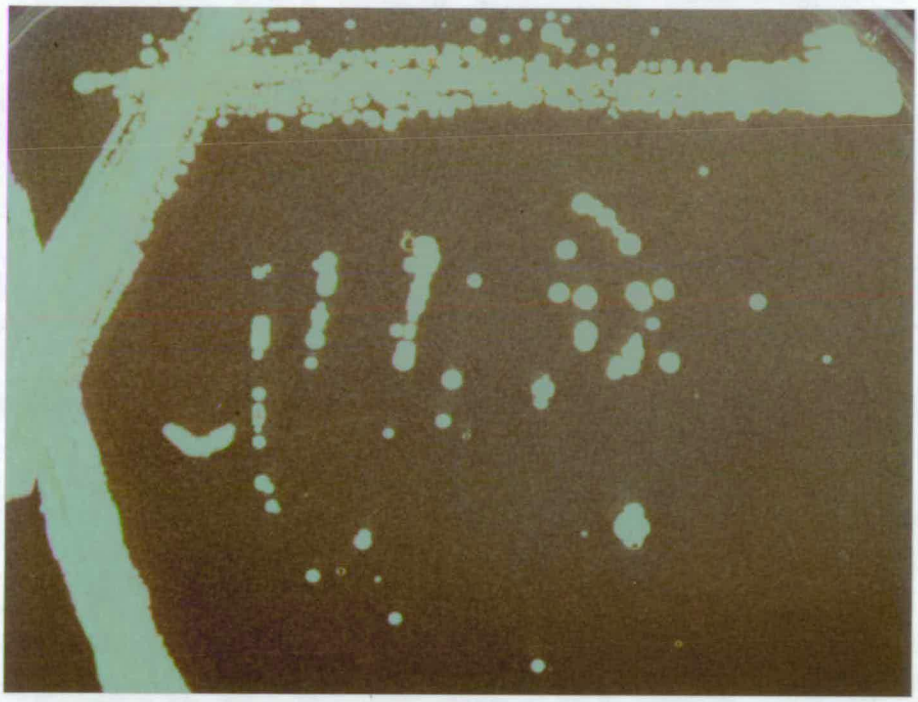
expressed form white colonies and those in which suppressor activity is absent form red colonies. Figure 5.6 shows the result of streaking Hutel 1 and Hutel 22 on non-selective media. The majority of colonies are white with Hutel 1 and red with Hutel 22. This reflects the instability of the 12 kb artificial chromosome in Hutel 22 which is lost in the absence of selection and the mitotic stability of the recombinant chromosome carrying the SUP4 gene in Hutel 1.

The next objective was to determine whether the human sequences in Hutel 1 and Hutel 22 were present at the ends of the new yeast chromosomes. One distinguishing feature of the end of a DNA molecule is that it presents an apparent site for all restriction enzymes. Figure 5.7 shows the results of hybridizing filters of various DNA digests of Hutel 1 and Hutel 22 with Try-4 and pBR322, and Tet-4 in the case of Hutel 22. Restriction maps constructed from this data have an apparent site for all the restriction enzymes used on the distal side of the Try-4 homology in both Hutel 1 and Hutel 22 (data not shown). In addition the restriction fragments detected by Try-4 are heterogeneous in all digests which is another characteristic of terminal restriction fragments. The terminal restriction fragments detected by Try-4 in Hutel 22 are more heterogeneous than those detected in Hutel 1. Telomeres of large chromosomes may be more homogeneous in yeast than those of small chromosomes, but a greater number of clones would need to be examined in order to confirm this idea. It may just be that the artificial chromosome in Hutel 1 represents a comparatively homogeneous population as it is mitotically stable and present in the cells at low copy number whereas the small artificial chromosome in Hutel 22 represents a heterogeneous

Figure 5.6 Stability of Hutel 1 and Hutel 22

Hutel 1 and Hutel 22 were streaked on non-selective medium and incubated at 30°C. White colonies contain the SUP4 gene carried on the pYAC4NEONOT vector, red colonies do not.

HUTEL 1



HUTEL 22

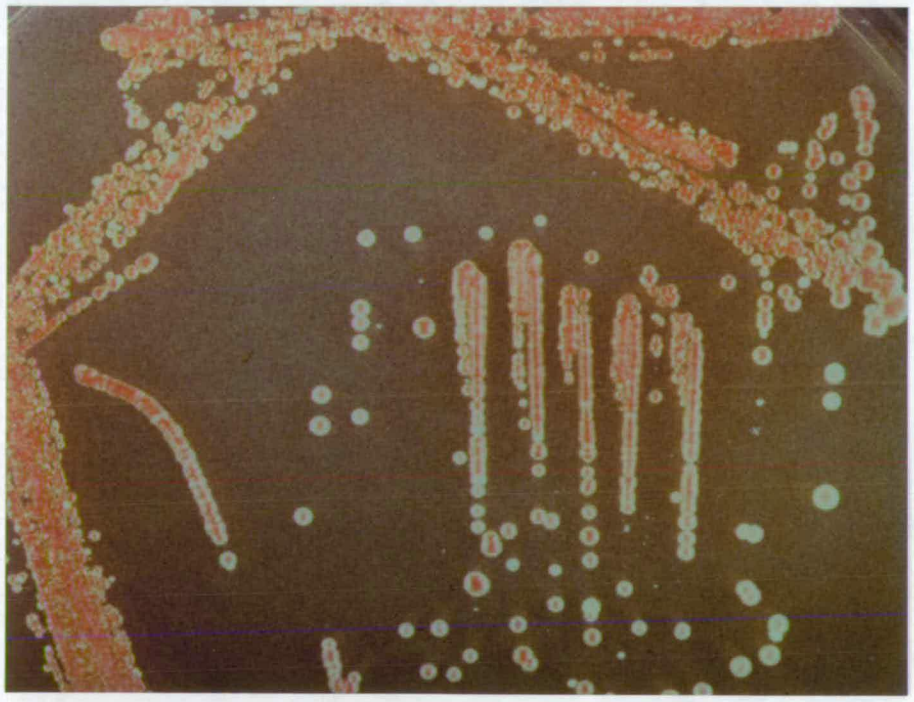
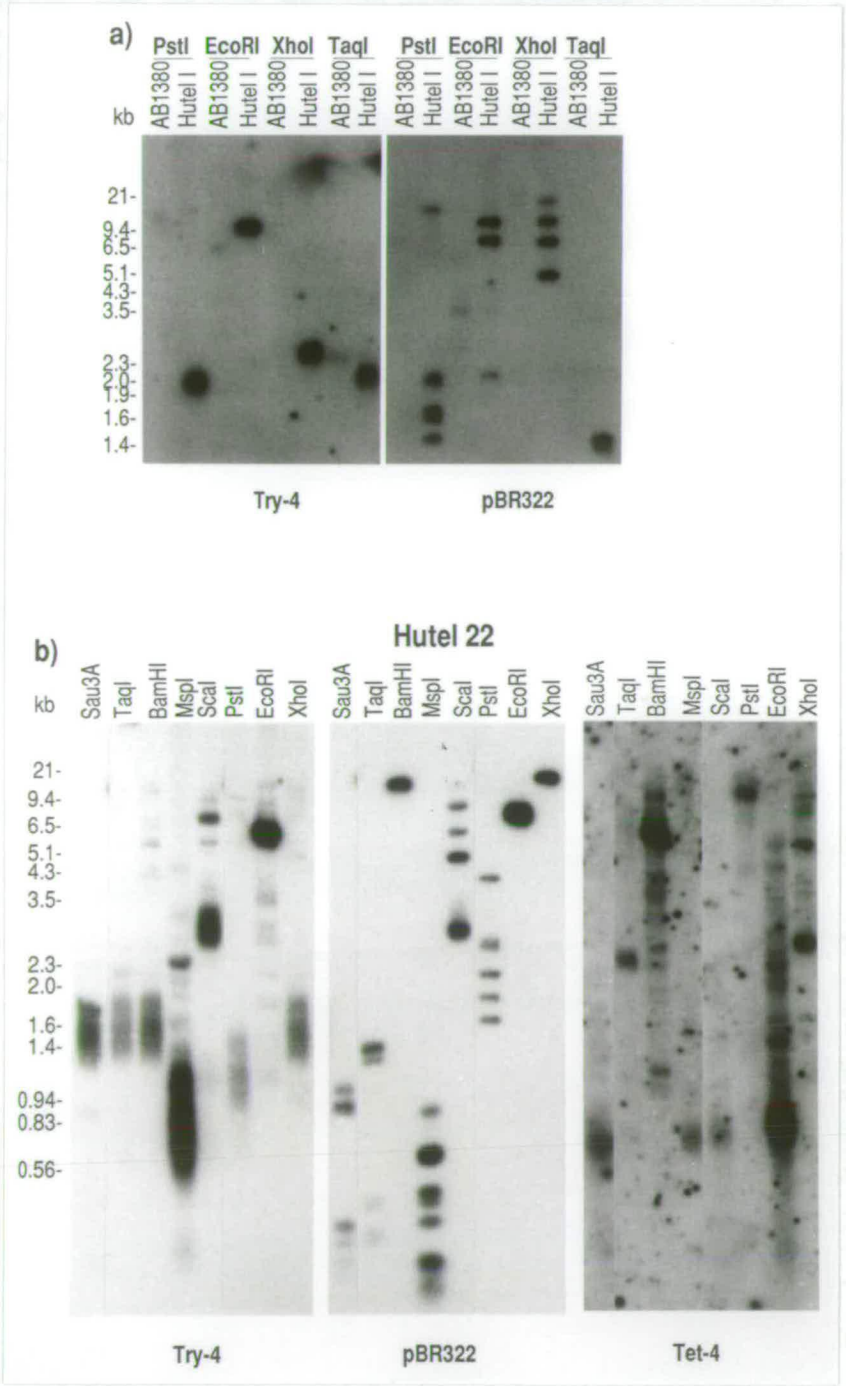


Figure 5.7 Hybridization Analysis of Hutel 1 and Hutel 22

Hutel 1, AB1380 and Hutel 22 DNAs were digested with the restriction enzymes as shown, the digests separated by electophoresis on 0.8% agarose gels and transferred to Hybond-N. The results of hybridizing the filters with ^{32}P-labelled pBR322 and Try-4, and Tet-4 in the case of Hutel 22 are shown, a) Hutel 1 and AB1380, b) Hutel 22. The conditions used are described in Figure 5.4.



population as it is mitotically unstable and is likely to be present at many copies per cell as would be expected for a small chromosome of its size (Murray and Szostak, 1983).

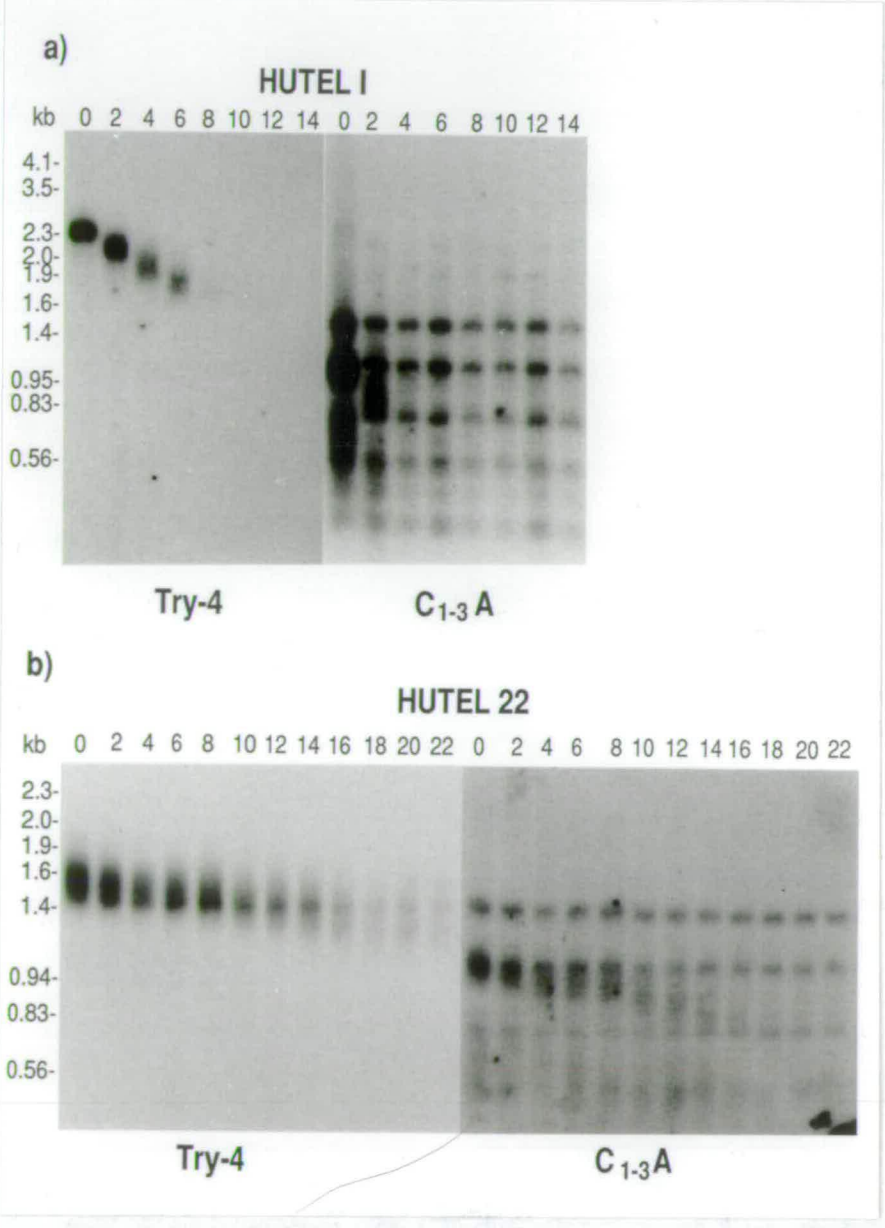
Heterogeneous fragments are detected by Tet-4 in digests of Hutel 22. This result is expected as Tet-4 is homologous to the Tetrahymena telomeric repeat at the other end of the transforming molecule. In contrast pBR322 detects discrete fragments in digests of Hutel 1 and 22 in addition to a few heterogeneous fragments. The discrete fragments detected are bounded by restriction sites and the heterogeneous fragments have a restriction site at only one end, the other end being the end of the chromosome.

If intact DNA is treated with the exonuclease Bal-31 prior to digestion with restriction enzymes terminal restriction fragments become shorter and then disappear completely with increasing time of Bal-31 digestion. Try-4 detects such Bal-31 sensitive fragments in both Hutel 1 and Hutel 22, shown in Figure 5.8. Hutel DNA was treated for increasing lengths of time with Bal-31 prior to digestion with the restriction enzyme Sau3A. The heterogeneous smear of hybridization detected by Try-4 becomes smaller and disappears after 8 mins of Bal-31 digestion for Hutel 1 and 18 minutes of Bal-31 digestion for Hutel 22. The filters were stripped and rehybridized with a probe homologous to the yeast telomeric repeat. This probe also detected a heterogeneous smear of hybridization which became smaller with increasing time of Bal-31 digestion. In addition it detected discrete fragments which were insensitive to Bal-31. Tracts of TG_{1-3} are found internally as well as at telomeres (Walmsley et al., 1984), so this result shows that the Bal-31 is degrading the DNA only from free ends.

The evidence presented in Figures 5.7 and 5.8 supports the

Figure 5.8 Try-4 Homologous Restriction Fragments are Bal-31 Sensitive

Hutel 1 and Hutel 22 DNAs were digested with Bal-31 for the times indicated (mins). The samples were then digested with the restriction enzyme Sau3A, separated by electrophoresis on 1% agarose gels and transferred to Hybond-N. Hutel 1 is shown in a) and Hutel 22 in b). The filters were first hybridized with ^{32}P-labelled Try-4 stripped and then hybridized with ^{32}P-labelled $C_{1-3}A$. $C_{1-3}A$ is an oligonucleotide, CCACACACCCACACACCACACCACACCCACACCCACACAC, which corresponds to the C-rich strand of cloned yeast telomeric repeat (gift of Dr Greider). The conditions used for both probes are as described in Figure 5.4 for Try-4.



conclusion that human sequences homologous to Try-4 are indeed located at the ends of the new yeast chromosomes in Hutel 1 and Hutel 22 and these sequences can provide telomere function.

5.5 Human Sequences in Hutel 1 and Hutel 22 are Telomeric in Humans

Most, but not all, of the sequences homologous to the Trypanosoma telomeric repeat in humans are telomeric as not all the cross-hybridizing sequences are Bal-31 sensitive (see Figure 4.8). Therefore it was important to establish whether the human sequences in Hutel 1 and Hutel 22 were derived from human telomeres. To do this human sequences with no homology to Try-4 were isolated from Hutel 1 and Hutel 22 and used as hybridization probes.

Figure 5.7 shows that the heterogeneous smear of hybridization detected by Try-4 is smaller in PstI digests than in Sau3A digests in both Hutel 1 and Hutel 22. Therefore there is a PstI site in the cloned human DNA between the Sau3A site and the region homologous to Try-4 in both these clones. As there is a PstI site in pYAC4NEONOT in the ura3 gene there is a PstI fragment in these clones which contains human sequences and pYAC4NEONOT sequences including parts of ura3 (see Figure 5.5). Dr Cooke recovered this PstI fragment from a PstI library of Hutel 1 made in the vector pTZ19 using a ura3 fragment as a probe. This clone was named pHutel1. A PstI library of Hutel 22 was made in the vector Bluescribe and screened with the same ura3 fragment. A clone containing the human sequences and part of pYAC4NEONOT was recovered from the library and named pHutel22. As there is an XhoI site immediately adjacent to the BamHI cloning site in pYAC4NEONOT the human sequences from these two clones are on XhoI - PstI fragments.

The size of the XhoI - PstI fragment containing the human sequences was 262 bp in pHutel1 and 420 bp in pHutel22. These two fragments were isolated from LGT agarose gels and used as probes. These sequences do not cross-hybridize (data not shown) and will be referred to as XP1 and XP22.

Figure 5.9 shows the results of hybridizing XP1 and XP22 to the filters with the Bal-31 series of Hutel 1 and Hutel 22 used in Figure 5.8. The result of hybridizing Try-4 to these filters is also shown. Both XP1 and XP22 detect heterogeneous restriction fragments that become smaller with increasing time of Bal-31 digestion. Interestingly the rate of Bal-31 digestion slowed after 8 minutes for Hutel 1 and 18 minutes for Hutel 22 judging by the the size of the heterogeneous smear of hybridization seen in the later time points. If the hybridization pattern obtained with XP1 and XP22 is compared with that obtained with Try-4 it appears that sequences homologous to Try-4 are relatively rapidly degraded by the Bal-31 but the adjacent sequences are not. Similar differences in the rate of Bal-31 digestion of telomeric and immediately sub-telomeric sequences have been observed in Tetrahymena (Henderson et al., 1988). Rapid digestion of yeast telomeric repeats by Bal-31 has been observed (Shampay et al., 1984). Human duplex telomeric repeats appear to have a similar structure to those of these other eukaryotes which makes them particularly susceptible to Bal-31 digestion.

Figure 5.10 shows the results of hybridizing XP1 and Try-4 to a filter of human DNA treated with Bal-31 for different times and then digested with Sau3A. XP1 detects heterogeneous Bal-31 restriction fragments which are the same size as those detected by Try-4. In addition a number of discrete fragments are detected which are not

Figure 5.9 XP1 and XP22 Detect Terminal Restriction Fragments

The filters with the Bal-31 series of Hutel 1 and Hutel 22 DNAs subsequently digested with Sau3A used in Figure 5.8 were used. a) shows the result of hybridizing the Hutel 1 Bal-31 series with ^{32}P-labelled XP1 and also with ^{32}P-labelled Try-4. b) shows the result of hybridizing the Hutel 22 Bal-31 series with ^{32}P-labelled XP22 and also with ^{32}P-labelled Try-4. The conditions used are described in Figure 5.4.

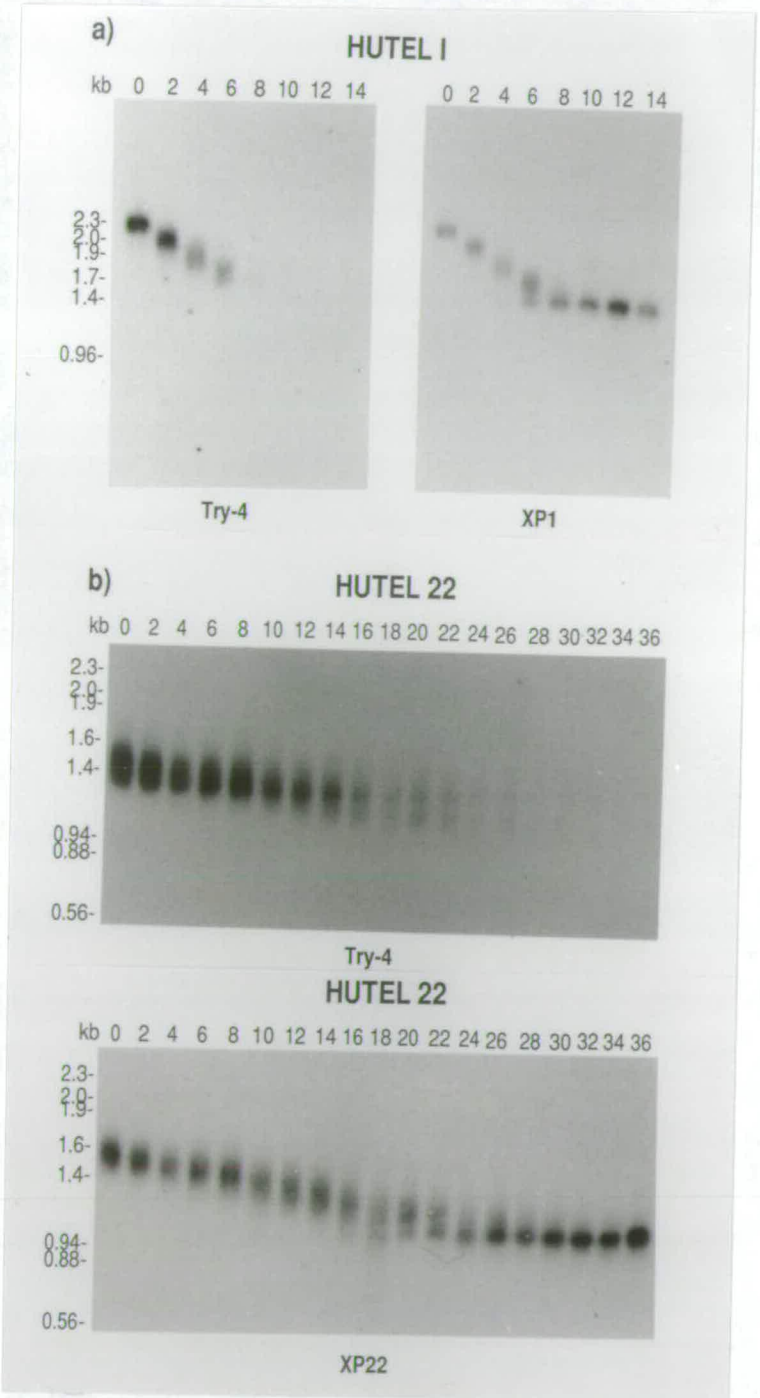
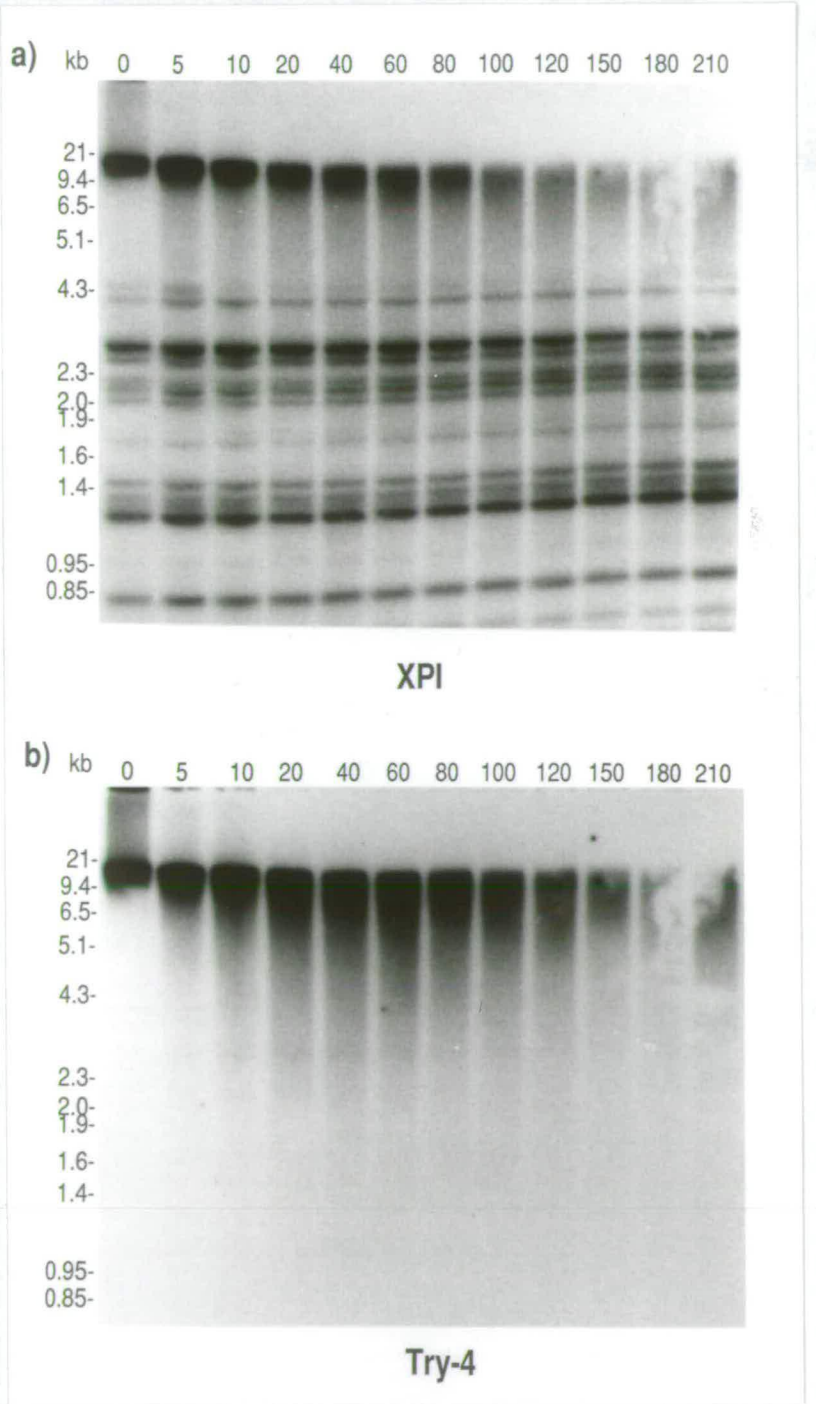


Figure 5.10 Terminal Location of XP1 in the Human Genome

Human sperm DNA was digested with Bal-31 for the times indicated (mins). The samples were then digested with the restriction enzyme Sau3A, 10µg aliquots separated by electrophoresis on a 1% agarose gel and then transferred to Hybond-N. The filter was first hybridized with a) ^{32}P-labelled XP1 stripped and then hybridized with b) ^{32}P-labelled Try-4. The conditions used are described in Figure 5.4.



degraded by Bal-31. Figure 5.11 shows the results of hybridizing XP1 and Try-4 to a filter of Sau3A and PstI digests of blood and sperm DNA from the same individual. XP1 detects heterogeneous Sau3A restriction fragments which are larger in sperm than in blood and which are the same size as fragments detected by Try-4. As expected from the restriction map of Hutel 1 in PstI digests XP1 detects only discrete fragments and Try-4 detects heterogeneous fragments. The human sequences in Hutel 1 must be immediately sub-telomeric in the human genome.

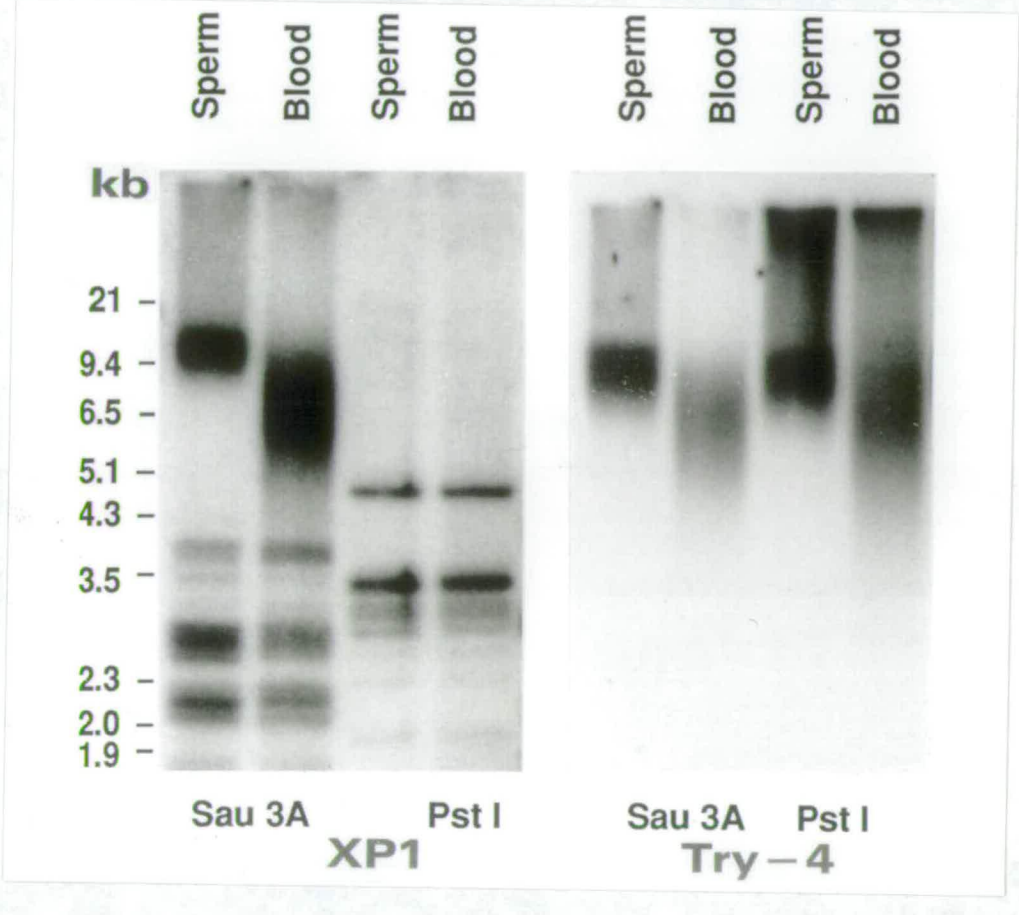
A similar analysis was done with XP22. All parts of XP22 which were used as hybridization probes contained sequences which were highly repeated in the human genome and which detected a generalised smear with all restriction enzymes. Therefore it proved impossible to establish conclusively that the human sequences in Hutel 22 were derived from a telomeric location in human DNA.

5.6 Discussion

The experiments described in this chapter demonstrate that human telomeric sequences are similar to those of Tetrahymena (Shampay et al., 1984) and Oxytricha (Pluta et al., 1984) in that they can provide telomere function in yeast. This result supports the idea that human telomeres have a very similar structure to those of lower eukaryotes suggested by the high degree of sequence conservation between human telomeres and other species discussed in Chapter 4. All these results suggest that arrays of simple repeats, whose sequence must be very

Figure 5.11 Terminal Location of XP1 in the Human Genome

Blood and sperm DNAs from the same individual were digested with the restriction enzyme Sau3A or PstI, separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The filter was first hybridized with ^{32}P-labelled XP1, stripped and then hybridized with ^{32}P-labelled Try-4. The conditions used are described in Figure 5.4.



similar to that of the Trypanosoma telomeric repeat, are found at the ends of human chromosomes. The number of the repeats is variable which leads to heterogeneity of terminal restriction fragments.

The Sau3A fragments of human DNA used in the cloning experiment were between 5 and 15 kb whereas the human sequences cloned in Hutel 1 were ~2.3 kb and those in Hutel 22 were ~1.5 kb. In addition hybridization analysis showed that the human sequences represented heterogeneous terminal restriction fragments. The best explanation of these observations is that the human sequences have been considerably shortened before being stabilised by the addition of variable numbers of yeast terminal repeats resulting in the heterogeneity observed. The human sequences were probably degraded by nucleases before the healing reaction in the yeast. This implies that although human terminal repeats are a suitable substrate for the addition of yeast terminal repeats they are not themselves recognised as a telomere per se in yeast. On yeast chromosomes tracts of the terminal repeat TG_{1-3} vary from 250 to 400 bp in length (Walmsley and Petes, 1985). Judging by the size of terminal restriction fragments detected by Try-4 the tract of human telomeric repeats has been shortened so that it is about this size. Yeast may not tolerate long terminal tracts of repeats of the kind found at the ends of human chromosomes.

The human telomeric and sub-telomeric sequences in both Hutel 1 and Hutel 22 display differential sensitivity to Bal-31 in that the rate of Bal-31 digestion slows once the terminal repeats have been degraded. The degradation of Tetrahymena telomeres by Bal-31 has the same kinetics (Henderson et al., 1988). Immediately sub-telomeric sequences are very A-T rich in Tetrahymena and as Bal-31 is known to degrade A-T rich DNA faster than G-C rich DNA this result is

surprising. This indicates that the repeats found at the telomeres of these organisms confers a special structure on the duplex DNA which results in it being particularly susceptible to Bal-31. This unusual structure may be important for telomere function. It is also possible that the sub-telomeric sequences are particularly refractory to Bal-31 digestion.

Preliminary analysis of the human sub-telomeric sequences isolated from Hutel 1 has shown that these sequences are not restricted to telomeres but are found internally on chromosomes as well. The organisation of these sequences on chromosomes and its sequence will be investigated in Chapter 6.

Cloning of human telomeres in yeast using a combination of selection for function and homology to the Trypanosoma telomeric repeat has provided a powerful means by which telomeric and sub-telomeric sequences can be isolated. Single-copy sub-telomeric sequences which can be linked to the most distal markers currently available will define the length of the map. The methodology used here for the isolation of human telomeres should be applicable to other species whose telomeres cross-hybridize to telomeric sequences from lower eukaryotes.

CHAPTER 6

ANALYSIS OF HUTEL 1 PRO-TERMINAL SEQUENCES

6 ANALYSIS OF HUTEL 1 PRO-TERMINAL SEQUENCES

6.1 Introduction

The human pro-terminal sequence XP1, whose isolation was described in Chapter 5, detects both telomeric and internal fragments. This conclusion is reached because this sequence detects both heterogeneous and discrete restriction fragments, of which only the heterogeneous restriction fragments are sensitive to the exonuclease Bal-31. In this chapter the sequence composition of XP1 and the hybridization properties of the XP1 fragment and other pro-terminal sequences derived from Hutel 1 will be investigated.

6.2 Sequencing of XP1

The XP1 fragment was sub-cloned into M13 and single-stranded template made which was then sequenced using the dideoxy chain termination method (Sanger et al., 1977).

The sub-clone pHutel1 was digested with XhoI and PstI, the digest run on a 1% LGT gel and the 262 bp XP1 fragment cut out and purified. This fragment was then digested with Sau3A. M13 mp18 and M13 mp19 vector DNAs were digested with BamHI and PstI. Ligations of vector alone and vector plus XP1 were set up and transformed into JM101 as described in Materials and Methods. Single-stranded template was prepared from recombinant white plaques and sequenced using the dideoxy chain termination method of Sanger. The consensus sequence obtained from several runs of both strands is shown in Figure 6.1a.

Figure 6.1 Sequence of XP1

a) Sequence of the human part of XP1.
b) Sequence composition of the XP1 sequence.
c) Dinucleotide frequencies of the XP1 sequence.

All sequence analysis used the STADEN programs.

a) Sequence

```
        10         20         30         40         50
GATCTTGATA GTACTGAGGA GGATTGGTCA GGCATTTTGT AGAATGTTAT

        60         70         80         90        100
TTTACCTCCT TGTCTGCAAG AGGGCCCGGC AGTGTCCGCA GCTGCCAGCA

       110        120        130        140        150
GGCGGGCGTG CTGCCACTAC GATGTGAGCA AGAGGGCCCT GCAATGTCCC

       160        170        180        190        200
TAGCTGCCAG CAGCGTGCCG CCACTATACT GCAAGCAAGA GAGCCCTCGC

       210        220        230        240        250
GTGCCCCGCG CCAGCAGTGG GCGCTGGACA CCACTGTCAC CTACAGGGCC


CTGCAG
```

b) Sequence Composition

Base	T	C	A	G
Number	51	76	51	78
Percentage	19.9%	29.7%	19.9%	30.5%

c) Dinucleotide Frequencies

	T		C		A		G	
	OBS	EXPECTED	OBS	EXPECTED	OBS	EXPECTED	OBS	EXPECTED
T	3.92	4.00	3.14	5.96	3.92	4.00	9.02	6.04
C	7.06	5.96	9.41	8.88	8.63	5.96	4.71	9.00
A	3.53	4.00	4.31	5.96	2.35	4.00	9.80	6.04
G	5.49	6.04	12.94	9.00	5.10	6.04	6.67	9.12

The base composition and dinucleotide frequencies found in XP1 are shown in Figure 6.1b and c. The sequence has a G + C content of 60%. This is considerably higher than that found in bulk DNA which is 40%. High G + C content is one characteristic of CpG-rich islands. These are regions of the genome which are unmethylated, have high G + C content and have no suppression of the dinucleotide CpG (Bird, 1986). One characteristic of these regions is that the frequencies of the two dinucleotides CpG and GpC are the same. A comparison of the frequencies of the two dinucleotides CpG and GpC in the XP1 sequence shows that the frequency of GpC is 12.94% and the frequency for CpG is 4.71%. This suppression of CpG is less than that found in non-island DNA, but it is unlikely that the XP1 sequence is an island-like sequence. One reason for this is that CpG-rich islands are often found associated with the 5' end of genes (Bird, 1987). Preliminary evidence suggests that XP1 does not have this characteristic. When XP1 was examined for open reading frames stop codons were found in all six frames (data not shown), although it is possible that if XP1 is part of an island it could be outside the coding region because islands usually overlap upstream sequences. The length of the human DNA in the XP1 sequence is only 256 bp, so it would be necessary to determine the sequence of flanking DNA and perform the statistical analysis on that before any firm conclusions can be made about any special properties of this sequence. High C + G content may be a characteristic of telomeric regions. Recognition sites for rare-cutting enzymes with two CpG dinucleotides in their recognition sequences are frequent in the distal part of the pseudoautosomal region, indicating that there is a high density of unmethylated CpG dinucleotides in this region (Petit et al., 1988). The clustering of these sites is not reminiscent of

CpG-rich islands (Bird, 1986). No sequences with significant homology to the XP1 sequence were found when the merged GENBANK and EMBL databases were searched with the XP1 sequence (data not shown).

The XP1 sequence was searched for internally repeated sequences. Figure 6.2 shows the results of this analysis. Using a minimum repeat length of 9 bases three regions of homology are revealed. If the minimum repeat length is reduced to 8 bases a fourth copy of the same repeat can be seen in addition to other shorter repeats. Figure 6.2 shows the XP1 sequence arranged so that these repeats are aligned. The most striking repeat motifs are two almost perfect 36 bp repeats and a third, more degenerate copy.

Although, like the telomeric repeats, the XP1 sequence has a high G + C content the G's and C's are not segregated onto different strands as they are in the telomeric repeats.

6.3 Hybridization of XP1 to Blood and Sperm Digests

The result of hybridizing various digests of blood and sperm DNAs from the same individual with XP1 is shown in Figure 6.3. This is interesting for a number of reasons. Discrete restriction fragments are detected in addition to heterogeneous telomeric fragments. PstI is the only enzyme with which no heterogeneous fragments are detected. In Hutel 1 there is a PstI site on the telomeric side of the XP1 fragment, so XP1 might be expected to detect only discrete PstI restriction fragments. If XP1 homologous sequences are present at more than one chromosome end there must always be at least one PstI site between the XP1 homology and the end. Fewer fragments are detected when an enzyme with a 6 bp as opposed to a 4 bp recognition sequence

Figure 6.2 XP1 Sequence Analysed for Repeats

Sequence of XP1 aligned to show repeats. Dashes indicate where gaps have been introduced into the sequence and identical bases are in capital letters and underlined.

CENTROMERE<-------gatcttgatagtactgaggaggattggtcaggcattttgtagaa
tgttattttacctccttgtc

tGCAAGAGGGCCCgGCAGTGTCCgcAGCTGCCAGCAG-gcgggCGTGCtGCCACTAcgatgtg

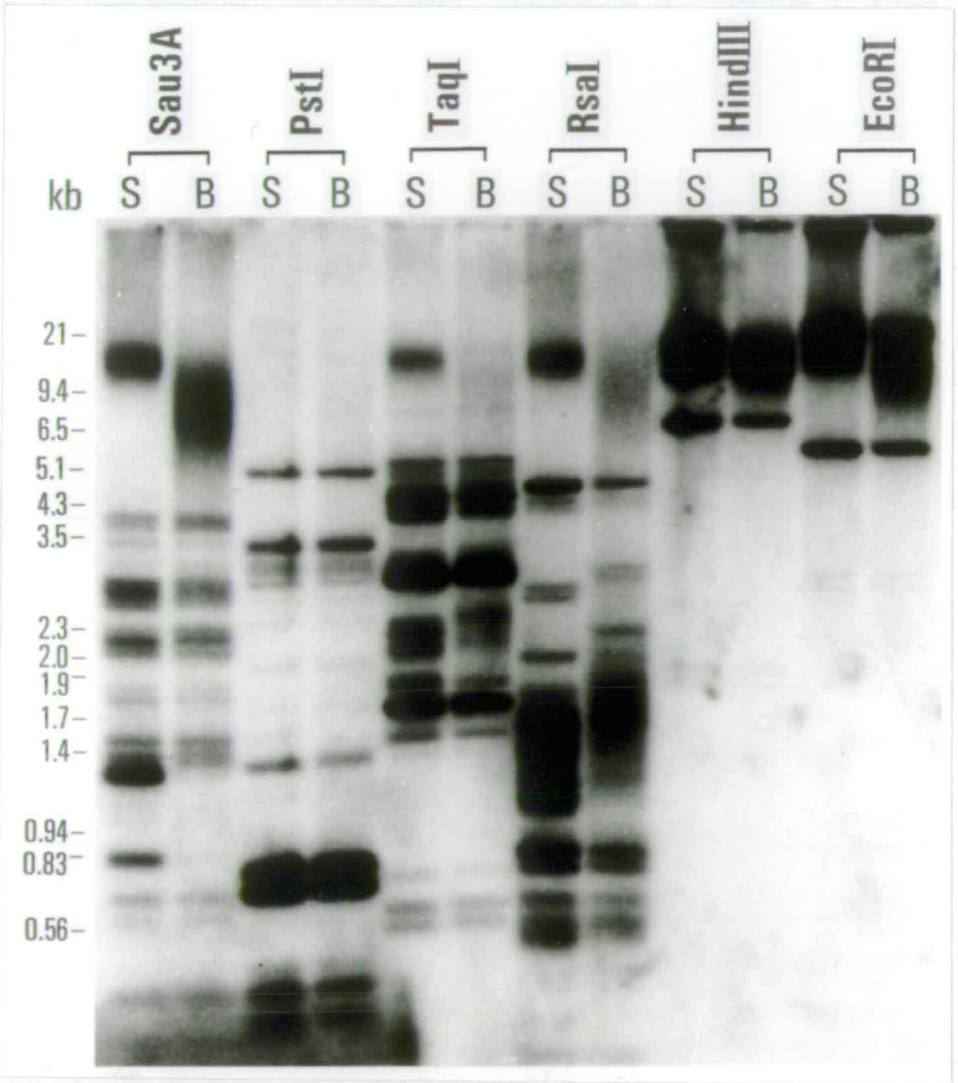
AGCAAGAGGGCCCTGCAaTGTCCCtAGCTGCCAGCAG------CGTGCcGCCACTAtactgca

AGCAAGAGaGCCCTcgcGTGcCCCg--C-GCCAGCAGtgggcgctggacaCCACTGtc

AcCtAcAGGGCCCTGCAG-------->TELOMERE

Figure 6.3 Hybridization of XP1 with Blood and Sperm Digests

Blood and sperm DNAs from the same individual were digested with the restriction enzymes shown, separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The filter was hybridized with ^{32}P-labelled XP1 as described in Materials and Methods. The filter was washed at 0.1 x SSC 0.1% SDS at ^{65}°C and exposed to XAR-5 film. S = sperm DNA and B = blood DNA.



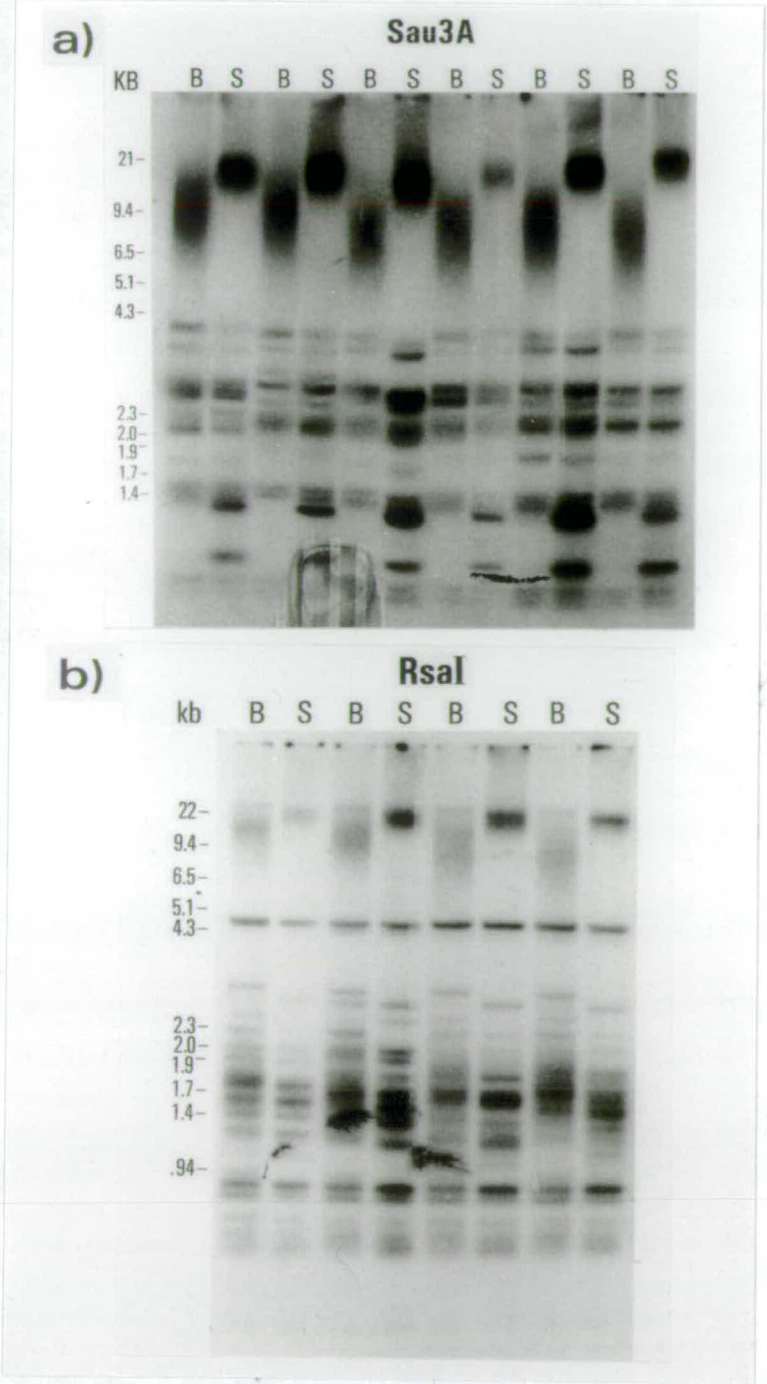
is used, with the exception of PstI. This may indicate that regions of homology to XP1 are found only at a few sites in the genome, perhaps in tandem arrays, or that more copies of this sequence lie within a few kb of the terminal repeats and are not separated from them by these restriction enzymes.

Intriguingly there appear to be differences between the discrete fragments detected in blood and sperm digests with some restriction enzymes. With Sau3A some hybridizing fragments in the sperm DNA digest appear to be increased in intensity, consistent with an amplification of a sequence containing this Sau3A fragment. With RsaI although basically the same hybridization pattern is detected in blood and sperm DNA digests a group of fragments are larger in the blood digests. A similar phenomenon is seen with TaqI, although the effect is less marked. The same differences between blood and sperm DNA have been found with all individuals so far examined. This is shown in Figure 6.4a for Sau3A and Figure 6.4b for RsaI.

A number of other enzymes have been tested to determine if they too have different XP1 hybridization patterns in blood and sperm DNA digests, but no differences have been found with any of these enzymes (data not shown). The three restriction enzymes Sau3A, RsaI and TaqI all share a common feature; they either contain a CpG dinucleotide in their recognition site, as in the case of TaqI, or they can overlap a CpG sequence, as with Sau3A and RsaI. Most occurrences of the CpG dinucleotide are methylated in mammalian genomes (Bird, 1987). This methylation is known to block cleavage by some restriction enzymes. The restriction enzyme Sau3A is blocked by such methylation whereas its isoschizomer MboI is not. Sau3A will not cut GAT^{m5}C but MboI will cut GAT^{m5}C (Nelson and McClelland, 1989). The methylation status of

Figure 6.4 Hybridization of XP1 with Sau3A and RsaI Digests

Blood and sperm DNAs from the same individual were digested with the restriction enzyme Sau3A, shown in a), or RsaI, shown in b). The digests were separated by electrophoresis on 0.8% agarose gels and transferred to Hybond-N. The filters were hybridized with ^{32}P-labelled XP1 using the conditions described in Figure 6.3.
S = sperm DNA and B = blood DNA.



different tissues is known to vary. In sperm satellite sequences are undermethylated whereas dispersed repetitive elements and some, but not all, gene sequences are heavily methylated (Sanford et al., 1985). Sau3A and MboI digests of blood and sperm were hybridized with the XP1 fragment to see if the blood/sperm hybridization differences found with Sau3A are also found with MboI. As can be seen in Figure 6.5 both blood and sperm MboI digests have the same hybridization pattern and that pattern is the same as that of Sau3A digested sperm DNA. Therefore methylation differences between blood and sperm DNA are responsible for the apparent amplification of some Sau3A fragments in sperm DNA as compared to blood DNA. The sperm DNA appears to be undermethylated with respect to blood DNA from the same individual. Probably most chromosomes have these Sau3A sites containing a ^{m5}C but a few do not. Other enzymes known to be sensitive to methylation were tested in a similar way. The results of comparing the isoschizomers HpaII and MspI are shown in Figure 6.6. HpaII is sensitive to methylation whereas MspI is not. The patterns of XP1 hybridization seen are similar to those seen with Sau3A and MboI in that hybridization differences between blood and sperm digests are apparent with HpaII but not with MspI. There are differences between the hybridization patterns of sperm DNA MspI and HpaII digests indicating that some restriction fragments detected by XP1 in these digests are bounded by methylated sites in the sperm DNA. Therefore although sperm DNA is undermethylated when compared to blood DNA at HpaII sites in regions homologous to XP1 it is not completely unmethylated.

The restriction enzymes TaqI and RsaI are reported to be insensitive to methylation (Nelson and McClelland, 1989). Therefore the differences between the hybridization patterns of blood and sperm

Figure 6.5 Hybridization of XP1 with Sau3A or MboI Digests

Blood and sperm DNAs from the same individual were digested with the restriction enzyme Sau3A or MboI, separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The filter was hybridized with ^{32}P-labelled XP1 using the conditions described in Figure 6.3. S = sperm DNA and B = blood DNA.

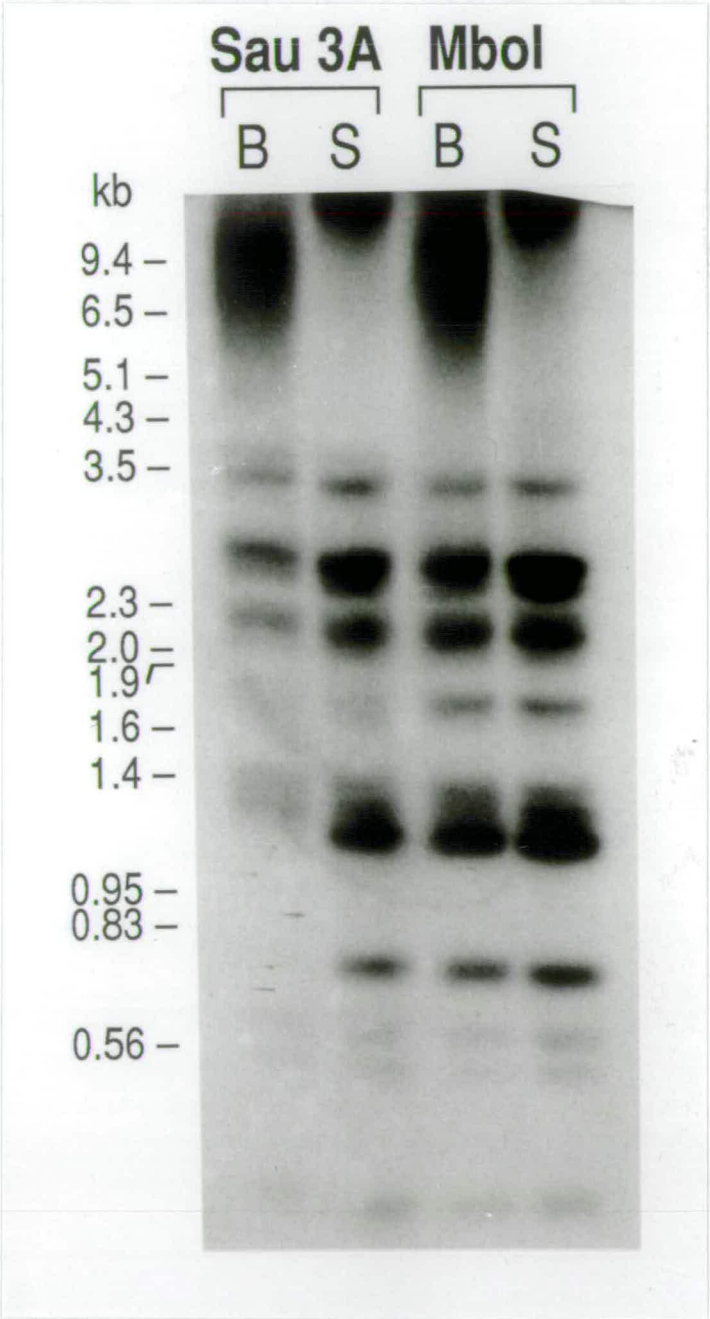
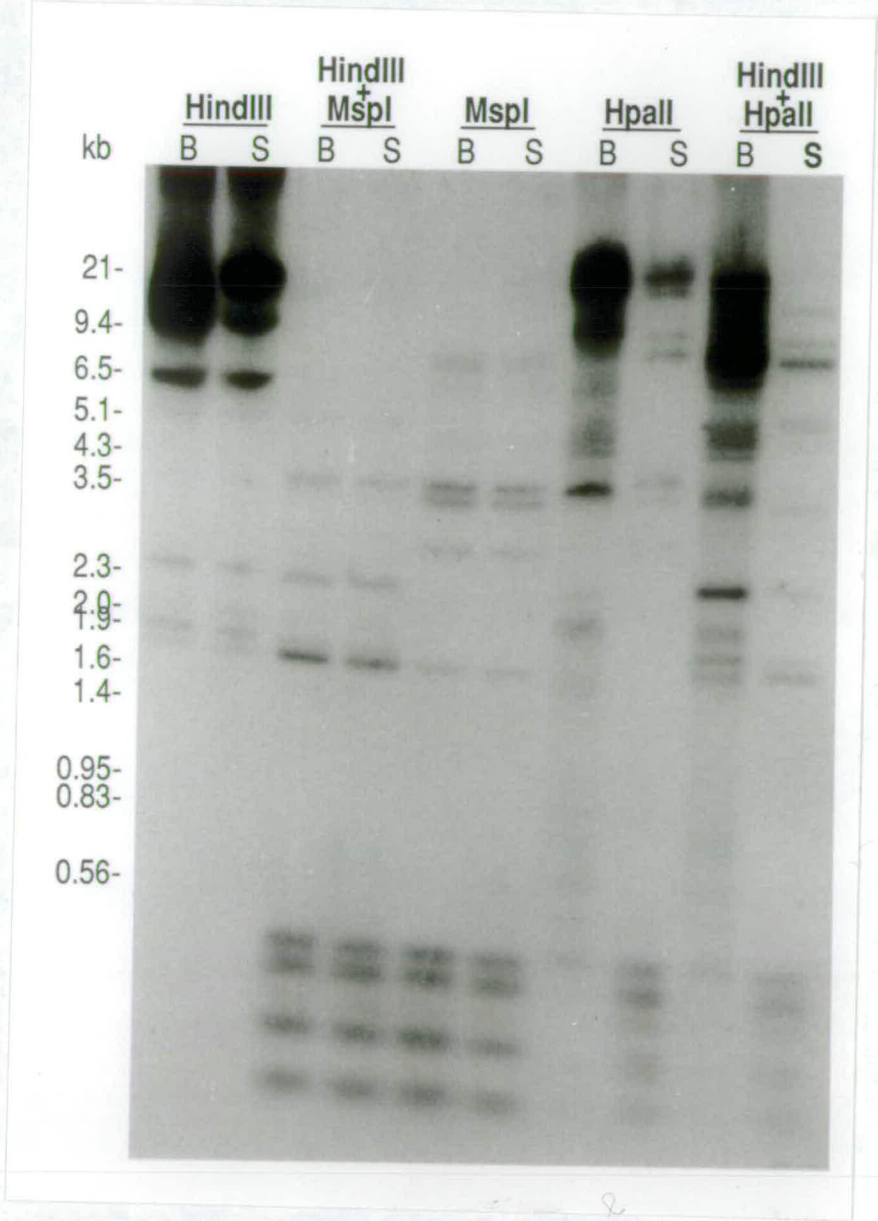


Figure 6.6 Hybridization of XP1 with HpaII and MspI Digests

Blood and sperm DNAs from the same individual were digested with the restriction enzymes as shown, separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The filter was hybridized with ^{32}P-labelled XP1 using the conditions described in Figure 6.3. S = sperm DNA and B = blood DNA.



DNAs observed with these enzymes cannot be accounted for by methylation. The same differences in the hybridization pattern of blood and sperm DNA digests is seen with the isoschizomer of TaqI, TthHB81, which is also reported to be insensitive to methylation (data not shown). There are no isochizomers available for RsaI. However the restriction enzyme ScaI has the recognition sequence AGTACT, the central four bases of which is a RsaI site. The XP1 hybridization patterns of ScaI blood and sperm DNA digests do show some differences. There appears to be an extra hybridizing fragment of ~20 kb present in the blood digest (data not shown). The molecular basis of these differences is unknown at present.

Preliminary experiments done by Dr Cooke have indicated that the XP1 hybridizing fragments, which exhibit the methylation based Sau3A differences, are probably telomeric. Telomeric sequences can be separated from bulk genomic DNA by isopycnic centrifugation on a Cs_2SO_4 gradient in the presence of silver ions (Brown, 1989). BamHI digested human blood and sperm DNAs were fractionated over such a column. The DNA responsible for the Sau3A methylation based differences fractionated with the telomeric fraction. Similar analysis for the hybridizing fragments which exhibit blood/sperm differences with RsaI and TaqI did not give clear-cut results but these are likely to be telomeric as well because in another experiment most of the XP1 cross-hybridizing sequences co-fractionated with the telomeric sequences.

6.4 Evolutionary Conservation of XP1 Cross-hybridizing Sequences

To determine if sequences homologous to XP1 are conserved a filter of Sau3A digests of DNAs from various species was hybridized with XP1. The result is shown in Figure 6.7. XP1 detected cross-hybridizing sequences in chimpanzee and human but no other species examined. Therefore XP1 does not appear to be highly conserved. This is in contrast to the terminal telomeric repeat, TTAGGG, which is found in many species (see Figure 4.9). The simple repeats which confer telomere function are well-conserved whereas the telomere-adjacent sequences, which probably perform different functions, are not highly conserved.

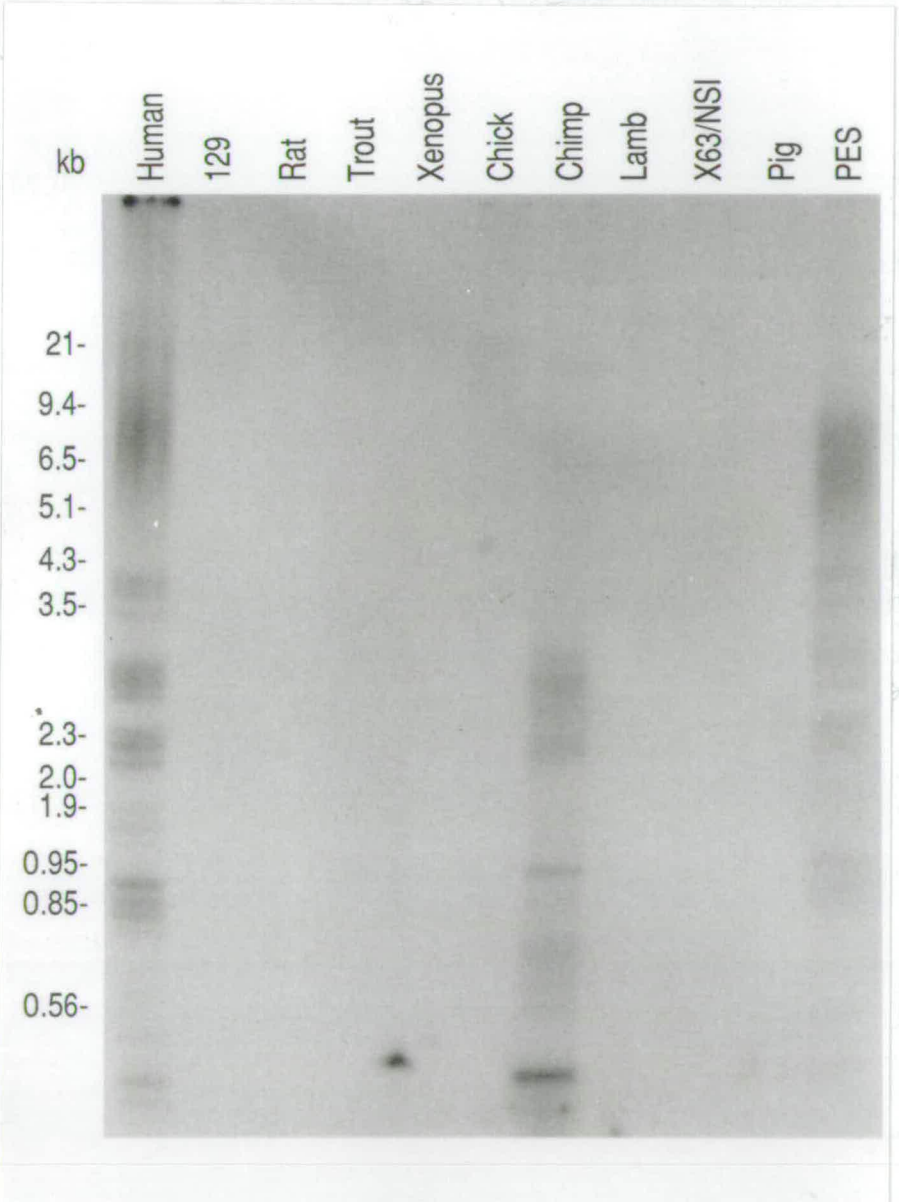
6.5 Other Human Pro-terminal Sequences From Hutel 1

The human sequences from Hutel 1 were sub-cloned by Dr Cooke into the vector pBS and named pHutel-2-end. The human sequences were cloned in the following way. Hutel 1 DNA was briefly treated with Bal-31 and an adaptor, blunt at one end and with an EcoRI sticky end at the other end, was ligated. This DNA was then digested with XhoI and ligated with SalI/EcoRI-digested pBS DNA, transformed into DH5α cells and the desired clone selected by homology to XP1 and Try-4. A single NaeI site was mapped between the PstI site and the terminal repeats in pHutel-2-end.

In order to determine if other sequences in pHutel-2-end have the same hybridization characteristics as XP1, fragments were isolated from pHutel-2-end and used as hybridization probes. The result of hybridizing the 333 bp PstI/NaeI fragment, which is immediately

Figure 6.7 Cross-Hybridization of XP1 to Other Species

DNAs from the species shown were digested with the restriction enzyme Sau3A, separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The filter was hybridized with ^{32}P-labelled XP1 using the conditions described in Figure 6.3. 129 is a mouse strain, DNA extracted from male liver. X63/NS1 is a mouse cell-line and PES is a human lymphoblastoid cell-line.



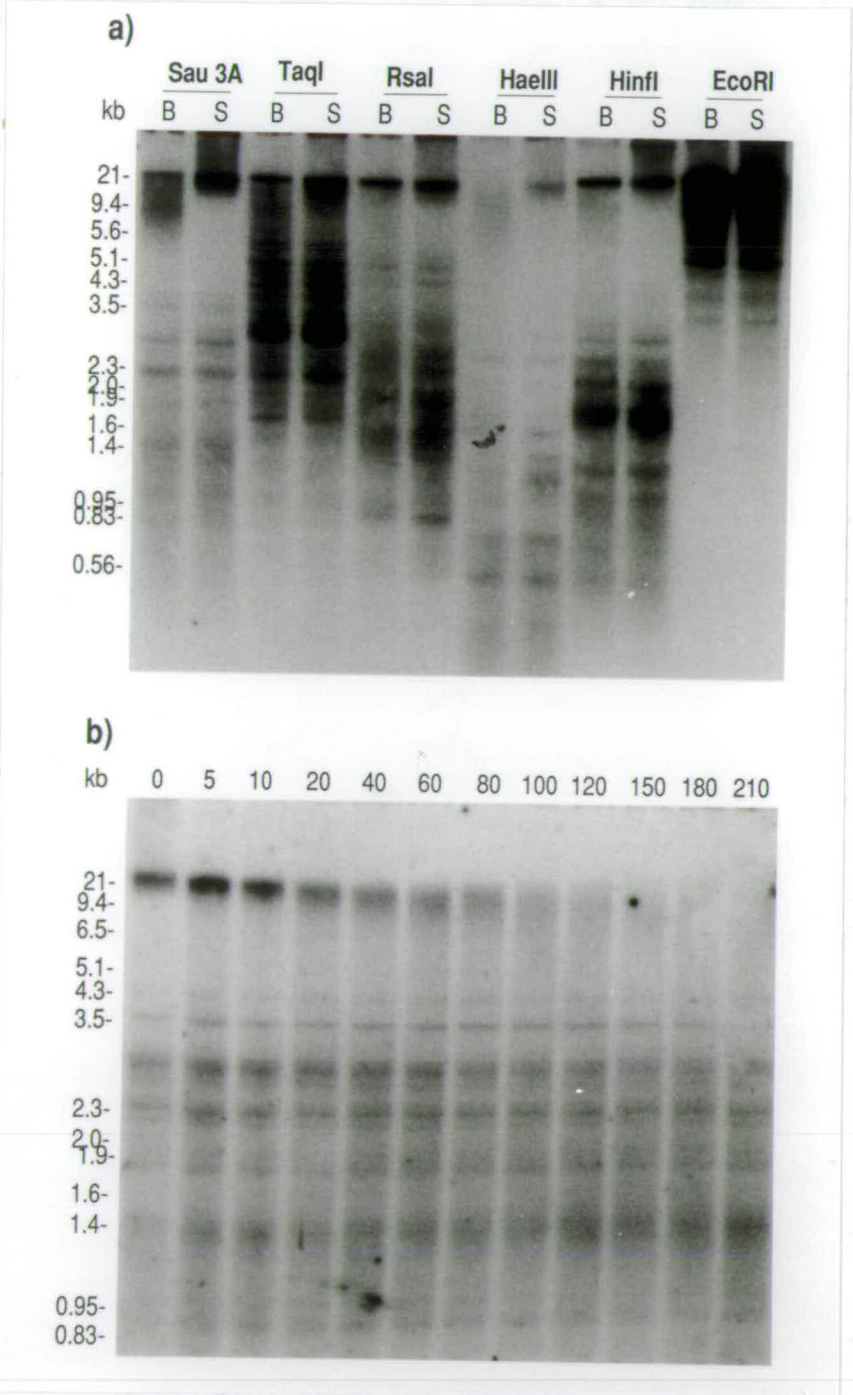
adjacent to the XP1 fragment, to digests of blood and sperm DNAs from the same individual is shown in Figure 6.8a and the result of hybridizing it to a human DNA Bal-31 series is shown in Figure 6.8b. This fragment is the same as XP1 in that it detects both heterogeneous and discrete restriction fragments and only the heterogeneous restriction fragments are sensitive to Bal-31. Some, but not all, of the fragments are detected by XP1. It is different from XP1 in that no blood/sperm differences, other than the size of hybridizing terminal restriction fragments, are detected with Sau3A, RsaI or TaqI. This fragment does not contain any homology to the terminal repeat sequence, TTAGGG, by sequence analysis (Cooke, pers. comm.).

Sequences between the NaeI site and the terminal repeats were also examined. As there are no convenient restriction sites between the NaeI site and the terminal repeats deletion derivatives were generated which no longer contained any of the terminal repeats. These were derived in the following way. First the 1.5 kb EcoRI/NaeI fragment from pHutel-2-end, which contains the terminal repeats and adjacent sequences, was transferred to the vector pGem7-7Zf(+). This was done as pGem7-7Zf(+) has convenient restriction sites flanking the SmaI and EcoRI cloning sites which can be used to generate nested deletions. EcoRI/NaeI digested pHutel-2-end DNA was run on a 1% LGT gel and the 1.5 kb fragment excised, ligated to EcoRI/SmaI digested pGem7-7Zf(+) vector, transformed into XL1-Blue cells by electroporation and plated onto Amp X-Gal plates. DNA was prepared from white colonies and a clone containing the correct insert was identified by restriction analysis and comparison to pHutel-2-end digests. As the EcoRI site is adjacent to the terminal repeats nested deletions were generated from this site. The clone was digested with

Figure 6.8 Hybridization Analysis of PstI-NaeI Fragment

a) Blood and sperm DNAs from the same individual were digested with the restriction enzymes shown, separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. B = blood DNA and S = sperm DNA.

b) Human sperm DNA was digested with Bal-31 for the times indicated (mins). The samples were then digested with the restriction enzyme Sau3A, 10µg aliquots separated by electrophoresis on a 1% agarose gel and then transferred to Hybond-N.
The filters were hybridized with ^{32}P-labelled PstI - NaeI fragment using the conditions described for XP1 in Figure 6.3.



the restriction enzymes SphI and EcoRI. SphI is near the EcoRI site in the polylinker and because it generates an end with a 3' overhang it is resistant to Exo III. ExoIII will degrade all types of free ends except those with a 3' extension. Nested deletions were produced as described in Materials and Methods, samples being taken at 5' time intervals from 0' up to 30'. DNA from two examples of each time point were prepared and hybridized with Try-4 to determine at which time point all the terminal repeats had been removed (data not shown). From this analysis the 20' time point was the earliest at which there was no detectable hybridization to Try-4. This was confirmed by sequence analysis (Cooke, pers. comm.). This clone was named END2TER20' and contains a ~700 bp insert which can be excised using the enzymes HindIII and ApaI. This HindIII/ApaI fragment was hybridized to blood/sperm digests but it only gave a weak signal. However it did appear to hybridize principally to heterogeneous restriction fragments in addition to some discrete ones (data not shown).

6.6 Discussion

XP1 has a high G + C content of 60%, which is 50% higher than the 40% found in bulk DNA. Sequences distal to XP1 have an even higher G + C content of 85% (Cooke, pers. comm.). It seems that the G + C content of telomere-adjacent sequences increases as the telomere is approached. High G + C content is one characteristic of CpG-rich islands. However, another is that the sequences are unmethylated in all tissues which is not the case with XP1. The XP1 sequence contains two copies of an almost perfect 36 bp repeat and a third more

degenerate copy. These internal repeats may account for the complex hybridization patterns that are seen with this sequence in spite of its comparative short length.

All the pro-terminal sequences investigated in this chapter detect internal as well as terminal restriction fragments, but only XP1 detects different hybridization patterns with blood and sperm DNAs using some restriction enzymes. The restriction enzymes with which these differences are seen either contain a CpG dinucleotide or their restriction sites can overlap a CpG dinucleotide. Methylation of overlapping CpG sequences accounts for the differences seen with Sau3A, as no differences are seen with its isoschizimer MboI which is not sensitive to methylation. Sensitivity to methylation has not been reported for the other two enzymes with which differences are seen, TaqI and RsaI. This suggests that these differences have a different molecular basis. The XP1 hybridization pattern of other tissues has been determined and some have the "sperm" pattern of hybridization with Sau3A, RsaI and TaqI digests (Cooke, pers. comm.). In most cases a tissue has either the "blood" pattern or the "sperm" pattern of hybridization with all three enzymes. If the RsaI and TaqI differences are not accounted for by methylation then the cause appears to be found in tandem with the appropriate methylation status. No differences have been seen with restriction enzymes that have a 6 bp recognition site other than ScaI, even those which are sensitive to methylation. This may be because the differences are close to telomeres and distal to the final site for these enzymes, with the result that the XP1 hybridizing fragments exhibiting blood/sperm differences are in the heterogeneous smear of hybridization and are therefore not apparent. However such an interpretation is difficult to

reconcile with the result that no XP1 hybridization differences are seen with the restriction enzyme PstI, a site for which is always present between XP1 homologous sequences and the telomere. As no differences are seen when sequences adjacent to XP1 are hybridized to blood and sperm DNA Sau3A, RsaI or TaqI digests the methylation, in the case of Sau3A, and the modifications responsible for the RsaI and TaqI differences must be very localised.

As yet the exact chromosomal location of any of these internal fragments has not been determined and it is not known if they are dispersed or clustered. It is possible that these pro-terminal sequences are found near more than one chromosomal end. This possibility will be investigated in the next chapter.

CHAPTER 7

DISTRIBUTION AND ARRANGEMENT OF PRO-TERMINAL SEQUENCES

7 DISTRIBUTION AND ARRANGEMENT OF PRO-TERMINAL SEQUENCES

7.1 Introduction

Sequences homologous to XP1 are present at more than one chromosome end and experiments described in this chapter attempt to determine how many, and which chromosome ends, contain these sequences. The human sequences in Hutel 1 consist of the terminal repeats and only the immediately sub-telomeric sequences because the cloning enzyme used was Sau3A which cuts frequently in the human genome. So that more telomere adjacent sequences, and additional telomeres, could be analysed another end library was made using BamHI, which cuts less frequently than Sau3A, as the cloning enzyme. End clones which cross-hybridize to XP1 were isolated from this library and the distribution and hybridization properties of the sequences flanking the XP1 homology were examined.

7.2 Construction of Human BamHI Telomere Library

In order to isolate more human telomeres, and longer arrays of telomere-adjacent sequences, new end libraries were made in the vector pYAC4NEONOT using BamHI as the cloning enzyme. Libraries were constructed using blood and sperm DNAs from the same individual so that telomeric regions derived from blood and sperm could be compared.

In the Sau3A end library construction described in Chapter 5 telomeric sequences were enriched for by size-fractionation of Sau3A digested DNA. As the average size of restriction fragments generated by BamHI is large, size-fractionation would not enrich greatly for

telomeric sequences. Instead BamHI digested blood and sperm DNAs were separated by isopycnic centrifugation on a Cs_2SO_4 gradient in the presence of silver ions. As the telomeric DNA is very GC-rich compared to bulk DNA it can be greatly enriched on such a gradient.

The vector pYAC4NEONOT was digested with NotI, dephosphorylated using BAP, and test ligations were done to check that it could no longer ligate to itself. This DNA was then digested with BamHI and test ligations were done to check that it could ligate to itself.

Ligation of prepared vector to fractionated BamHI digested human DNA and subsequent transformation into yeast were as outlined for the Sau3A library described in Chapter 5 except that the yeast strain used was YPH274 (Heiter, unpublished). The preparation of telomeric DNAs, the ligation and the transformation were done by Dr Cooke.

2000 Ura^+ colonies, 1000 from each transformation, were screened with Try-4 and XP1. 46 colonies derived from the blood library hybridized with Try-4, of which 23 also hybridized with XP1. 12 colonies derived from the sperm library hybridized with Try-4, of which 6 hybridized with XP1. There were no colonies which hybridized with XP1 and not with Try-4. Three of the blood derived clones, B4G7, B5F3 and B5H8; and two of the sperm derived clones, S3E6 and S3G5, hybridized strongly with XP1. These clones were chosen for further study.

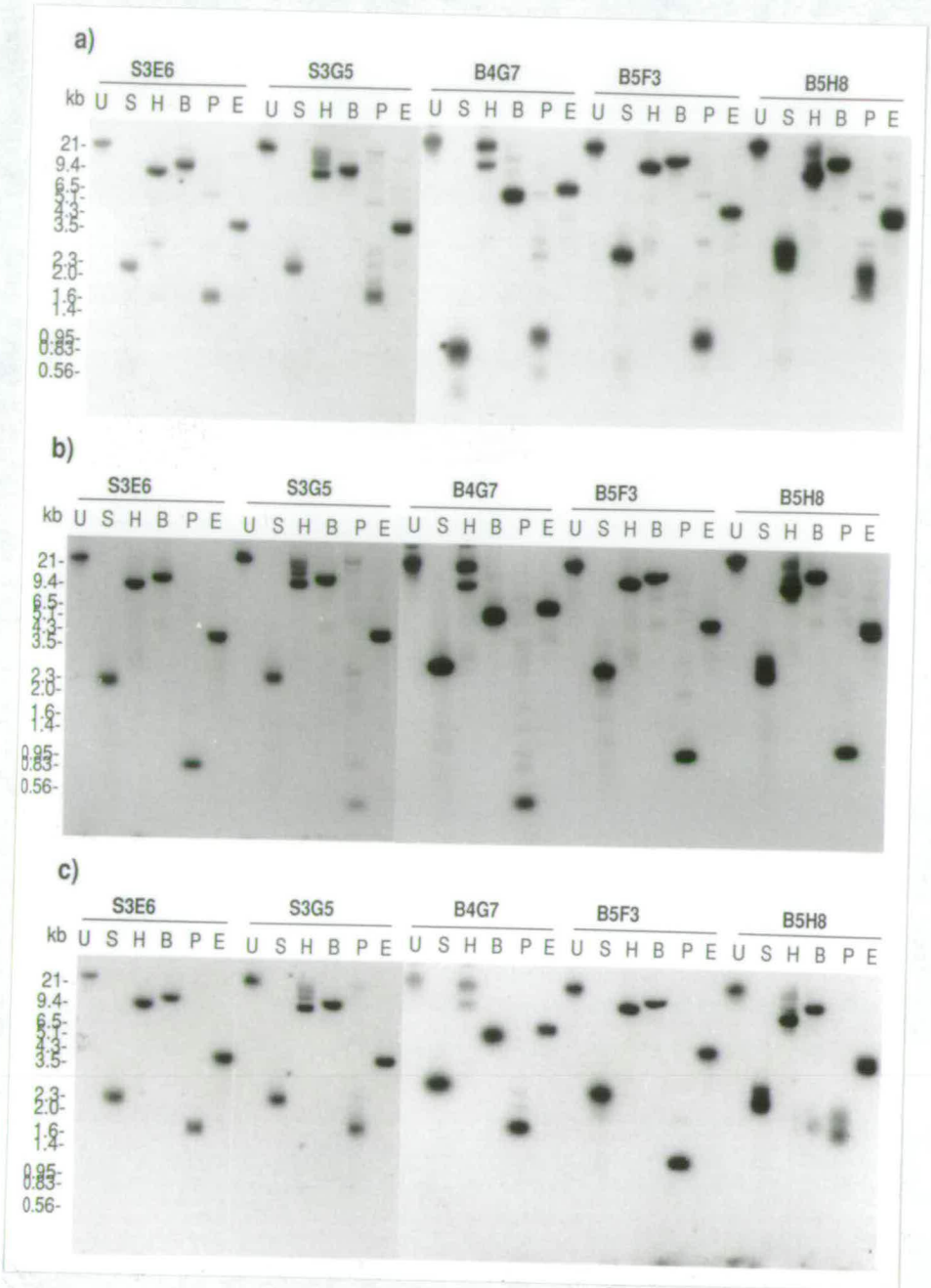
7.3 Analysis of Yeast Clones Containing XP1 Homology

DNA was prepared from the five yeast clones as described in Materials and Methods. Figure 7.1 shows the results of hybridizing a filter containing various digests of these clones with Try-4, XP1 and

the 333bp PstI - NaeI fragment derived from pHutel-2-end described in Chapter 6. Try-4 detected heterogeneous restriction fragments in all digests, consistent with it detecting terminal restriction fragments containing variable numbers of terminal repeats. The size of the heterogeneous fragments detected did not vary considerably between the blood and the sperm clones when the same restriction enzyme was used, nor did the intensity of hybridization. This implied that the terminal array of repeats had been shortened to less than 1 kb irrespective of whether the clones were derived from blood or sperm DNA. A similar abbreviation was observed in the clones Hutel 1 and Hutel 22. The size of the BamHI fragments cloned was 5.3 kb in B4G7 and between 9 and 10 kb for the other clones. B4G7 may contain telomeric sequences derived from a telomere where the BamHI site is comparatively close to the telomere or sequences may have been deleted during cloning. There are differences between the restriction maps of the different clones indicating that they were derived from different chromosomal ends. XP1 detected heterogeneous restriction fragments with some enzymes and discrete fragments with other enzymes. In all the clones there is a PstI site between the XP1 homologous sequences and the terminal repeats, as predicted by the result that XP1 detects only discrete restriction fragments in PstI digests of human DNA. All the clones contained sequences which cross-hybridized to the PstI - NaeI fragment derived from Hutel 1. These were located on the telomeric side of the XP1 cross-hybridizing sequences. This suggests that telomeric regions which contain XP1 cross-hybridizing sequences always contain similar adjacent sequences which are arranged in the same orientation with respect to the telomere.

Figure 7.1 Hybridization Analysis of Yeast Clones

DNAs from yeast clones B4G7, B5F3, B5H8, S3E6 and S3G5 were digested with the restriction enzymes shown, the digests separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The results of hybridizing the filter with the following ^{32}P-labelled probes are shown a) Try-4, b) XP1 and c) 333 bp PstI - NaeI fragment from pHute1-2-end. The hybridization conditions used are described in Materials and Methods. The washing conditions used with Try-4 were 4 x SSC 0.1% SDS at 65°C and the washing conditions used with XP1 and the PstI - NaeI fragment were 0.1 x SSC 0.1% SDS at 65°C. The filters were exposed to XAR-5 film. U = uncut, S = Sau3A, H = HindIII, B = BamHI, P = PstI and E = EcoRI.



Pulsed field gel analysis of the five clones showed that the Try-4 and XP1 cross-hybridizing sequences were present on short, linear chromosomes (data not shown). As described in Chapter 5 the mitotic stability of the artificial chromosomes can be tested easily because of the presence of the SUP4 gene on pYAC4NEONOT. Colonies containing artificial chromosomes are white and those which have lost them are red. When the five clones were streaked on non-selective medium the majority of the colonies were red in all cases, showing that the artificial chromosomes carrying the SUP4 gene had been lost. This is the expected result as short linear chromosomes are unstable in yeast.

7.4 Sub-cloning of Human Telomeric Sequences

In order to examine the human telomeric and sub-telomeric sequences further it was necessary to sub-clone the human sequences from the yeast clones into plasmids. The strategy used was as follows. To make the terminus susceptible to cloning it is necessary to remove the 3' overhang to generate flush ends. To do this the yeast DNA was treated briefly with Bal-31. An adaptor was ligated to this DNA as described in Materials and Methods. The adaptor used was formed by the two oligonucleotides, 5'-AGCTTGGCACCGCTCCGT-3'and 5'-ACGGAGCGGTGCCA-3' (synthesised by John Inglis) which can anneal to form an adaptor which is flush at one end and has a HindIII sticky end at the other end. The two oligonucleotides were examined by computer analysis to check that they could not form any secondary structure. After ligation to the adaptor the yeast DNA was digested with BamHI and the digested DNA separated by electrophoresis on a 1% LGT gel. Gel slices containing BamHI fragments of the size of those which contained the human

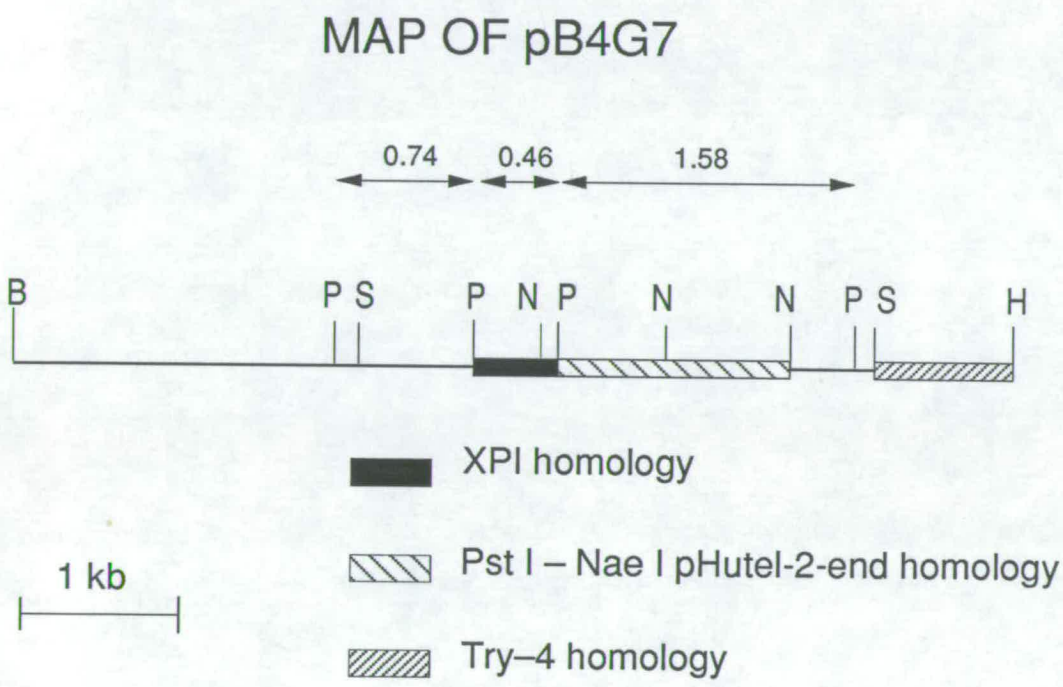
telomeric fragments were excised and the DNA purified from the agarose using agarase. The size-selected BamHI digested DNA was ligated to the vector pBS which had been digested with BamHI and HindIII and purified over a column so that the small BamHI - HindIII fragment from the linker region was eliminated. These ligations were transformed into the strain XL1-Blue by electroporation and plated onto agar plates containing ampicillin. Ampicillin resistant colonies were screened with Try-4 and positive clones were colony purified. These were then screened with XP1. One sub-clone was obtained from the B4G7 yeast which contained a BamHI - HindIII insert of the correct size and was named pB4G7. Other clones prepared from the other yeast clones contained sequences homologous to Try-4 but not XP1, sequences appeared to have been deleted and to have undergone rearrangment during the sub-cloning. As only 5.3 kb of human telomeric sequences were present in B4G7 they are probably more readily cloned than the larger fragments present in the other clones. It may be difficult to clone long arrays of telomeric sequences stably in plasmids because of their repeated nature.

7.5 Analysis of pB4G7

A restriction map of pB4G7 is shown in Figure 7.2. Sequences homologous to XP1, the 333bp PstI - NaeI fragment derived from pHutel -2-end and Try-4 are indicated. The orientation and order of the sequences homologous to these various probes is the same as that found in pHutel-2-end. One way that pB4G7 differs from pHutel-2-end is that there is a Sau3A site between the XP1 homology and the terminal repeats.

Figure 7.2 Restriction Map of pB4G7

Shown below is the restriction map of pB4G7 as deduced from double-digestion and hybridization analysis with Try-4, XP1 and the 333 bp PstI - NaeI fragment from pHutel-2-end. Restriction fragments homologous to these sequences are indicated. There are additional Sau3A sites which have not been mapped between the BamHI site and the Sau3A sites shown. The hybridization and washing conditions used were as described in Figure 7.1. B = BamHI, H = HindIII, N = NaeI, P = PstI and S = Sau3A.



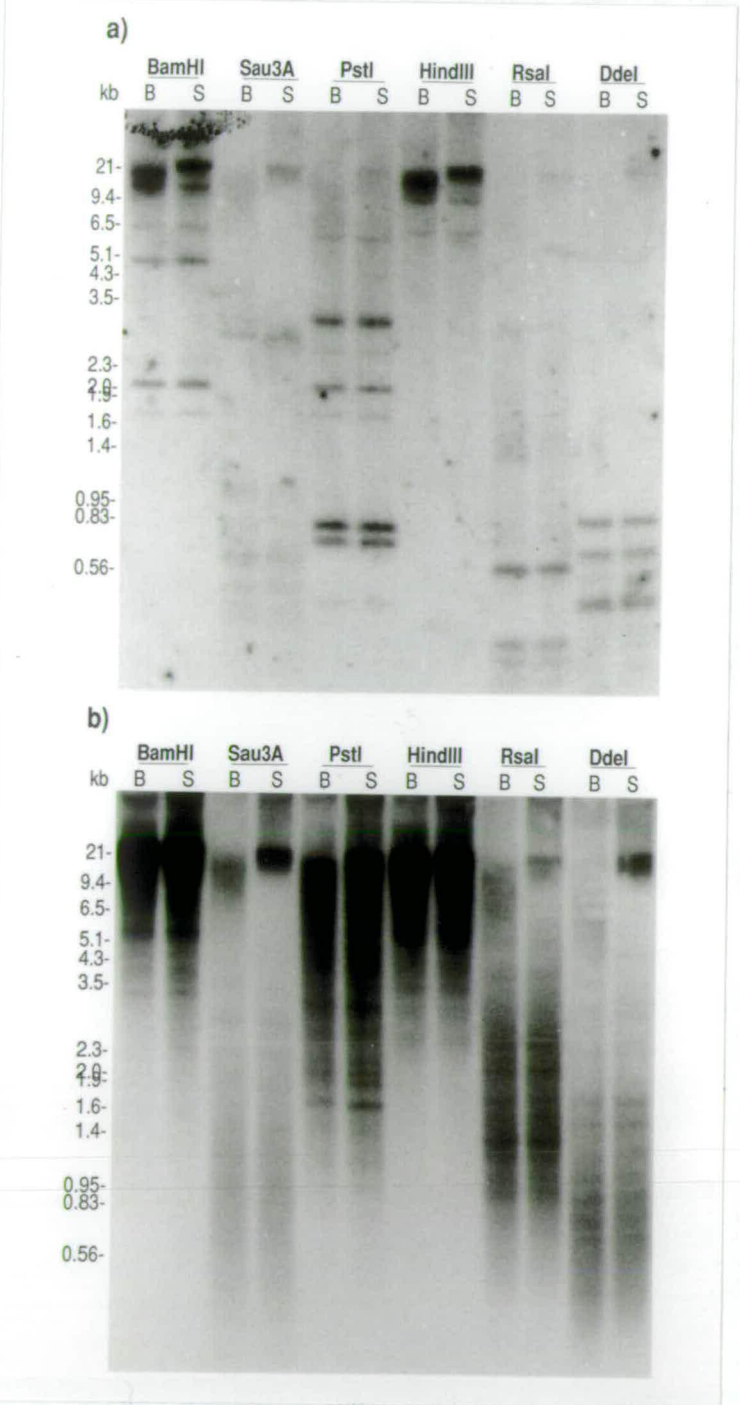
In order to determine whether any yeast telomeric repeats had been added to the human telomeric repeats restriction digests of pB4G7 and Dirt3, which contains at least 58 tandem repeats of TTAGGG (Forney et al., 1987), were hybridized with an oligonucleotide corresponding to the G-rich strand of terminal repeats isolated from the yeast S. cerevisiae. The sequence of this oligonucleotide is 5'-GTGTGTGGGTGTGGGTGTGGTGTGGTGTGTGGGTGTGTGG-3' (a gift of Dr Greider). This oligonucleotide hybridized to pB4G7 restriction fragments which also hybridized with Try-4 and it did not hybridize to Dirt3 showing that it does not cross-hybridize to tandem repeats of TTAGGG. This suggested that the human telomeric repeat TTAGGG had acted as a primer for the addition of yeast telomeric repeats. This interpretation was confirmed by sequence analysis by Dr Cooke. This result confirms that human telomeres, like those of Tetrahymena, can support telomere function in yeast.

7.6 Hybridization Analysis of pB4G7 Sequences

Figure 7.3 shows the result of hybridizing the 0.74 kb PstI fragment and the 1.58 kb fragment from pB4G7 which flank the 0.46 kb fragment containing homology to XP1 with various digests of blood and sperm DNAs from the same individual. Both probes detect both discrete and heterogeneous restriction fragments. They do not, however, detect any differences between blood and sperm digests when the same restriction enzyme is used, except for differences that can be accounted for by the longer array of telomeric repeats found on sperm telomeres. As these fragments are immediately adjacent to the XP1 homology in pB4G7 the methylation and/or other modifications which are

Figure 7.3 Hybridization of XP1-Adjacent Sequences with Blood and Sperm Digests

Blood and sperm DNAs from the same individual were digested with the restriction enzymes shown, separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The results of hybridizing the filter with the two PstI fragments from pB4G7 flanking the XP1 homology are shown; a) ^{32}P-labelled 0.74 kb fragment and b) ^{32}P-labelled 1.58 kb fragment. The hybridization and washing conditions used were as described for XP1 in Figure 7.1.
S = sperm DNA and B = blood DNA.



responsible for the differences between blood and sperm digests seen with XP1 must be highly localised. The centromere-proximal BamHI - PstI fragment from pB4G7 was also used as a hybridization probe. This fragment detected an intense heterogeneous smear of hybridization indicating that it contains sequences which are highly repeated in the human genome. These may be the major SINE repeat Alu, which is a 300 bp sequence repeated approximately a million times in the human genome (Rogers, 1985).

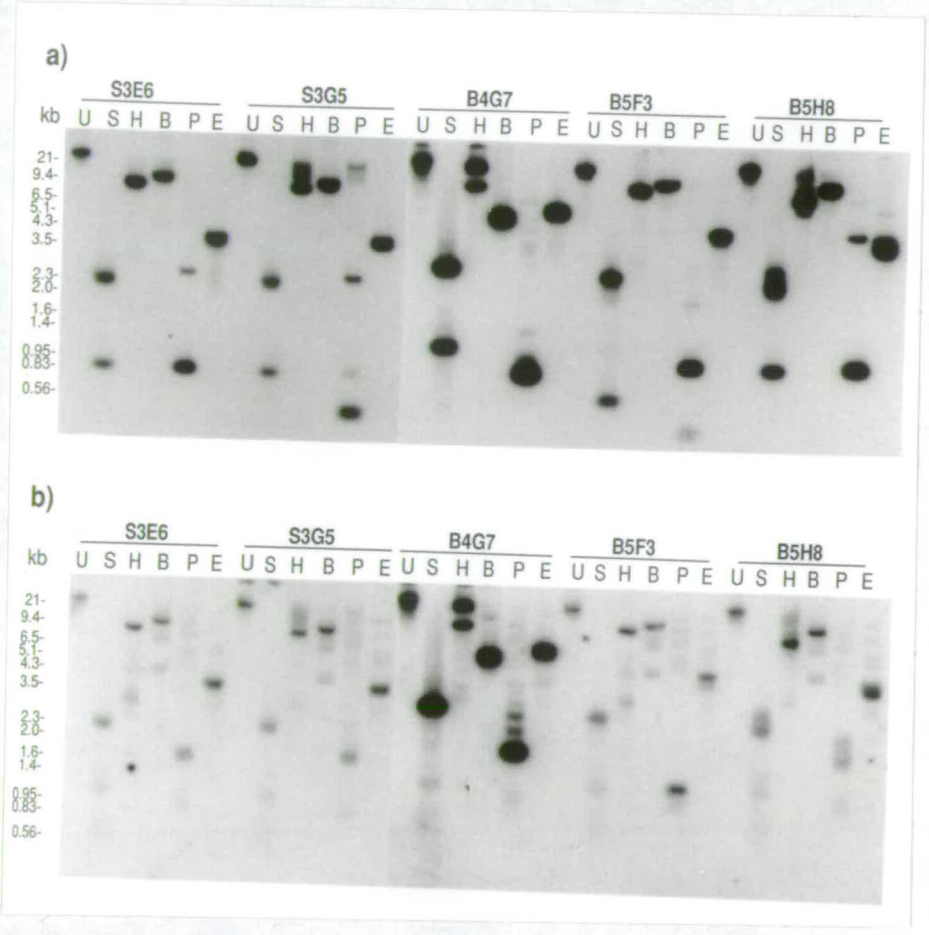
7.7 Distribution of Sub-Telomeric Sequences

Of the end clones from the BamHI end library which hybridize with Try-4 only half hybridize with XP1. This suggests that sequences homologous to XP1 are present adjacent to telomeres on approximately half the chromosome ends. Figure 7.4 shows that both the 0.74 kb PstI fragment and the 1.58 kb PstI fragment hybridize to the five yeast clones which contained sequences homologous to XP1. The intensity of hybridization of the 1.58 kb fragment to pB4G7 is greater than that to the other clones indicating that the extent of homology to this fragment in these other clones is not great. This result supports the suggestion that XP1 and adjacent sequences have a conserved distribution at telomeres, but there are some variations.

A number of different experiments were carried out to determine which chromosomes ends contain sequences homologous to XP1. XP1 was hybridized with Sau3A and RsaI digests of DNA from mouse/human hybrids which contain only the human X or Y chromosome and of DNA from a human lymphoblastoid cell-line. XP1 only hybridized with DNA from the human

Figure 7.4 Hybridization of XP1-Adjacent Sequences with Yeast Clones

The results of hybridizing the filter containing the various digests of the yeast clones B4G7, B5F3, B5H8, S3E6 and S3G5 used in Figure 7.1 with the two PstI fragments flanking the XP1 homology in pB4G7 are shown; a) ^{32}P-labelled 0.74 kb fragment and b) ^{32}P-labelled 1.58 kb fragment. The hybridization and washing conditions used were as described for XP1 in Figure 7.1. U = uncut, S = Sau3A, H = HindIII, B = BamHI, P = PstI and E = EcoRI.



lymphoblastoid cell line suggesting that the terminal parts of the X and Y chromosomes do not contain sequences homologous to XP1 (data not shown).

The three PstI fragments of 0.74 kb, 0.46 kb and 1.58 kb from pB4G7 were used for in situ hybridization to metaphase spreads. pB4G7 was digested with PstI and the digest run on a 1% LGT gel. The three PstI fragments were excised and the DNA purified. To eliminate any of the large PstI fragment which contains the terminal repeats and the part of the insert containing repeated DNA which had not been separated away on the first gel this process was repeated twice. Try-4 did not hybridize to the purified fragments (data not shown) showing that they were not contaminated with the PstI fragment containing the terminal repeats. When the purified fragments were hybridized to various digests of human DNA they detected terminal restriction fragments, in addition to discrete fragments, but they did not appear to contain any highly repeated sequences and so were a suitable probe for in situ hybridization (data not shown). The in situ hybridization was carried out by Judy Fantes. The fragments hybridize principally to a few chromosome ends and they did not appear to hybridize consistently to any internal sites. The most consistent hybridization was to both ends of chromosome 20 and to the tip of the short arms of chromosomes 4 and 22. Table 7.1 shows the accumulated data from 19 metaphase spreads which were examined. From the proportion of end clones that contain sequences homologous to XP1 it might have been expected that a greater number of chromosome ends would have been detected. There are two main reasons why only a few ends were detected. Either only a few chromosomal ends have homology to XP1 and adjacent sequences and the use of BamHI in the generation of the end

Table 7.1 *in situ* Hybridization of PstI Fragments of pB4G7

The three PstI fragments from pB4G7 described in the text were used for *in situ* hybridization. The table below shows the accumulated data from 19 metaphase spreads which were examined. The spreads were first hybridized with biotinylated pb4G7 PstI fragments visualised by FITC-avidin and were subsequently DAPI stained so that the individual chromosomes could be identified. A position was counted as positive only when there was hybridization to both sister chromatids.

CHROMOSOME	NUMBER OF POSITIVE TELOMERES	
	p	q
1	-	1
2	-	-
3	-	1
4	12	-
5	-	1
6	3	1
7	1	-
8	-	-
9	2	2
10	-	-
11	1	-
12	-	-
13	-	1
14	-	-
15	-	-
16	-	-
17	1	-
18	-	-
19	-	-
20	2	3
21	-	-
22	4	1
X	-	-
Y	-	-

INTERSTITIAL SITES	NUMBER OF POSITIVE SIGNALS
mid 19 or 20q	1
mid 11q	1
mid 2p	1
13cen	1
mid 6q	1
mid 4p	1

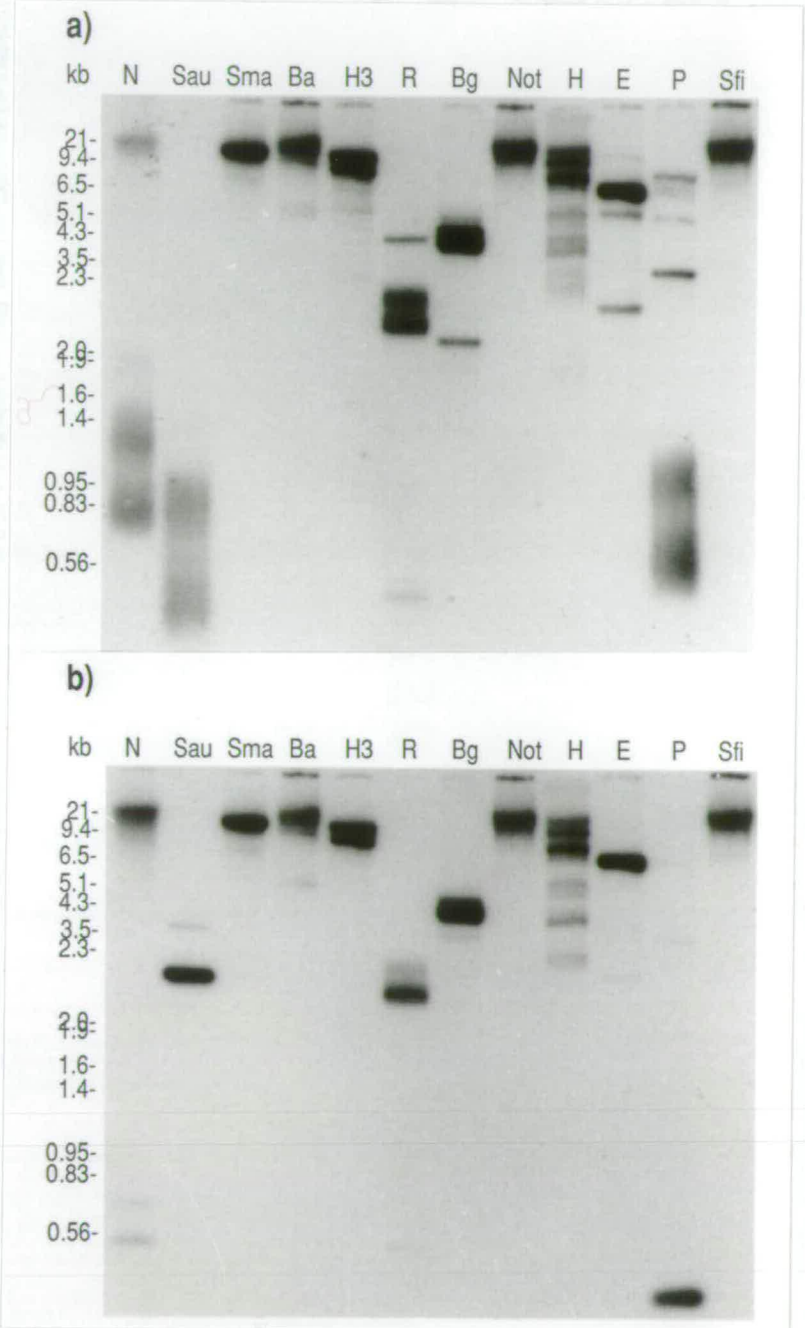
clones biased towards these, or 50% of chromosomal ends have homology to XP1 and adjacent sequences but contain different amounts and the probe used only detects those ends with extensive homology. It would be interesting to determine how many end clones generated using other enzymes contain sequences homologous to XP1. Once more human telomeres have been analysed longer telomere-adjacent sequences can be isolated and used in in situ hybridization which might provide clearer results. Mouse/human hybrid cell-lines containing single human chromosomes could also be used.

7.8 Sequences Homologous to XP1 are Present at the 4p Telomere

A yeast clone Y88BT, containing the terminal BssHI fragment from human chromosome 4p, has been generated (Bates et al., submitted). Figure 7.5 shows the result of hybridizing various digests of Y88BT DNA with Try-4 and XP1. Y88BT contains sequences homologous to both probes. Try-4 detects heterogeneous restriction fragments in all digests as expected. XP1 detects both heterogeneous and discrete restriction fragments so it is located immediately adjacent to the telomere. There is only one region of homology to XP1 which is contained on a ~400 bp PstI fragment. This suggests that XP1 is not part of a tandemly repeated array and, on this chromosome end at least, there are no other regions of homology to XP1 up to 100 kb from the end. The same filter was hybridized with the two PstI fragments which flank the XP1 homology in pB4G7 (data not shown) and they too detected cross-hybridizing sequences which also have a terminal location. This result confirms the in situ result which suggested that XP1 and adjacent sequences were present at the tip of 4p.

Figure 7.5 Hybridization of Try-4 and XP1 to Y88BT

Y88BT, a yeast clone containing the terminal 100 kb BssHI fragment from human chromosome 4p (Bates et al., submitted), was digested with the restriction enzymes shown, the digests separated by electrophoresis on a 0.8% agarose gel and transferred to Hybond-N. The results of hybridizing the filter with a) ^{32}P-labelled Try-4 and b) ^{32}P-labelled XP1 are shown. The hybridization and washing conditions used were as described in Figure 7.1. N = NaeI, Sau = Sau3A, Sma = SmaI, Ba = BamHI, H3 = HindIII, R = RsaI, Bg = BglII, Not = NotI, H = HinfI, E = EcoRI, P = PstI and Sfi = SfiI.



7.9 pHutel-2-end and pB4G7 do not Contain Arrays of TTGGGG Repeats

Work by Allshire *et al.* (1989) has suggested that human telomeres contain at least three type of repeat arranged in an ordered fashion. The proposed arrangment is a mixture of blocks of TTGGGG and TGAGGG repeats adjacent to a block of TTGGGG repeats which is adjacent to the terminal TTAGGG repeats. The human terminal sequences in the clones pHutel-2-end and pB4G7 were examined to see if they contained arrays of TTGGGG repeats. A filter containing various digests of pHutel-2-end, pB4G7, Dirt3, which contains tandem TTAGGG repeats, and pYAC4NEONOT, which contains ~700 bp of tandem TTGGGG repeats, were hybridized with Tet-4, the $(TTGGGG)_4$ oligonucleotide. 300 ng of unlabelled Try-4 was included in the hybridization to compete out cross-hybridization between Tet-4 and tandem arrays of TTAGGG repeats. Tet-4 hybridized only to pYAC4NEONOT (data not shown). The human telomeric regions contained in the clones pHutel-2-end and pB4G7 do not contain arrays of TTGGGG repeats and do not conform to the suggested arrangement of G-rich repeats suggested by Allshire *et al.*. It may be that TTGGGG repeats were present on the original human telomeres and were deleted either during the yeast cloning or during the sub-cloning into plasmids, but this is unlikely as such repeats are maintained on artificial chromosomes in yeast and in the plasmid pYAC4NEONOT. A greater number of human telomeres need to be analyzed in order to determine how many human telomeres do contain arrays of these other G-rich repeats in addition to the terminal TTAGGG repeats which are present on all chromosome ends.

7.10 Discussion

The analysis of pB4G7 which showed that an array of yeast telomeric repeats had been added to the human TTAGGG repeats confirms the suggestion that human telomeres can support telomere function in yeast. The human telomeric repeat appears to have acted as a primer for the addition of these repeats. The terminal array of TTAGGG repeats is abbreviated to less than 1 kb on cloning into *S. cerevisiae*. This occurs irrespective of whether the human telomere is derived from blood DNA or from sperm DNA where the terminal array of repeats is ~5 kb longer. Possible reasons for this abbreviation have been discussed already in Section 5.6.

The human end clones which contained homology to Try-4 and to XP1 all contained homology to sequences immediately adjacent to XP1 and these sequences were always arranged in the same fashion. Methylation, and any other modifications, responsible for the differences detected by XP1 between blood and sperm DNA digests from the same individual must be highly localised as no differences are detected with sequences adjacent to XP1. XP1 and associated sequences appear to represent one class of telomere-adjacent sequences as not all human telomeres contain these sequences. Analysis of end clones from the BamHI end library which do not contain any homology to XP1 will probably reveal additional classes of telomere-adjacent sequences.

The presence of shared sequences between different chromosomal ends, in addition to the terminal repeats, raises the possibility that exchange of sequences may occur between different chromosome ends. This is a well documented phenomenon in yeast (Horowitz *et al.*, 1984).

A greater number of human telomeres need to analysed in order to establish if such exchanges similar to those which occur between yeast chromosomal ends do happen.

A recent paper from Cheng *et al.* (1989) describes the characterization of a human telomere isolated by complemention in yeast in a similar way to that described in this thesis. The major difference was that the cloning enzyme used was EcoRI. The structural organisation of the telomere described is similar to that of the telomere cloned in pB4G7. There is a 4 kb sequence immediately adjacent to the terminal repeats which is found at most, but not all, human telomeres. The most centromere-proximal region cloned contained AT-rich sequences and an *Alu* repeat. This organisation of sequences is the same as that found in pB4G7. The telomere-adjacent sequences cloned by Cheng *et al.* may represent a different class of telomere-adjacent sequences to those which cross-hybridize to XP1.

As telomere-adjacent sequences vary between different chromosome ends they are unlikely to perform a crucial role in telomere function. They may, however, be involved in the various telomere-telomere associations which have been observed.

CHAPTER 8

DISCUSSION

8 DISCUSSION

The work described in this thesis, and experiments by other workers, show that telomeres of human chromosomes share many conserved features with those of lower eukaryotes. In particular there are many similarities in the organisation and type of DNA sequences found at and near the ends of chromosomes. Generally telomeric regions can be divided into two main components, terminal regions consisting of arrays of simple repeats which are necessary for replication and maintenance of telomeres and sub-terminal regions which in some species consist of more complex repeats. These sub-terminal regions are often recombinogenic and may be involved in the various telomere-telomere associations that occur.

The arrangement of the terminal sequences is well conserved between humans and lower eukaryotes. In all cases there is an array of a variable number of simple, G/C rich repeats where the G's and C's are usually segregated onto different strands, the G-strand runs 5' - 3' towards the telomere and at the terminus there is a 3' extension of this strand. The exact sequence of the telomeric terminal repeat varies between species but the main repeat sequence found at human telomeres is the same as that found in Trypanosoma.

Immediately pro-terminal regions on human chromosomes consist of more complex G/C rich repeats. The pro-terminal sequences analysed so far are present on some, but not all, chromosomes. There appear to be several classes of these pro-terminal sequences which are shared between different chromosome ends. A similar arrangement of pro-

terminal sequences is found in the yeast S. cerevisiae (Chan and Tye, 1983a). The most proximal human pro-terminal sequences are A/T rich and may contain Alu repeats.

8.1 Terminal Sequences

The first suggestion that the DNA sequences found at human telomeres shared common features with those of lower eukaryotes came from analysis done with the sequence 29C1, derived from close to the pseudoautosomal telomere, which showed that terminal restriction fragments were heterogeneous (Cooke et al., 1985). This heterogeneity suggested that variable numbers of repeats were present at different pseudoautosomal telomeres. The nature of these repeats was unknown, but the results of experiments which used 29C1 to monitor the effect of various nucleases on the end suggested that a 3' extension is found at the terminus (Cooke, 1987). Human telomeres therefore shared at least two common features with those of lower eukaryotes, an array of a variable number of repeats with a 3' extension at the terminus.

The experiments described in Chapter 4, and in Allshire et al. (1988), show that the Tetrahymena telomeric repeat TTGGGG cross-hybridizes to human telomeres. Therefore human telomeres share not only the same arrangement of repeated DNA at the terminus but also the sequence of the telomeric repeat is conserved. In addition other simple telomeric repeats were shown to cross-hybridize to human telomeres. A human repetitive sequence TTAGGG, which is identical to the telomeric repeat of Trypanosoma, cross-hybridizes to human telomeres (Moyzis et al., 1988) as does the Arabidopsis telomeric repeat TTTAGGG (Richards and Ausubel, 1988).

These cross-hybridization experiments showed that the sequence and arrangement of the terminal repeats found at human telomeres are conserved with those of lower eukaryotes. The structure of human telomeres is conserved to the extent that human telomeres can be expected to function in yeast. This is in fact the case and this ability to complement telomere function allowed human telomeric sequences to be isolated as described in Chapter 5. Analysis of these human sequences in yeast showed that, as expected, an array of TTAGGG repeats was present at telomere ends orientated such that the G-rich strand runs 5' - 3' towards the terminus. Sequence analysis showed that the principle repeat sequence in this array was TTAGGG but variants on this sequence are also present, for example TTGGGG, (Cooke, pers. comm.). Variation in the telomeric repeat sequence is a feature shared with other species. In Plasmodium the telomeric repeat sequence is TTTAGGG or TTCAGGG and other variants are occasionally present (Ponzi et al., 1985; Vernick and McCuthan, 1988). The length of the repeat unit can also vary, for example in the yeast S. cerevisiae telomeric repeats have an irregular sequence of TG_{1-3} (Shampay et al., 1984).

The human terminal repeats were considerably abbreviated during the cloning procedure into yeast and possible reasons for this have been discussed already in Chapter 5. Only ~1 kb of human terminal repeats were present in the yeast clones whereas in human cells the terminal "barren" region, presumably consisting of the simple telomeric repeats, extends for upwards of 5 kb. Therefore it is likely that the human telomeric sequences which were cloned represent the most proximal part of the terminal array. The distal part of the array is the region most likely to be important for telomere replication and

maintenance. Therefore alterations in the sequence of the most proximal repeats, by processes such as mutation and strand-slippage, might not be detrimental and could therefore be maintained. Also the most proximal repeats in the array would be less often corrected by recombination or de novo telomere addition. Therefore mutations in this region could have possibly accumulated with time. It would be interesting to determine if the sequence of the most distal repeats is more uniform.

Hybridization (Chapter 7) and sequence analysis (Cooke, pers. comm.) showed that an extensive tract of yeast telomeric repeat had been added to the human telomeric repeat present at the end of the artificial chromomsome in the clone B4G7. Therefore human telomeres are capable of being recognised as a functional telomere in yeast, the human telomeric repeat acting as a primer for the addition of yeast-type telomeric repeats. This result demonstrates that structural and functional aspects of telomere organisation are conserved between lower and higher eukaryotes.

The terminal array of telomeric repeats is about 5 kb longer in germ-line cells than in somatic cells (Cooke and Smith, 1986; Cross et al., 1989). It is possible that the extra telomeric repeats present on germ-line chromosomes have an important germ-line function prehaps during meiosis, pairing and synapsis of chromosomes during zygotene usually starts at the telomeres (Chandley, 1986). It is also possible that the difference is merely a consequence of the fact that the mechanism for maintaining the telomeric repeats is less efficient in somatic as opposed to germ-line cells.

Recently a human terminal transferase, or telomerase, activity has been identified in crude extracts of HeLa cells (Morin, submitted). This activity is a ribonucleoprotein and is similar to the telomerase activity of Tetrahymena described by Greider and Blackburn. The human telomerase activity adds many copies of the repeat TTAGGG onto oligonucleotide primers corresponding to the G-rich strand of various telomeric repeats. This suggests that the method by which telomeres are replicated, in addition to the sequence and structure of telomeres, is conserved between lower and higher eukaryotes. In addition the fact that the sequence added was TTAGGG confirms that this sequence is indeed the human telomeric repeat.

The hybridization of the Arabidopsis telomeric repeat, TTTAGGG, is probably due to cross-hybridization of this repeat to TTAGGG (Chapter 4 and Allshire et al., 1989). The experiments described in Allshire et al. suggested that blocks of repeats consisting of the sequence TTGGGG and TGAGGG are present at human telomeres immediately adjacent to to the terminal array of TTAGGG repeats. Blocks of these repeats have not been found in the human telomere clones examined so far. This apparant conflict of results will only be resolved once a greater number of human telomeres have been cloned and sequenced.

8.2 Sub-Telomeric Regions

The strategy of cloning human telomeres in yeast has allowed the immediately sub-telomeric regions, in addition to the terminal repeats, to be isolated. Examination of the pro-terminal sequences cloned has yielded the surprising result that such sequences are shared between different chromosomal ends. Pro-terminal sequences

related to XP1 may be present on up to half the telomeres, judging by the proportion of yeast end clones which contain cross-hybridizing sequences and the results of *in situ* experiments. However, in contrast to the terminal repeats the pro-terminal sequences examined so far are not found in species more phylogenetically diverged than the chimpanzee.

The human telomeric clones which have been studied in this thesis share related pro-terminal sequences. From hybridization experiments the different clones share extensive regions of homology but there are differences in the restriction maps of the different clones suggesting that they were derived from different chromosomal ends. They do share some common features however, for example in all cases there is a terminally located PstI site.

Sequence analysis has shown that the pro-terminal sequences fall into two distinct regions. The immediately pro-terminal sequences are G/C rich, the G/C content rising as the terminal repeats are approached (Cooke, pers. comm.). In contrast the sequence of the most proximal pro-terminal region examined is A/T rich and may contain *Alu* repeats (Chapter 7 and Cooke, pers. comm.).

The pro-terminal region best characterised, the XP1 fragment discussed in Chapter 6, represents the most proximal part of the G/C rich region. XP1 has a G/C content of 60% and contains 3.5 copies of an imperfect 36 bp repeat. Hybridization and Bal-31 analysis showed that sequences homologous to XP1 are present at internal as well as telomeric locations. XP1 shows different hybridization patterns of the discrete internal fragments between sperm and blood DNA digests from the same individual when some restriction enzymes are used. Some of these differences are due to under-methylation of sperm as compared to

blood DNA but the molecular basis of other differences remains unexplained. Whatever the modifications are which are responsible for these differences they must be highly localised as sequences immediately adjacent to XP1 do not exhibit similar hybridization differences. It is possible that these different sperm and blood hybridization patterns which are not accounted for by methylation indicate the presence of as yet undescribed modifications of germ-line DNA which may have functional significance.

As pro-terminal sequences are shared between several chromosomal ends it is possible that exchanges could occur between different ends. Such exchanges have been described in S. cerevisiae (Horowitz et al., 1984). As yet it is not known if XP1 homologous sequences are restricted to certain chromosomal ends or if they can be present on different ends. Analysis of the distribution of XP1 homologous sequences in a number of different cells from different origins by in situ hybridization will answer this question. If XP1 homologous sequences are not restricted to certain ends this would suggest that exchanges can occur between different chromosomal ends.

There are probably several different classes of pro-terminal sequences which are present at chromosomal ends. The number of these classes and their sequence composition await further analysis. However, as it is clear that different chromosomal ends have sub-telomeric regions which contain different sequences, it is unlikely that these regions are important for telomere maintenance. The terminal repeats, which are present at all telomeres, are more likely to fulfil this function. The sub-telomeric regions are probably important for other aspects of telomere function, such as telomere-telomere interactions. The fact that the most proximal part of the

sub-telomeric region is A/T rich is interesting because A/T rich DNA is known to associate with both the nuclear matrix (Flickinger, 1986) and to the nuclear scaffold (Mirkovitch et al., 1984), both structures with which telomeres are known to associate.

8.3 Future Work

The isolation of human telomeres is important for many reasons. Sub-telomeric sequences which can recognize specific chromosomal ends will be important in closing the genetic map because they will provide markers for the ends of chromosomes. A particular problem for which this is important is the cloning of the gene responsible for Huntington's disease (the HD gene). A polymorphic DNA marker genetically linked to the gene has been isolated (Gusella et al., 1983). This locus has been localised to chromosome 4p16, which is the last cytologically detectable band before the telomere (Wang et al., 1985). Linked markers have only been isolated from the centromeric side of the gene and it has so far proved impossible to obtain flanking markers from the telomeric side (Gilliam et al., 1987). The isolation of human 4p telomeric sequences by complementation in yeast from an individual homozygous for HD (Bates et al., submitted) should provide suitable flanking markers which will be useful in the isolation of the HD gene.

Telomere-binding proteins have been found in many species, for example the protein heterodimer which binds tenaciously to Oxytrichia telomeres (Price and Cech, 1989). It would be interesting to study any proteins which are associated with human telomeres. Telomeric sequences could be used in a gel retardation assay to determine if

there are any proteins which preferentially bind to them (Strauss and Varshavsky, 1984). If there are proteins bound to human telomeres they may protect the end from degradation by exonucleases. Whether Bal-31 can degrade telomeres in nuclei could be investigated and also whether there are any differences during the cell-cycle in the sensitivity of telomeres to exonucleases by examining telomeres in nuclei isolated from a synchronous cell population at different stages.

Telomeres would provide one of the necessary components of an artificial chromosome. The other components necessary are an origin of replication, a centromere and possibly a nuclear retention factor similar to that of Epstein-Barr virus which consists of 20 tandem copies of a 30 bp repeat and allows linked DNA to be maintained for long periods after introduction into cells. Recently the isolation of human sequences which can act as origins of replication has been reported (Krysan et al., 1989). Other workers have reported similar findings but with very different DNA sequences (Yates and Guan, pers. comm.). Therefore it is still unclear what origins of replication in human cells actually consist of. The construction of such an artificial chromosome would provide a useful vector for investigating meiotic and mitotic behaviour of chromosomes and for investigating the effect of mutations on that behaviour. It might also provide a means of introducing cloned genes into human cells.

Telomeres of mouse chromosomes, like those of human chromosomes, consist of arrays of simple repeats which cross-hybridize to the Trypanosoma telomeric repeat TTAGGG (see Chapter 4). The number of repeats present at mouse telomeres is far greater than that present at human telomeres, irrespective of the tissue source. Telomeres of mouse chromosomes have a repeat array of a least 20 kb in cell-lines and the

somatic tissues which have been examined, as judged by the resolution capacity of conventional gels. In spite of this difference between human and mouse telomeres the pseudoautosomal telomere of the human X chromosomes present in the X-MGU mouse/human hybrids has retained the characteristic human length of the terminal array and has not been converted to a mouse-type of array length (see Chapter 3). Initial pulsed-field gel data has supported this finding and has shown that the repeat array of the mouse telomeres in the mouse and the hybrid cell-lines is very large. The human X chromosome pseudoautosomal telomere does not appear to be rearranged in any gross way in the mouse/human hybrids and as it is stable it is probably functional even though the terminal array is considerably shorter than that found on the mouse telomeres. Why mouse telomeres have such a long terminal array and whether the human telomere in a mouse cell background is truely functional requires further investigation.

8.4 Conclusions

I have described experiments designed to determine the structure of human telomeres and to isolate the DNA sequences found in this region of the genome which had not been described previously.

The initial cross-hybridization experiments revealed that the type of DNA sequences present at human telomeres were very similar to those of lower eukaryotes and that the organisation of those sequences was conserved. These findings suggested a method for cloning human telomeres by complementation of telomere function, as it was already known that telomeric sequences from several organisms, which were similar to those found at human telomeres, could complement for

telomere function in yeast. Human telomeres too could successfully complement for telomere function in yeast and this allowed the isolation of human telomeric sequences.

Analysis of the human telomeric sequences showed that human telomeres consist of two distinct regions. An array of a variable number of simple TTAGGG repeats is found at the extreme end of all human chromosomes. These repeats are probably important for telomere maintenance and replication. The sequence of these repeats and their organisation are evolutionary conserved which implies that only a limited number of sequences can fulfil telomere functions. Sub-telomeric regions consist of sequences which can be shared between different chromosomal ends and these sequences may be important in telomere-telomere interactions and be involved in recombination between ends. The sequence of these pro-terminal regions appears to be species-specific. In conclusion I think that it is fair to state that we have finally come to the end !.

REFERENCES

REFERENCES

Agard, D. A. and Sedat, J. W. (1983). Three-dimensional architecture of a polytene nucleus. Nature 302, 676-681.

Allshire, R. C., Dempster, M. and Hastie, N. D. (1989). Human telomeres contain at least three types of G-rich repeat distributed non-randomly. Nucl. Acids Res. 17, 4611-4627.

Allshire, R. C., Gosden, J. R., Cross, S. H., Cranston, G., Rout, D., Sugawara, N., Szostak, J. W., Fantes, P. A. and Hastie, N. D. (1988). Telomeric repeat from T. thermophila cross hybridizes with human telomeres. Nature 332, 656-659.

Anand, R. (1986). Pulsed field gel electrophoresis: a technique for fractionating large DNA molecules. Trends Genet. 2, 278-283.

Baker, B. S., Carpenter, A. T. C., Esposito, M. S., Esposito, R. E. and Sandler, L. (1976). The genetic control of meiosis. Ann. Rev Genet. 10, 53-134.

Baroin, A., Prat, A. and Caron, F. (1987). Telomeric site position heterogeneity in macronucular DNA of Paramecium primaurelia. Nucl. Acids Res. 15, 1717-1728.

Baroudy, B. M., Venkatesan, S. and Moss, B. (1982). In-completely base-paired flip-flop terminal loops link the two DNA strands of the Vaccinia virus genome into one uninterrupted polynucleotide chain. Cell 28, 315-324.

Bates, G. P., MacDonald, M. E., Baxendale, S., Sedlacek, Z., Romano, D., Whaley, W. L., Allitto, B. A., Poustka, A., Gusella, J. F. and Lehrach, H. (1989). A 4p telomere clone likely to contain the mutant form of the Huntington's disease locus. (submitted).

Bellis, M., Pagès, M. and Roizès, G. (1987). A simple and rapid method for preparing yeast chromosomes for Pulsed Field Gel Electrophoresis. Nucl. Acids Res. 15, 6749.

Berman, J., Tachibana, C. Y. and Tye, B-K. (1986). Identification of a telomere-binding activity from yeast. Proc. Natl. Acad. Sci. USA 83, 3713-3717.

Bernards, A., Michels, P. A. M., Lincke, C. R. and Borst, P. (1983). Growth of chromosome ends in multiplying trypanosomes. Nature 303, 592-597.

Biessmann, H. and Mason, J. M. (1988). Progressive loss of DNA sequences from terminal chromosome deficiences in Drosophila melanogaster. EMBO J. 7, 1081-1086.

Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. Nature 321, 209-213.

Bird, A. P. (1987). CpG islands as gene markers in the vertebrate nucleus. Trends Genet. 3, 342-347.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7, 1513-1523.

Blackburn, E. H. and Challoner, P. B. (1984). Identification of a telomeric DNA sequence in Trypanosoma brucei. Cell 36, 447-457.

Blackburn, E. H. and Chiou, S. S. (1981). Non-nucleosomal packaging of a tandemly repeated DNA sequence at termini of extrachromosomal DNA coding for rDNA in Tetrahymena. Proc. Natl. Acad. Sci. USA 78, 2263-2267.

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

Blackburn, E. H. and Szostak, J. W. (1984). The molecular structure of centomeres and telomeres. Ann. Rev. Biochem. 53, 163-194.

Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. and Falkow, S. (1977). Construction and characterisation of new cloning vehicles II. A multipurpose cloning system. Gene 2, 95-113.

Brewer, B. J. and Fangman, W. L. (1987). The localization of replication origins on ARS plasmids in S. cerevisiae. Cell 51, 463-471.

Brown, W. R. A. (1988). A physical map of the human pseudoautosomal region. EMBO J. 7, 2377-2385.

Brown, W. R. A. (1989). Molecular cloning of human telomeres in yeast. Nature 338, 774-776.

Buchman, A. R., Kimmerly, W. J., Rine, J. and Kornberg, R. D. (1988a). Two DNA-binding factors recognize specific sequences at silencers, upstream activating sequences, autonomously replicating sequences and telomeres in Saccharomyces cerevisiae. Mol. Cell Biol. 8, 210-225.

Buchman, A. R., Lue, N. F. and Kornberg, R. D. (1988b). Connections between transcriptional activators, silencers, and telomeres as revealed by functional analysis of yeast DNA-binding protein. Mol. Cell Biol. 8, 5086-5099.

Budarf, M. and Blackburn, E. H. (1986). Chromatin structure of the telomeric region and 3'-nontranscribed spacer of the Tetrahymena ribosomal RNA genes. J. Biol. Chem. 261, 363-369.

Budarf, M. and Blackburn, E. H. (1987). S1 nuclease sensitivity of a double-stranded telomeric DNA sequence. Nucl. Acids Res. 15, 6273-6292.

Bullock, W. O., Fernandez, J. M. and Short, J. M. (1987). XL1-Blue - a high-efficiency plasmid transforming recA Escherichia coli strain with beta-galactosidase selection. Biotechniques 5, 376-379.

Burgers, P. M. J. and Percival, K. J. (1987). Transformation of yeast spheroplasts without cell fusion. Anal. Biochem. 163, 391-397.

Burke, D. T., Carle, G. F. and Olson, M. V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artifical chromosome vectors. Science 236, 806-812.

Burmeister, M. and Lehrach, H. (1989). Isolation of large DNA fragments from agarose gels using agarase. Trends Genet. 5, 41.

Carson, M. J. and Hartwell, L. (1985). CDC17 : an essential gene that prevents telomere elongation on yeast. Cell 42, 249-257.

Chan, C. S. M. and Tye, B-K. (1983a). A family of Saccharomyces cerevisiae repetitive autonomously replication sequences have very similar genomic environments. J. Mol. Biol. 168, 505-523.

Chan, C. S. M. and Tye, B-K. (1983b). Organization of DNA sequences and replication origins at yeast telomeres. Cell 33, 563-573.

Chandley, A. C. (1986). A model for effective pairing and recombination at meiosis based on early replicating sites (R-bands) along chromosomes. Human Genet. 72, 50-57.

Chandley, A. C., Goetz, P., Hargreave, T. B., Joseph, A. M. and Speed, R. M. (1984). On the nature and extent of XY pairing at meiotic prophase in man. Cytogenet. Cell Genet. 38, 241-247.

Cheng, J. F., Smith, C. L. and Cantor, C. R. (1989). Isolation and characterization of a human telomere. Nucl. Acids Res. 17, 6109-6127.

Cheung, M. K., Drivas, D. T., Littau, V. C. and Johnson, E. M. (1981). Protein tightly bound near the termini of the Physarum extrachromosomal rDNA palindrome. J. Mol. Biol. 91, 309-314.

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

Claflin, L. and Williams, K. (1978). Mouse myeloma-spleen cell hybrids : enhanced hybridization frequencies and rapid screening procedures. Current Topics in Microbiol. & Immunol. 81, 107-109.

Cooke, H. and Cross, S. (1988). PYAC-4 Neo a yeast artifical chromosome vector which codes for G418 resistance in mammalian cells. Nucl. Acids Res. 16, 11817.

Cooke, H. J. (1987). Telomeres of the human X and Y chromosomes. Development 101(Suppl.), 101-106.

Cooke, H. J., Brown, W. R. A. and Rappold, G. A. (1985). Hypervariable telomeric sequences from the human sex chromosomes are pseudoautosomal. Nature 317, 687-692.

Cooke, H. J., Fantes, J. and Green, D. (1983). Structure and evolution of human Y chromosome DNA. Differentiation 23(Suppl.) S48-S55.

Cooke, H. J. and Smith, B. A. (1986). Variability at the telomeres of the human X/Y pseudoautosomal region. Cold Spring Harbor Symp. Quant. Biol. 51, 213-219.

Cross, S. H., Allshire, R. C., McKay, S. J., McGill, N. I. and Cooke, H. J. (1989). Cloning of human telomeres by complementation in yeast. Nature 338, 771-774.

Darling, S. M., Banting, G. S., Pym, B., Wolfe, J. and Goodfellow, P. N. (1986). Cloning an expressed gene shared by the human sex chromosomes. Proc. Natl. Acad. Sci. USA 83, 135-139.

Dawson, D. and Herrick, G. (1984). Telomeric properties of C4A4-homologous sequences in micronuclear DNA of Oxytricha fallax. Cell 36, 171-177.

Dawson, D. S., Murray, A. W. and Szostak, J. W. (1986). An alternative pathway for meiotic chromosome segregation in yeast. Science 234, 713-717.

Dunn, B., Szauter, P., Pardue, M. L. and Szostak, J. W. (1984). Transfer of yeast telomeres to linear plasmids by recombination. Cell 39, 191-201.

Edwards, C. A. and Firtel, R. A. (1984). Site-specific phasing in the chromatin of the rDNA in Dictyostelium discoideum. J. Mol. Biol. 180, 73-90.

Ellis, N. A., Goodfellow, P. J., Pym, B., Smith, M., Palmer, M., Frischauf, A. M. and Goodfellow, P. N. (1989). The pseudoautosomal boundary in man is defined by an Alu repeat sequence inserted on the Y chromosome. Nature 337, 81-84.

Emery, H. S. and Weiner, A. M. (1981). An irregular satellite sequence is found at the termini of the linear extrachromosomal rDNA in Dictyostelium discoideum. Cell 26, 411-419.

Feinberg, A. P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132, 6-13.

Feinberg, A. P. and Vogelstein, B. (1984). Addenum: a technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137, 266-267.

Flickinger, R. A. (1986). Localization of AT-rich sequences at the nuclear matrix. Cell Biol. Int. Rep. 10, 415-420.

Forney, J., Henderson, E. R. and Blackburn, E. H. (1987). Identification of the telomeric sequence of the acellular slime molds Didymium iridis and Physarum polycephalum. Nucl. Acids Res. 15, 9143-9152.

Gilliam, T. C., Tanzi, R. E., Haines, J. L., Bonner, T. I., Faryniaz, A. G., Hobbs, W. J., MacDonald, M. E., Cheng, S. V., Folstein, S. E., Conneally, P. M., Wexler, N. S. and Gusella, J. F. (1987). Localization of the Huntington's disease gene to a small segment of Chromosome 4 flanked by D4S10 and the telomere. Cell 50, 565-571.

Gottschling, D. E. and Cech, T. R. (1984). Chromatin structure of the molecular ends of Oxytricha macronuclear DNA - phased nucleosomes and a telomeric complex. Cell 38, 501-510.

Gottschling, D. E. and Zakian, V. A. (1986). Telomere proteins : specifice recognition and protection of the natural termini of Oxytricha macronuclear DNA. Cell 47, 195-205.

Gray, H. B., Ostrander, D. A., Hodnett, J. L., Legerski, R. J. and Robberson, D. L. (1975). Extracellular nucleases of Pseudomonas Bal-31. I. Characterisation of single-strand specific deoxyriboendonuclease and double-strand deoxyriboexonuclease activities. Nucl. Acids Res. 2, 1459-1492.

Greider, C. W. and Blackburn, E. H. (1985). Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell 43, 405-413.

Greider, C. W. and Blackburn, E. H. (1987). The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. Cell 51, 887-898.

Greider, C. W. and Blackburn, E. H. (1989). A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. Nature 337, 331-337.

Grunstein, M. and Hogness, D. S. (1975). Colony hybridization : a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA 72, 3961-3965.

Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchil, A. Y., Young, A. B., Shoulson, I., Bonilla, E. and Martin, J. B. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. Nature 306, 234-238.

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. J. Mol. Biol. 166, 557-580.

Henderson, E., Hardin, C. C., Walk, S. K., Tinoco J/r., I. and Blackburn, E. H. (1987). Telomeric DNA oligonucleotides form novel intermolecular structures containing guanine.guanine base pairs. Cell 51, 899-908.

Henderson, E., Larson, D., Melton, W., Shampay, J., Spangler, E., Greider, C., Ryan, T. and Blackburn, E. H. (1988). Structure, synthesis, and regulation of telomeres. Cancer Cells 6, 453-461.

Henderson, E. R. and Blackburn, E. H. (1989). An overhanging 3' terminus is a conserved feature of telomeres. Mol. Cell Biol. 9, 345-348.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28, 351-359.

Hofmann, J. F. X., Laroche, T., Brand, A. H. and Gasser, S. M. (1989). RAP-1 factor is necessary for DNA loop formation in vitro at the silent mating type locus HML. Cell 57, 725-735.

Horowitz, H., Thorburn, P. and Harber, J. E. (1984). Rearrangements of highly polymorphic regions near telomeres of Saccharomyces cerevisiae. Mol Cell Biol 4, 2509-2517.

IBI Catalogue (86/87). Life on the Edge (Vol II).

Jäger, D. and Philippsen, P. (1989). Stabilization of dicentric chromosomes in Saccharomyces cerevisiae by telomere addition to broken ends or by centromere deletion. EMBO J. 8, 247-254.

Jeffreys, A. J., Wilson, V. and Thein, S. L. (1985). Hypervariable 'minisatellite' regions in human DNA. Nature 314, 67-73.

Johnson, E. M. (1980). A family of inverted repeat sequences and specific single-strand gaps at the termini of the Physarum rDNA palindrome. Cell 22, 875-886.

Klobutcher, L. A., Swanton, M. T., Donini, P. and Prescott, D. M. (1981). All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. Proc. Natl. Acad. Sci. USA 78, 3015-3019.

Köhler, G., Howe, S. C. and Milstein, C. (1976). Fusion between immunoglobulin-secreting and non-secreting myeloma cell lines. Eur. J. Immunol. 6, 292-296.

Koller, P. C. and Darlington, C. D. (1934). The genetical and mechanical properties of the sex chromosomes 1. Rattus norvegicus. Journal Genetics 29, 159-173.

Krysan, P. J., Haase, S. B. and Calos, M. P. (1989). Isolation of human sequences that replicate autonomously in human cells. Mol. Cell Biol. 9, 1026-1033.

Larson, D. D., Spangler, E. A. and Blackburn, E. H. (1987). Dynamics of telomere length variation in Tetrahymena thermophila. Cell 50, 477-483.

Lathe, R., Kieny, M. P., Skory, S. and Lecocq, J. P. (1984). LABORATORY METHODS Linker tailing: unphosphorylated linker oligonucleotides for joining DNA termini. DNA 3, 173-182.

Lejenue, J., Dutrillaux, B., Rethoré, M. O. and Prieur, M. (1973). Comparison de la structure fine des chromatides d'Homo sapiens et de Pan troglodytes. Chromosoma 43, 423-444.

Levis, R. W. (1989). Viable deletions of a telomere from a Drosophila chromosome. Cell 58, 791-801.

Lundblad, V. and Szostak, J. W. (1989). A mutant with a defect in telomere elongation leads to senesence in yeast. Cell 57, 633-643.

Lustig, A. J. and Petes, T. D. (1986). Identification of yeast mutants with altered telomere structure. Proc. Natl. Acad. Sci. USA 83, 1398-1402.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982). Molecular Cloning : A Laboratory Manual. Cold Spring Harbor, N.Y.

Mathog, D., Hochstrasser, M., Gruenbaum, Y., Saumweber, H. and Sedat, J. (1984). Characteristic folding pattern of polytene chromosomes in Drosophila salivary gland nuclei. Nature 308, 414-421.

Matsumato, T., Fukui, K., Niwa, O., Sugawara, N., Szostak, J. W. and Yanagida, M. (1987). Identification of healed terminal DNA fragments in linear minichromosomes of Schizosaccharomyces pombe. Mol. Cell Biol. 7, 4424-4430.

McCarroll, R. M. and Fangman, W. L. (1988). Time of replication of yeast centromeres and telomeres. Cell 54, 505-513.

McClintock, B. (1938). The production of homozygous deficient tissues with mutant characteristics by means of the abberrant mitotic behaviour of ring-shaped chromosomes. Genetics 23, 315-376.

McClintock, B. (1941). The stability of broken ends of chromosomes in Zea Mays. Genetics 26, 234-282.

McClintock, B. (1942). The fusion of broken ends of chromosomes following nuclear fusion. Proc. Natl. Acad. Sci. USA 28, 458-463.

Merchlinsky, M. and Moss, B. (1986). Resolution of linear minichromosomes with hairpin ends from circular plasmids containing Vaccinia virus concatemer junctions. Cell 45, 879-884.

Messing, J. (1983). New M13 vectors for cloning. Methods Enzymol. 101, 20-78.

Mirkovitch, J., Mirault, M. E. and Laemmli, U. K. (1984). Organization of the higher-order chromatin loop : specific DNA attachment sites on the nuclear scaffold. Cell 39, 223-232.

Moens, P. B. , Ed. (1987). Meiosis. Academic Press.

Morin, G. B. (1989). The human telomere transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. (submitted).

Mortimer, R. K. and Schild, D. (1980). Genetic map of Saccharomyces cerevisiae. Microbiol. Reviews 44, 519-571.

Moses, M. J. (1977). Synaptonemal complex karyotyping in spermatocytes of the chinese hamster (Cricetulus griseus). Chromosoma 60, 99-125.

Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratliff, R. L. and Wu, J. R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. USA 85, 6622-6626.

Müller-Hill, B., Crapo, L. and Gilbert, W. (1968). Mutants that make more lac repressor. Proc. Natl. Acad. Sci. USA 59, 1259-1264.

Müller, U., Fontaine, D., Adinolfi, M., Fraccaro, M. and Lalande, M. (1989). Analysis of complex Y chromosome abberrations using a single DNA probe (Y-367). Cytogenet. Cell Genet. 50, 161-164.

Murray, A. W., Claus, T. E. and Szostak, J. W. (1988). Characterisation of two telomeric DNA processing reactions in yeast. Mol. Cell Biol. 8, 4642-4650.

Murray, A. W. and Szostak, J. W. (1983). Construction of artificial chromosomes in yeast. Nature 305, 189-193.

Murray, A. W. and Szostak, J. W. (1985). Chromosome segregation in mitosis and meiosis. Ann. Rev. Cell Biol. 1, 289-315.

Murray, A. W. and Szostak, J. W. (1986). Construction and behaviour of circularly permuted and telocentric chromosomes in Saccharomyces cerevisiae. Mol. Cell Biol. 6, 3166-3172.

Nelson, M. and McClelland, M. (1989). Effect of site-specific methylation on DNA modification methyltransferases and restriction endonucleases. Nucl. Acids Res. 17(Suppl.), r389-r415.

Newlon, C. S. and Burke, W. (1980). Replication of small chromosomal DNAs in yeast. ICN-UCLA Symp. Mol. Cell Biol. 19, 399-409.

Norrander, J., Kempe, T. and Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene 26, 101-106.

Oka, Y. and Thomas, Jr., C. A. (1987). The cohering telomeres of Oxytricha. Nucl. Acids Res. 15, 8877-8898.

Petit, C., Levilliers, J. and Weissenbach, J. (1988). Physical mapping of the human pseudoautosomal region; comparison with genetic linkage map. EMBO J. 7, 2369-2376.

Pluta, A. F., Dani, G. M., Spear, B. B. and Zakian, V. A. (1984). Elaboration of telomeres in yeast : recognition and modification of termini from Oxytricha macronuclear DNA. Proc. Natl. Acad. Sci. USA 81, 1475-1479.

Pluta, A. F., Kaine, B. P. and Spear, B. B. (1982). The terminal organisation of macronuclear DNA in Oxytricha fallax. Nucl. Acids Res. 10, 8145-8154.

Pluta, A. F. and Zakian, V. A. (1989). Recombination occurs during telomere formation in yeast. Nature 337, 429-433.

Pologe, L. G. and Ravetch, J. V. (1988). Large deletions result from breakage and healing of P. falciparum chromosomes. Cell #55#, 869-874.

Ponzi, M., Pace, T., Dore, E. and Frontali, C. (1985). Identification of a telomeric DNA sequence in Plasmodium berghei. EMBO J. 4, 2991-2995.

Porteous, D. J. (1987). Chromosome mediated gene transfer : a functional assay for complex loci and an aid to human gene mapping. Trends Genet. 3, 177-182.

Price, C. M. and Cech, T. R. (1987). Telomeric DNA-protein interactions of Oxytricha macronuclear DNA. Genes & Development 1, 783-793.

Price, C. M. and Cech, T. R. (1989). Properties of the telomere DNA-binding protein from Oxytricha nova. Biochemistry 28, 769-774.

Raghuraman, M. K., Dunn, C. J. and Cech, T. R. (1989). Oxytricha telomeric nucleoprotein complexes reconstituted with synthetic DNA. Nucl. Acids Res. 17, 4235-4253.

Rappold, G. A. and Lehrach, H. (1988). A long-range restriction map of the pseudoautosomal region by partial digest PFGE analysis from the telomere. Nucl. Acids Res. 16, 5361-5377.

Richards, E. J. and Ausubel, F. M. (1988). Isolation of a higher eucaryotic telomere from Arabidopsis thaliana. Cell 53, 127-136.

Robertson, A. E., Wolffe, A. P., Houser, L. J. and Olins, D. E. (1989). The 5S RNA gene minichromosome of Euplotes. Nucl. Acids Res. 17, 4699-4712.

Rogers, J. H. (1985). The origin and evolution of retroposons. Int. Rev. Cytol. 93, 188-279.

Roth, M. and Prescott, D. M. (1985). DNA intermediates and telomere addition during genome reorganization in Euplotes crassus. Cell 41, 411-417.

Rouyer, F., Simmler, M-C., Johnsson, C., Vergnaud, G., Cooke, H. J. and Weissenbach, J. (1986a). A gradient of sex linkage in the pseudoautosomal region of the human sex chromosomes. Nature 319, 291-295.

Rouyer, F., Simmler, M-C., Vergnaud, G., Johnsson, C., Levilliers, J., Petit, C. and Weissenbach, J. (1986b). The pseudoautosomal region of the human sex chromosomes. Cold Spring Harbor Symp. Quant. Biol. 51, 221-228.

Rubin, G. M. (1978). Isolation of a telomeric DNA sequence from Drosophila melanogaster. Cold Spring Harbor Symp. Quant. Biol. 42, 1041-1046.

Ruddle, F. H. and Creagan, R. P. (1975). Parasexual approaches to the genetics of man. Ann Rev. Genetics 9, 407-487.

Runge, K. W. and Zakian, V. A. (1989). Introduction of extra telomeric DNA sequences into Saccharomyces cerevisiae result in telomere elongation. Mol. Cell Biol. 9, 1488-1497.

Sanford, J. P., Chapman, V. M. and Rossant, J. (1985). DNA methylation in extra-embryonic lineages of mammals. Trends Genet. 1, 89-93.

Sanger, F. and Coulson, A. R. (1978). The use of thin acrylamide gels for DNA sequencing. FEBS Letters 87, 107-110.

Sanger, F., Micklens, S. and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Seawright, A., Fletcher, J., Fantes, J. A., Morrison, H., Porteous, D. J., Li, S. S. L., Hastie, N. D. and van Heyningen, V. (1988). Analysis of WAGR deletions and related translocations with gene-specific DNA probes using FACS-selected cell hybrids. Somatic Cell & Mol. Genet. 14, 21-30.

Sen, D. and Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature 334, 364-366.

Shampay, J. and Blackburn, E. H. (1988). Generation of telomere-length heterogeneity in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA 85, 534-538.

Shampay, J., Szostak, J. W. and Blackburn, E. H. (1984). DNA sequences of telomeres maintained in yeast. Nature 310, 154-157.

Shippen-Lentz, D. and Blackburn, E. H. (1989). Telomere terminal transferase activity from Euplotes crassus adds large numbers of TTTTGGGG repeats onto telomeric primers. Mol. Cell Biol. 9, 2761-2764.

Shoeman, R. L., Wadle, S., Scherbarth, A. and Traub, P. (1988). The binding in vitro of the intermediate filament protein vimentin to synthetic oligonucleotides containing telomere sequences. J. Biol. Chem. 263, 18744-18749.

Shore, D. and Nasmyth, K. (1987). Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. Cell 51, 721-732.

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

Spangler, E. A., Ryan, T. and Blackburn, E. H. (1988). Developmentally regulated telomere addition in Tetrahymena thermophila. Nucl. Acids Res. 16, 5569-5585.

Stillman, B. (1983). The replication of Adenovirus DNA with purified protein. Cell 35, 7-9.

Strauss, F. and Varshavsky, A. (1984). A protein binds to a satellite DNA repeat at three specific sites that would be brought into proximity by DNA folding in the nucleosome. Cell 37, 889-901.

Sugawara, N. and Szostak, J. W. (1986). Telomeres of Schizosaccharomyces pombe. Yeast 2, S373-S373.

Szostak, J. W. and Blackburn, E. H. (1982). Cloning yeast telomeres on linear plasmid vectors. Cell 29, 245-255.

Traverse, K. L. and Pardue, M. L. (1988). A spontaneously opened ring chromosome of Drosphila melanogaster has acquired He-T DNA sequences at both new telomeres. Proc. Natl. Acad. Sci. USA 85, 8116-8120.

Van der Ploeg, L. H. T., Lin, A. Y. C. and Borst, P. (1984). Structure of the growing telomeres of trypanosomes. Cell 36, 459-468.

van Heyningen, V., Barron, L., Brock, D. J. H., Crichton, D. and Lawrie, S. (1982). Monoclonal antibodies to human α-foetoprotein. Analysis of the behaviour of three different antibodies. J. Immunol. Methods 50, 123-131.

Vernick, K. D. and McCutchan, T. F. (1988). Sequence and structure of a Plasmodium falciparum telomere. Mol. & Biochem. Parasitol. 28, 85-94.

Vieira, J. and Messing, J. (1982). The pUC plasmids an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universial primers. Gene 19, 259-269.

Walmsley, R. W., Chan, C. S. M. and Tye, B-K. (1984). Unusual DNA sequences associated with the ends of yeast chromosomes. Nature 310, 157-160.

Walmsley, R. W. and Petes, T. D. (1985). Genetic control of chromosome length in yeast. Proc. Natl. Acad. Sci. USA 82, 506-510.

Wang, H. S., Greenberg, C. R., Kalonsek, D., Gusella, J. F., Horsman, D. and Hayden, M. R. (1985). Sub-regional assignment of the linked marker (D4S10) for Huntington disease by *in situ* hybridization. Cytogenet. Cell Genet. **40**, 772.

Watson, J. D. (1972). Origin of concatameric T4 DNA. Nature New Biol. **239**, 197-201.

Weiss, B., Jacquemin-Sablon, A., Live, T. R., Fareed, G. C. and Richardson, C. C. (1968). Enzymatic breakage and joining of deoxyribonucleic acid. J. Biol. Chem. **243**, 4543-4555.

White, M. J. D. (1978). Chain processes in chromosomal speciation. Syst. Zool. **27**, 285-298.

Yanisch-Peron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains : nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**, 103-119.

Young, B. S., Pession, A., Traverse, K. L., French, C. and Pardue, M. L. (1983). Telomeric regions in *Drosophila* share complex DNA sequences with pericentric heterochromatin. Cell **34**, 85-94.

Yunis, J. J. and Prakash, O. (1982). The origin of man : a chromosomal pictoral legacy. Science **215**, 1525-1529.

Zahler, A. M. and Prescott, D. M. (1988). Telomere terminal transferase activity in the hypotrichous ciliate *Oxytricha nova* and a model for replication of the ends of linear DNA molecules. Nucl. Acids Res. **16**, 6953-6972.

Zakian, V. A. and Blanton, H. M. (1988). Distribution of telomere-associated sequences on natural chromosomes in *Saccharomyces cerevisiae*. Mol. Cell Biol. **8**, 2257-2260.

APPENDIX

PUBLISHED PAPERS

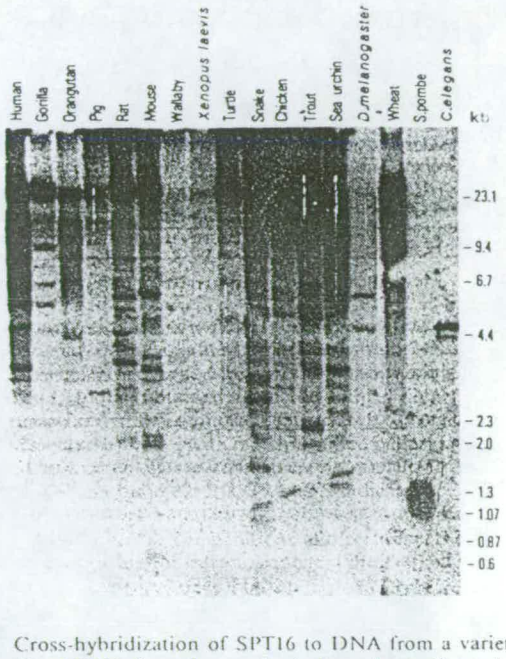


Fig. 1 Cross-hybridization of SPT16 to DNA from a variety of species. Total DNA from the species indicated was digested with an excess of the restriction endonuclease *Eco*RI (3 units per μg, 37 °C, overnight). The amount of DNA loaded on to a 0.8% agarose gel was ~10 μg, except for the *D. melanogaster* and *S. pombe* samples, of which 1 μg and 300 ng were loaded respectively. After electrophoresis the gel was blotted to Hybond and hybridized to ^{32}P-labelled *Bam*HI/*Hind*III telomere fragment purified from pucSPT16. Hybridization was overnight at 68 °C in 7% SDS, 0.5 M Na_2HPO_4, *p*H 7.2. Filters were washed 3 times at 68 °C in 0.1 × SSC, 0.1% sodium pyrophosphosphate and 0.1% SDS. Size markers (kb) are indicated.

Telomeric repeat from *T. thermophila* cross hybridizes with human telomeres

Robin C. Allshire*, John R. Gosden*, Sally H. Cross*, Gwen Cranston*, Derek Rout*, Neal Sugawara†, Jack W. Szostak†, Peter A. Fantes‡ & Nicholas D. Hastie*

* MRC Clinical and Population Cytogenetics Unit, Western General Hospital, Edinburgh EH4 2XU, UK
† Massachusetts General Hospital, Boston, Massachusetts 02114, USA
‡ Department of Zoology, University of Edinburgh, Edinburgh EH9 3JT, UK

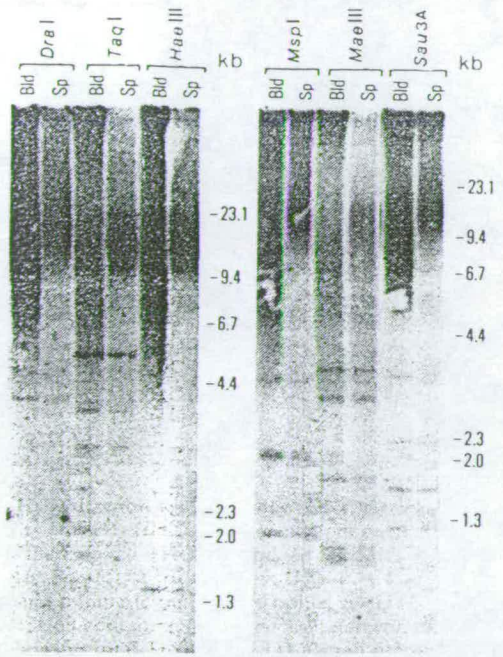
The ends (telomeres) of eukaryotic chromosomes must have special features to ensure their stability and complete replication. Studies in yeast[1–3], protozoa[4–6], slime moulds[7,8] and flagellates[8,10] show that telomeres are tandem repeats of simple sequences that have a G-rich and a C-rich strand. Mammalian telomeres have yet to be isolated and characterized, although a DNA fragment within 20 kilobases of the telomeres of the short arms of the human sex chromosomes has been isolated[11]. Recently we showed that a chromosome from the fission yeast *Schizosaccharomyces pombe* could, in some cases, replicate as an autonomous mini-chromosome in mouse cells[12]. By extrapolation from other systems[1,3,13,14], we reasoned that mouse telomeres could be added to the *S. pombe* chromosome ends in the mouse cells. On setting out to test this hypothesis we found to our surprise that the telomeric probe used (containing both the *S. pombe* and *Tetrahymena thermophila* repeats) hybridized to a series of discrete fragments in normal mouse DNA and DNA from a wide range of eukaryotes. We show here that the sequences hybridizing to this probe are located at the telomeres of most, if not all, human chromosomes and are similar to the *Tetrahymena* telomeric-repeat component of the probe.

The telomeric tandem repeat sequence from *S. pombe* has recently been cloned[3]. Owing to the cloning strategy, the resultant plasmid pSPT16 contained ~100 base pairs (bp) of the *T. thermophila* telomeric repeat, $(C_4A_2)_N$ and ~300bp of the *S. pombe* telomeric repeat, $(C_{1-6}G_{0-1}T_{0-1}GTA_{1-2})_N$.

Figure 1 shows the hybridization of this probe to a variety of eukaryotic DNAs. In a number of cases there is a variable smear in addition to a series of discrete bands. This is a provocative finding because individual telomeres in several organisms show considerable size heterogeneity[1–11]. The total signal seen in these organisms is variable and may be proportional to the number of linkage groups, although the signal in some species (for example, *Xenopus laevis* and the wallaby) is relatively weak.

A strong signal has also been seen in other organisms (tobacco, petunia and crab, data not shown), but we have not detected any hybridization of the probe to DNA extracted from *S. cerevisiae* or *E. coli* (data not shown). These hybridizations were all washed at high stringency (0.1 × SSC; 68 °C). There must therefore be a considerable degree of similarity between the probe and the sequences detected.

Next we tested whether these cross-hybridizing sequences could be located at, or near, human telomeres. Firstly, we looked for differences in the pattern of hybridization observed between sperm and somatic DNA from the same individual. Cooke and Smith[15] showed that telomeres of the human sex chromosomes were several kilobase pairs (kb) longer in sperm DNA relative to somatic DNA. Figure 2 shows the digestion pattern seen when the SPT16 probe was hybridized to sperm and somatic



ig. 2 Comparison of germline and somatic DNA samples from single individual. Blood (Bld) and sperm (Sp) DNAs (~10 μg) ere digested with the restriction enzymes indicated and analysed y blot-hybridization (conditions and probe as in Fig. 1). Size markers (kb) are indicated.

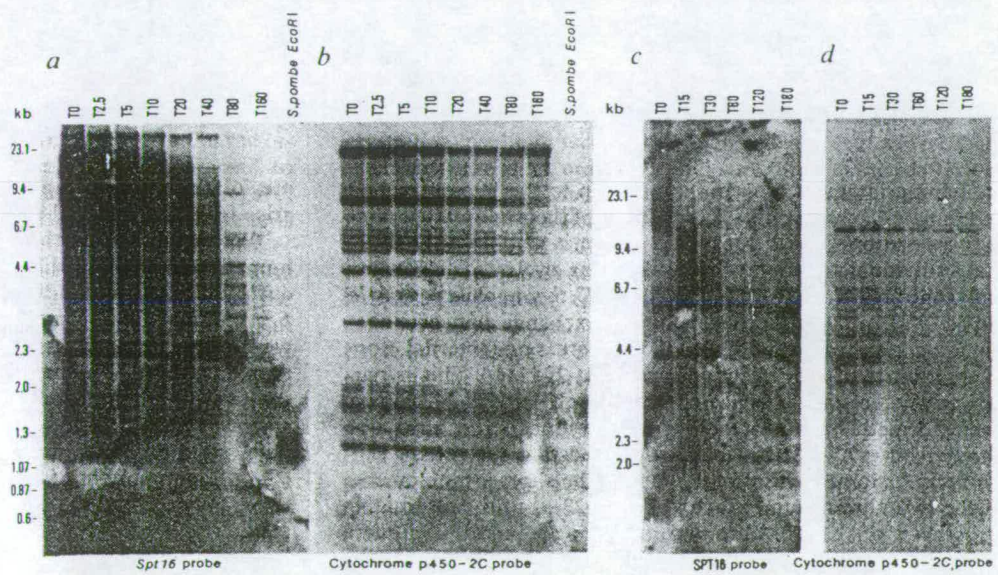
IA from the same individual using a variety of restriction ymes. The smear in all cases is on average of higher molecular ght and more homogeneous in sperm DNA than in somatic IA, indicating that the probe could be recognizing human meres. A very similar pattern is seen with a range of restric- enzymes. This indicates that the region detected by SPT16 evoid of these restriction enzyme sites and probably consists simple repetitive sequence. This region ranges from 5–20 kb lood and 10–20 kb in sperm.

he definitive test of telomere location is sensitivity to diges- by the exonuclease *Bal*31, because the telomeres of lower aryotes[1,3,5,6,9,10] and the human sex chromosomes[15] are all ssible to digestion by this enzyme. Figure 3*a* and *c* shows ridization of the SPT16 probe to human placenta or sperm A that has been digested for increasing periods of time with *Bal*31 and then with the restriction endonucleases MspI and TaqI respectively. It is noticeable that the diffuse smear of hybridization ranging in size from 5 to 20 kb in placenta DNA (Fig. 3*a*) has completely disappeared after 40–80 min incubation. By contrast, the size of most of the discrete bands remains undiminished for up to 160 min of *Bal*31 treatment. Panel 3*b* shows rehybridization of the same filter to a probe for the cytochrome P450-*C2C* gene family which resides at chromosome 10q23–24 (ref. 16), a long distance from the telomere. There is no decrease in size of the P450-*C2C*-hybridizing bands during the time course.

Thus the heterogeneous smear that is sensitive to *Bal*31 digestion must represent sequences close to the telomere. Using the telomeric probe for the human sex chromosomes[11] we have estimated that the rate of removal of nucleotides by *Bal*31 in Fig. 3*a* and *b* was ~3.8 kb per hour (data not shown). As the entire smear has almost gone within 40 min, the sequences detected must be within 2.5 kb of the end of the chromosomes in placenta. In Fig. 3*c* and *d* it can be seen that the large 10–20 kb smear of hybridization disappears within 60 min of *Bal*31 digestion, but that the cytochrome P450-*C2C* signal remains intact throughout. Clearly the more homogeneous smear in sperm DNA is also located close to the telomere. In Fig. 3 the discrete fragments hybridizing to the telomere probe appear to be no more sensitive to *Bal*31 digestion than the cytochrome P450-*C2C* genes. These discrete fragments are likely to reside at internal chromosomal sites. Whether such internal sites are at a telomeric location or not remains to be seen (see below). The prominence of these discrete bands in Fig. 3*a* compared with Fig. 2 is a result of the different hybridization conditions used (see figure legends).

If the SPT16 probe is hybridizing to telomeres, this should also be revealed by *in situ* hybridization to metaphase chromosome spreads. Figure 4 shows a representation of the grain distributions scored after hybridization of the telomere probe *in situ* to metaphase spreads of three different individuals (depicted by different symbols). Fifty-four per cent of the grains were located in the telomeric bands, compared with 5–15% for various single-copy probes mapped in our laboratory[16–18]. There are two other potentially interesting features revealed by these data. Firstly, there appears to be variation among the three individuals in the range of chromosomes to which the probe hybridizes strongly. For example, in male 2, 14.3% (8% of total grains scored) of telomeric grains (54% of total grains scored) are at 13qter, whereas in male 1, only 4.8% of grains are found at this telomere, and in female 1 there are none. In contrast, 9.5% of

g. 3 Bal31 sensitivity of the ffuse smear homologous to PT16 in human DNA. Human ale placental DNA (*a* and *b*) or erm DNA (*c* and *d*) was digested ith *Bal*31 for the times (*T*) indi- ted. The samples were then gested with *Msp*I (*a* and *b*) or aq1 (*c* and *d*), separated by elec- phoresis in agarose, blotted to trocellulose filters and hybridized ith ^{32}P-labelled *Bam*HI *Hind*III agment from pucSPT16 (*a* and . Hybridization conditions for trocellulose were as previously scribed[12]. Filters were washed as scribed in Fig. 1. After toradiographic exposure, the ters were washed in boiling distil- d water and were then rehybrid- ed with ^{32}P-labelled probe for the 50-*C2C* gene family[16]: the sulting autoradiographs are own in *b* and *d*. Size markers (kb) are indicated.



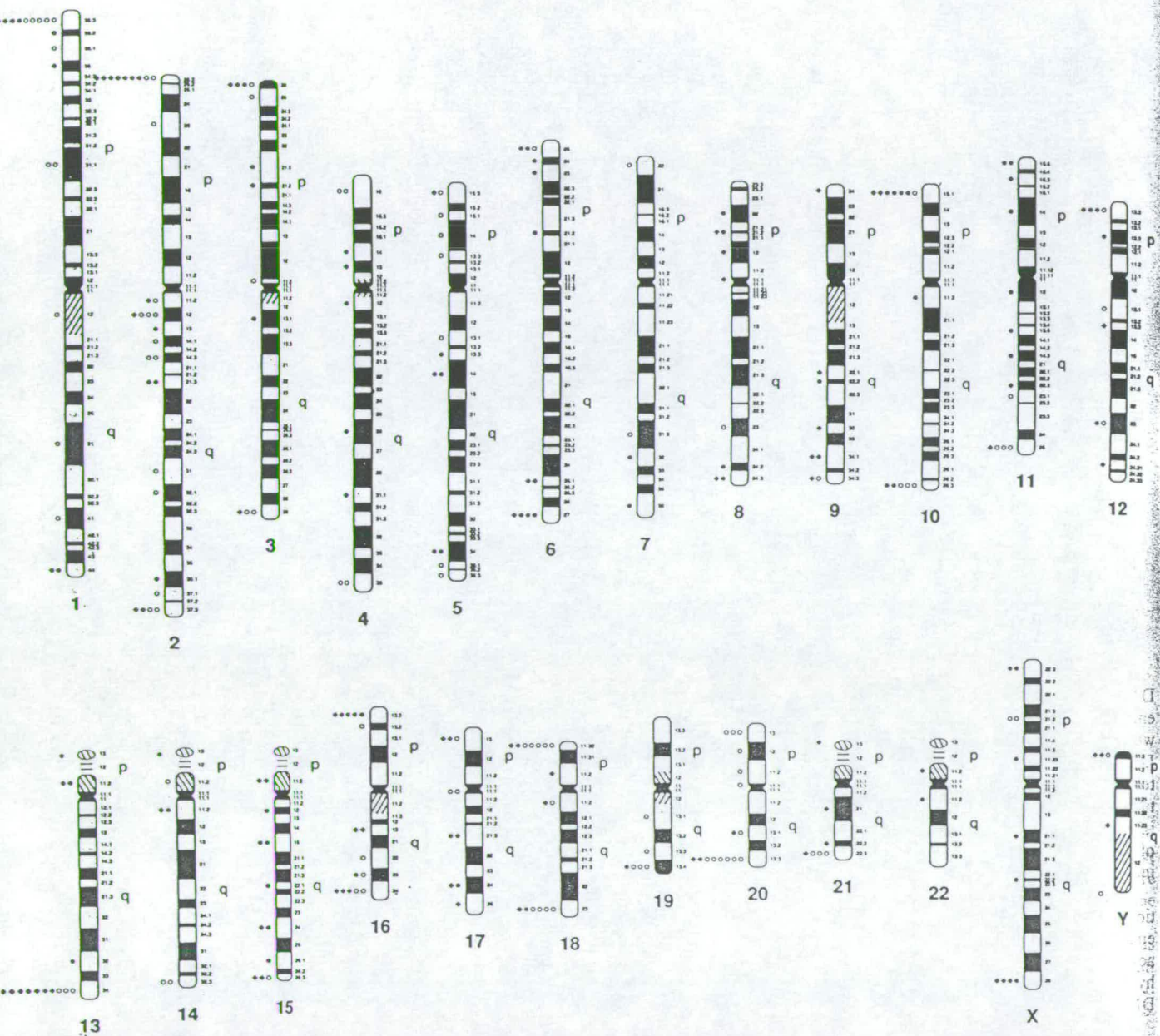


Fig. 4 *In situ* hybridization of SPT16 to metaphase spreads of human chromosomes. The positions of grains observed in 37 metaphase spreads prepared from two male individuals (male 1 and male 2 represented by ○ and ◆) and one female (female 1 represented by ●) are shown. A total of 274 grains were counted, 149 of which are located in the most telomeric bands. Metaphase spreads and ^{3}H-labelled SPT16 probe (specific activity $\sim 2\times10^8$ d.p.m. μg^{-1}) were prepared as described previously[12,17]. Hybridization and autoradiography were as described [12]. The final washes (4×5 min) were at 40 °C in 1×SSC.

the telomeric grains seen in male 1 map to the 20qter, male 2 and female 1 having only 2% and 2.7% of their telomeric grains at this position. If this is confirmed later it would imply either that individual telomeres are undergoing cycles of expansion and contraction (as demonstrated in *T. thermophila*[19] and *T. Brucei*[20,21]), or that there is telomere exchange between non-homologous chromosomes. Secondly, there is one internal, non-telomeric site of hybridization located at 2*q*11–14. This is particularly interesting because chromosome 2 is thought to have been formed by fusion of two ancestral acrocentric ape chromosomes[22–24]. It has been suggested that this was a centromeric fusion resulting in the loss of the short arm telomeres of the two chromosomes involved (ref. 22), but the alternative hypothesis of the two chromosomes being joined by telomere fusion[23,24], is supported by hybridization of the telomere probe at bands 2*q*11–2*q*14. It is also of interest that 2*q*11 and 2*q*13 are folate-sensitive fragile sites[25]. This fragility could possibly arise from the residual telomeric sequences.

We conclude that the SPT16 probe recognizes telomeres of humans and, in all likelihood, of other higher eukaryotes. Under our hybridization conditions, the telomeres of *T. thermophila*, *S. cerevisiae* and *S. pombe* failed to cross-hybridize, so the signal we observe probably reflects a real sequence identity, rather than an overall G-richness of the probe (Fig. 5). Furthermore, the use of other probes containing only the *S. pombe* or *T. thermophila* telomeric sequences has established that the signal detected in human and mouse DNA is due to the *T. thermophila* $(C_4A_2)_N$ component of the SPT16 probe (Fig. 5). The *S. pombe* telomeric repeat alone does not hybridize to total human, mouse or *S. cervisiae* DNA or to *T. thermophila* ribosomal DNA telomeres (Fig. 5*c*). Therefore human telomeres have a sequence similar to those of *T. thermophila*.

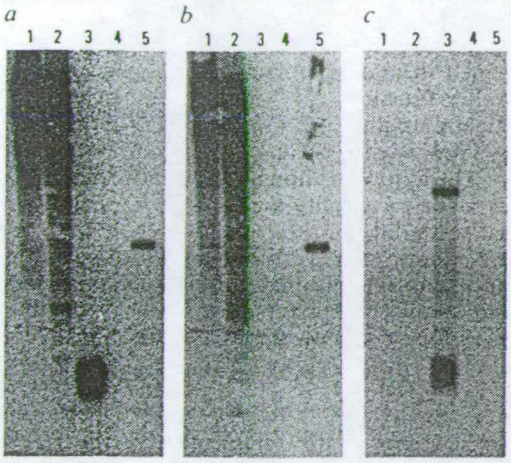


Fig. 5 Hybridization with *S. pombe* and *T. thermophila* telomeric probes. *a*, *Bam*HI/*Hind*III fragment from pucSPT16 containing both *T. thermophila* (100 bp) and *S. pombe* (300 bp) telomeric repeats; *b*, *Bam*HI/*Xho*I fragment from YAC4 containing the *T. thermophila* telomeric repeat (300 bp); *c*, *Sac*I/*Hind*III fragment from pNSU68 containing the *S. pombe* telomeric repeat (200 bp). Lanes showing total DNA from (1) human placenta, 10 μg; (2) mouse C_{127} cells, 10 μg; (3) *S. pombe* ED628, 300ng (4) *S. cerevisiae* DC5, 300ng digested with *Eco*RI and loaded on three identical 0.8% agarose gels. Lanes(5), a purified 3.5-kb telomeric *Bg*1II fragment (~80 pg) from *T. thermophila* rDNA. After electrophoresis, the gels were blotted to Hybond and hybridized overnight at 68 °C in 7% SDS, 0.5 M Na_2HPO_4, *p*H 7.2, to the ^{32}P-labelled purified telomeric probes. Filters in *a* and *b* were washed as in Fig. 1. Filter in *c* was washed in 1×SSC at 68 °C.

All known telomeres have one protruding strand which is G-rich. This must also be true of human telomeres, shown here to cross-hybridize to the $(C_4A_2 \cdot T_2G_4)_N$ *T. thermophila* telomeric repeat. Recently, it has been shown[26] that oligonucleotides of the *T. thermophila* telomere G-strand form novel intramolecular structures. Also, only oligonucleotides of the G-rich strand from the telomeres of five different organisms can act as elongation primers for the telomerase present in *T. thermophila* cell extracts[27]. Given that the *T. thermophila* telomeric repeat cross-hybridizes to human telomeres, we would also expect these other properties to be similar. As *T. thermophila* and humans are separated by a large evolutionary distance, it is possible that only a limited number of G-rich sequences can function as telomeres in eukaryotes.

Our results with this $(C_4A_2)_N$ probe point to a direct isolation of the telomeres of individual human chromosomes. Beyond defining their structural properties, cloned human telomeres could be useful in the construction of synthetic functional chromosomes, in the study of possible interactions between homologous and non-homologous chromosomes, to characterize chromosome rearrangements involving telomeres, and to define the genetic limits of each chromosome. In this last regard, it is of particular interest that the Huntington's chorea gene is thought to reside very close to the telomere of the short arm of chromosome 4 (ref. 28). No flanking markers on the distal side of this mutation have been identified. Thus the telomere of the short arm of chromosome 4 could allow a better definition of the locus and facilitate the isolation of the gene itself.

We thank Richard Meehan, Bob Hill, Adrian Bird, David Finnegan, Jane Kinnard, John Sulston, Anne Ramage and Steve Smith for various DNA samples; Howard Cooke and members of our laboratories for their invaluable advice and discussion, and especially Richard Buckland for pointing out the history of human chromosome 2; N. Davidson, S. Bruce and D. Stewart for photographic work; and Professor H. J. Evans for his encouragement and support. S.H.C. is supported by SERC studentship.

Received 27 January; accepted 7 March 1988.

1. Shampay, J., Szostak, J. W. & Blackburn, E. H. *Nature* **310**, 154–157
2. Walmsley, R. W., Chan, C. S. M., Tye, B. K. & Petes, T. D. *Nature* **310**, 157–160 (1984).
3. Sugawara, N. & Szostak, J. W. *Yeast* **2** (*supp.*) 373 (1986).
4. Blackburn, E. H. & Gall, J. G. *J. molec. Biol.* **120**, 33–53 (1978).
5. Klobutcher, L. A.,Swanton, M. A., Donini, P. & Prestcott, D. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3015–3019 (1981).
6. Ponzi, M., Pace, T., Dore, E. & Frontali, C. *EMBO. J.* **4**, 2991–2995 (1985).
7. Johnson, E. M. *Cell* **22**, 875–886 (1980).
8. Emery, H. S. & Weiner, A. M. *Cell* **26**, 411–419 (1981).
9. Blackburn, E. H. & Challoner, P. B. *Cell* **36**, 447–457 (1984).
10. Van der Ploeg, L. H. T., Lin, A. Y. C. & Borst, P. *Cell* **36**, 459–468 (1984).
11. Cooke, H. J., Brown, W. R. A. & Rappold, G. A. *Nature* **317**, 687–692 (1985).
12. Allshire, R. C. *et al. Cell* **50**, 391–403 (1987).
13. Pluta, A. F., Dani, G. M., Spear, B. B. & Zakian, V. A. *Proc. Acad. Sci. U.S. A.* **78**, 1475–1479 (1984).
14. Perrot, M., Barreau, C. & Bégueret, J. *Molec. cell. Biol.* **7**, 1725–1730 (1987).
15. Cooke, H. J. & Smith, B. A. *Cold Spring Harbor Symp. Quant. Biol.* **51**, 213–219 (1986).
16. Meehan, R. R. *et al. Am. J. hum. Genet.* **42**, 26–37 (1988).
17. Gosden, J. R., Middleton, P. G., Rout, D. & De Angelis, C. *Cytogenet. Cell Genet.* **43**, 150–153 (1986).
18. Liou, G. I. *et al. Somat. Cell. molec. Genet.* **13**, 315–323 (1987).
19. Larson, D. D., Spangler, E. A. & Blackburn, E. H. *Cell* **50**, 477–483 (1987).
20. Bernards, A., Michels, P. A. M., Lincke, C. R. & Borst, P. *Nature* **303**, 592–547 (1983).
21. Pays, E., Laurent, M., De Linte, K., Van Meirvenne, N. & Steinert, M. *Nucleic Acids Res.* **11**, 8137–8147 (1983).
22. De Grouchy, J., Turleau, C., Finaz, C. & Roubin, M. *Chromosome Variation in Human Evolution* (*ed.* Boyce, A. J.) 17–38 (Taylor & Francis, London, 1975).
23. Lejeune, J., Dutrillaux, B., Rethoré, M. O. & Prieur, M. *Chromosoma* **43**, 423–444 (1973).
24. Yunis, J. J. & Prakash, O. *Science* **215**, 1525–1530 (1982).
25. Sutherland, G. R. & Hecht, F. *Oxford Monographs on Medical genetics* **13**, (Oxford University Press, 1985).
26. Henderson, E. *et al. Cell* **51**, 899–908 (1987).
27. Greider, C. W. & Blackburn, E. H. *Cell* **51**, 887–898 (1987).
28. Gilliam, T. C. *et al. Cell* **50**, 565–571 (1987).

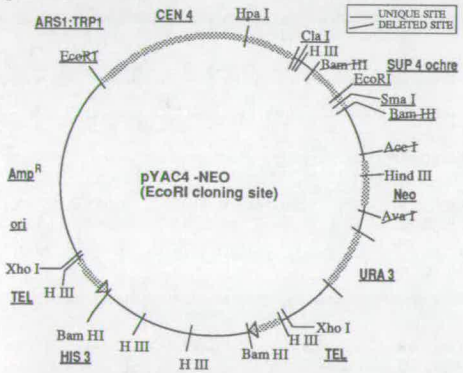
pYAC-4 Neo, a yeast artificial chromosome vector which codes for G418 resistance in mammalian cells

Howard Cooke and Sally Cross

MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, UK
Submitted November 21, 1988

Yeast artificial chromosome vectors allow DNA molecules in the 100-1000kb range to be cloned and propagated as linear molecules in S. cerevisiae (1). The increase in the size of clonable fragments is of use in constructing physical maps of genomes and in correlating the physical and genetic maps (2). There will be many circumstances in which it is desirable to reintroduce such yeast artificial chromosomes into mammalian cells, particularly where a large functional unit is involved or suspected to exist. To facilitate the selection of cells which have taken up YACs we have constructed a yeast artificial chromosome vector which contains a neomycin resistance gene driven by SV40 72 base pair repeats.

pYAC-4-Neo is derived from pYAC-4 (1). As a plasmid in E. Coli it confers Kanamycin and Ampicillin resistance, it carries Tetrahymena telomeric sequences which can function in yeast, a yeast centromere and ARS and the yeast markers URA3 TRP1 and SUP4. In mammalian cells the pSV2Neo (3) derived neomycin gene is expressed and allows selection for G418 resistance. pYAC-4-Neo was constructed by filling in the Sal 1 site of the parent vector and ligating this to an Acc1- Aval fragment of pSV2Neo after filling in these sites. The orientation of the fragment from pSV2Neo in the final construct was determined by size measurements of double digestion products. Cloning sites which allow colour and positive selection of recombinants are Eco Rl and Sma 1. Yeast transformation frequencies are comparable to the parent vector.



We constructed this vector in order to be able to reintroduce YAC clones into mammalian cells for functional analysis of large fragments of DNA. Others of the YAC series of vectors can be simply adapted to give Neomycin resistance by the same approach or, in the case of YAC vectors in which the Sma 1 site is unaltered, by substitution of the Cla 1/Sma 1 fragment from such vectors into pYAC-4-Neo.

Acknowledgements

We should like to thank D.T.Burke, G.F. Carle and M.V. Olsen for YAC vectors and yeast strains.

References

1) Burke D.T., Carle G.F. & Olsen M.V.: Science 236 806-812 1987
2) Coulson A., Waterston R., Kiff J., Sulston J. & Kohara Y.: Nature 335 184-186 1988
3) Southern P. & Berg P.: J. Mol Appl Genet 1 322-341 1982

Mouse lipocortin I cDNA

Tsuneaki Sakata*, Shoji Iwagami, Yuji Tsuruta, Ryuji Suzuki, Kanji Hojo, Kosaburo Sato and Hiroshi Teraoka

Shionogi Research Laboratories, Shionogi & Co., Ltd, Fukushima-ku, Osaka 553, Japan

Submitted November 18, 1988 Accession no. X07486

Mouse lipocortin I cDNA was cloned from a λgt10 cDNA library derived from mouse lung mRNA using rat lipocortin I cDNA fragment as a probe. Like the human (1) and rat (2) genes, this cDNA had an open reading frame of 1,038 nucleotides which encoded a protein of 346 amino acids. Nucleotide sequence homologies of the entire protein coding region between mouse and rat and between mouse and human lipocortin I were 93.9% and 86.8%, respectively. Its amino acid sequence homologies to rat and human were also 93.4% and 87.3%, respectively. As noted our previous report, the amino acid sequence homology between rat and human was 89% (2). Together with these results, the primary structures of lipocortin I were highly homologous among mouse, rat and human.

```
TACTTCTCTAAAA
10 20 30 40 50 60 70 80 90 100 110 120
ATGGCAATGGTATCAGAATTCCTCAAGCAGGCCCGTTTTCTTGAAAATCAAGAACAGGAATATGTTCAAGCTGTAAAATCATACAAAGGTGGTCCTGGGTCAGCAGTGACCCCCTACCCT
M A M V S E F L K Q A R F L E N Q E Q E Y V Q A V K S Y K G G P G S A V S P Y P
130 140 150 160 170 180 190 200 210 220 230 240
TCCTTCAATGTATCCTCGGATGTTGCTGCCTTGCACAAAGCTATCATGGTTAAAGGTGTGGATGAAGCAACCATCATTGACATTCTTACCAAGAGGACCAATGCTCAGCGCCAGCAGATC
S F N V S S D V A A L H K A I M V K G V D E A T I I D I L T K R T N A Q R Q Q I
250 260 270 280 290 300 310 320 330 340 350 360
AAGGCCGCGTACTTACAGGAGAATGGAAAGCCCTTGGATGAAGTCTTGAGAAAAGCCCTTACAGGCCACCTGGAGGAGGTTGTTTTGGCTATGCTAAAAACTCCAGCTCAGTTTGATGCA
K A A Y L Q E N G K P L D E V L R K A L T G H L E E V V L A M L K T P A Q F D A
370 380 390 400 410 420 430 440 450 460 470 480
GATGAACTCCGTGGTGCCATGAAGGGACTTGGAACAGATGAAGACACTCTCATTGAGATTTTGACAACAAGATCTAACGAACAAATCAGAGAGATTAATAGAGTCTACAGAGAAGAGCTG
D E L R G A M K G L G T D E D T L I E I L T T R S N E Q I R E I N R V Y R E E L
490 500 510 520 530 540 550 560 570 580 590 600
AAAAGAGATCTGGCCAAAGACATCACTTCAGATACATCTGGAGACTTTCGGAAAGCCTTGCTTGCTCTTGCCAAGGGTGACCGTTGTCAGGACTTGAGTGTGAATCAAGATTTGGCTGAT
K R D L A K D I T S D T S G D F R K A L L A L A K G D R C Q D L S V N Q D L A D
610 620 630 640 650 660 670 680 690 700 710 720
ACAGATGCCAGGGCTTTGTATGAAGCTGGAGAAAGGAGAAAGGGGACAGACGTGAATGTCTTCACCACAATTCTGACCAGCAGGAGCTTTCCTCATCTTCGCAGAGTGTTTCAGAATTAC
T D A R A L Y E A G E R R K G T D V N V F T T I L T S R S F P H L R R V F Q N Y
730 740 750 760 770 780 790 800 810 820 830 840
GGAAAGTACAGTCAACATGACATGAACAAAGCTCTGGATCTGGAACTGAAGGGTGACATTGAGAAGTGCCTCACAACCATCGTGAAGTGTGCCACCAGCACTCCAGCTTTCTTTGCCGAG
G K Y S Q H D M N K A L D L E L K G D I E K C L T T I V K C A T S T P A F F A E
850 860 870 880 890 900 910 920 930 940 950 960
AAGCTGTACGAAGCCATGAAGGGTGCCGGAACTCGCCATAAGGCATTGATCAGGATTATGGTCTCCCGTTCGGAAATTGACATGAATGAAATCAAAGTATTTTACCAGAAGAAGTATGGA
K L Y E A M K G A G T R H K A L I R I M V S R S E I D M N E I K V F Y Q K K Y G
970 980 990 1000 1010 1020 1030 1040
ATCTCTCTTTGCCAAGCCATCCTGGATGAAACCAAAGGAGACTATGAAAAAATCCTGGTGGCTCTGTGTGGTGGAAACTAGACATCCCAACTATTCTGCAAGGTTCTGAGGAGAATGTCT
I S L C Q A I L D E T K G D Y E K I L V A L C G G N *

CTTAGCTGTTTTCCTTCGATGGCATGGGCTTAAGTAGGAAAGTTGCTTTGGCAGATAAGTCTAATTACCTGCTTTGAATAATATAGCCTATAAATAGATTTTACATCATTACTCTGTACA

ATAGAGAAATACTTGTTTTGTTAATTATGTTTATCCCAAATTATAAAGCCCCATAAGCAAGTCACTTTGGTACCATTCCTGAGAAAGAAGTTTACATATAATAAAATAAAACAATTTTAT

AAGACAAAAAAAAAAAAAAAAA
```

* To whom correspondence should be addressed

References

1) Wallner, B. P., et al. (1986) Nature 320, 77-81.
2) Tamaki, M., et al. (1987) Nucleic Acids Res. 15, 7673.

Cloning of human telomeres by complementation in yeast

Sally H. Cross, Robin C. Allshire, Stewart J. McKay, Nio I. McGill & Howard J. Cooke

MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK

TELOMERES confer stability on chromosomes by protecting them from degradation and recombination and by allowing complete replication of the end[1]. They are genetically important as they define the ends of the linkage map. Telomeres of lower eukaryotes contain short repeats consisting of a G-rich and a C-rich strand, the G-rich strand running 5′–3′ towards the telomere and extending at the end[2–11]. Telomeres of human chromosomes share characteristics with those of lower eukaryotes[12,13], including sequence similarity as detected by cross-hybridization[14–16]. Telomeric repeats from many organisms can provide telomere function in yeast[2,17]. Here we describe a modified yeast artificial chromosome (YAC) vector with only one telomere which we used to clone human telomeres by complementation in yeast. YACs containing human telomeres were identified by hydridization to an oligonucleotide of the trypanosome telomeric repeat. A subcloned human fragment from one such YAC is immediately subtelomeric on at least one human chromosome.

Human DNA digested with restriction enzymes which have four base-pair (bp) recognition sequences displays a range of fragments which hybridize to oligonucleotides containing the trypanosome telomeric repeat sequence TTAGGG. In the case of *Sau*3A, this range of fragments is centred at about 10 kilobases (kb) in DNA isolated from peripheral blood cells. These fragments are lost after brief digestion of intact DNA with

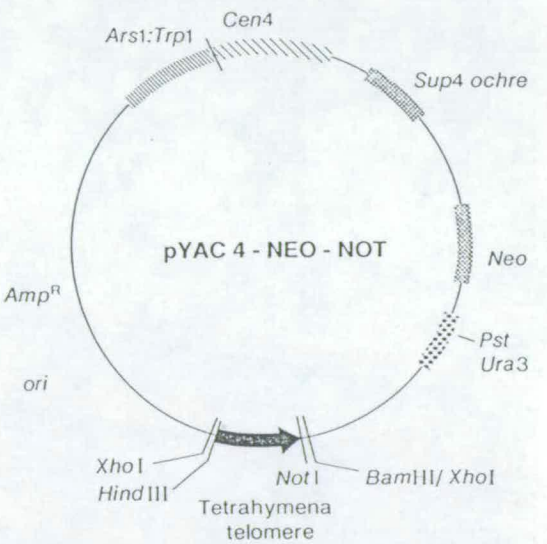


FIG. 1 Cloning of human sequences in pYAC4-Neo-Not. pYAC4-Neo[16] was cut to completion with *Bam*HI and partially with *Xho*I. The adaptor sequence 5′TCGAGGGATCCGCGGCCGC3′, base-paired with 15 of 19 bases, and having TCGA and GATC unpaired at the 5′ ends, was ligated as described[28]. The adaptor destroys the original *Bam*HI site and recreates the *Xho*I site. The *his3* gene and one of the tetrahymena telomeric repeats $(TTGGGG)_n$ in pYAC4-Neo are deleted. A *Not*I restriction site is introduced adjacent to the remaining tetrahymena telomeric repeat and a *Bam*HI site between the *Not*I and *Xho*I sites. The structure shown was confirmed by restriction site mapping. Human placental DNA between 5 and 15 kb after *Sau*3A digestion was recovered from an agarose gel and ligated to pYAC4-Neo-Not which had been digested with *Not*1, dephosphorylated and digested with *Bam*HI. This DNA was used to transform AB1380 spheroplasts.

exonuclease *Bal*31 (ref. 15). We have modified a YAC vector pYAC4 Neo (ref. 18) to give a vector (pYAC4-Neo-Not) which requires the addition of a second sequence with telomere function for its maintenance as a linear minichromosome in yeast. The presence of the *sup4* gene in the vector provides a colour test for the maintenance of this marker; AB1380 (ref. 19) yeast colonies are red in the absence and white in the presence of this gene. Details of the vector are given in the legend to Fig. 1. *Sau*3A-digested placental DNA from 5–15 kb was used as an enriched source of human telomeres for ligation to pYAC4-Neo-Not digested with *Not*I and *Bam*HI.

AB1380 was transformed[20] with the ligation mix and selected for Ura^+ transformants. Ura^+ transformants (240) were screened for the presence of human telomeric sequences on the basis of colony hybridization to the oligonucleotide $(TTAGGG)_4$. Two strongly hybridizing positive colonies were found (Hutel 1 and Hutel 22). Pulsed-field gel electrophoresis (PFG) comparison of AB1380 and Hutel 1 chromosomes showed that in this recombinant the 1.5 megabase (Mb) yeast chromosome 4 was replaced by two new chromosomes of 1.0 and 0.5 Mb respectively (Fig. 2*a*). Hybridization analysis shows that these chromosomes are the product of recombination between the transforming molecule and yeast chromosome 4 (Fig. 2*a*, *c*). In the case of Hutel 22, PFG analysis showed the presence of a 12-kb linear artificial chromosome, as shown in Fig. 2*b*. When plated on non-selective media, a high proportion of colonies were red as we expected, because a linear chromosome of this size should be relatively unstable[21]. In contrast, Hutel 1 gave less than 5 per cent red colonies when plated on non-selective media (data not shown).

We wished to know if the human sequences in Hutel 1 and Hutel 22 were present at the end of the new yeast chromosomes. First, the end of a DNA molecule presents as an apparent site for all restriction enzymes. Such a site was present at the distal end of the region of human DNA homologous to the $(TTAGGG)_4$ oligonucleotide probe for both Hutel 1 and Hutel 22 (data not shown). Two additional features should be exhibited by terminal restriction fragments with human homology. They should be decreased in size and eventually become undetectable when intact recombinant yeast DNA is digested with exonuclease *Bal*31. Second, these terminal fragments should be heterogeneous in size when compared with fragments of the genome which are bounded by two restriction sites[22]. Figure 3 shows that both of these criteria are satisfied for Hutel 1. We consider that the most likely explanation for this heterogeneity is the addition of yeast terminal repeats. We conclude that Hutel 1 and Hutel 22 contain human DNA fragments with homology to trypanosome telomeric repeats and that these sequences support telomere function in a yeast.

Although >95 per cent of sequences which hybridize to $(TTAGGG)_4$ are *Bal*31-sensitive in intact human DNA and hence presumably telomeric or closely subtelomeric, it was important to show that the human sequences that we cloned in yeast are telomeric in the human genome. We have cloned some of the human sequences from Hutel 1 into a plasmid (pHutel 1), as described in the legend to Fig. 4. This cloned DNA fragment has no homology to the $(TTAGGG)_n$ terminal repeats detectable by hybridization or by sequence analysis (data not shown). When hybridized to blots of human DNA treated with *Bal*31 for different times and then digested with *Sau*3A, this

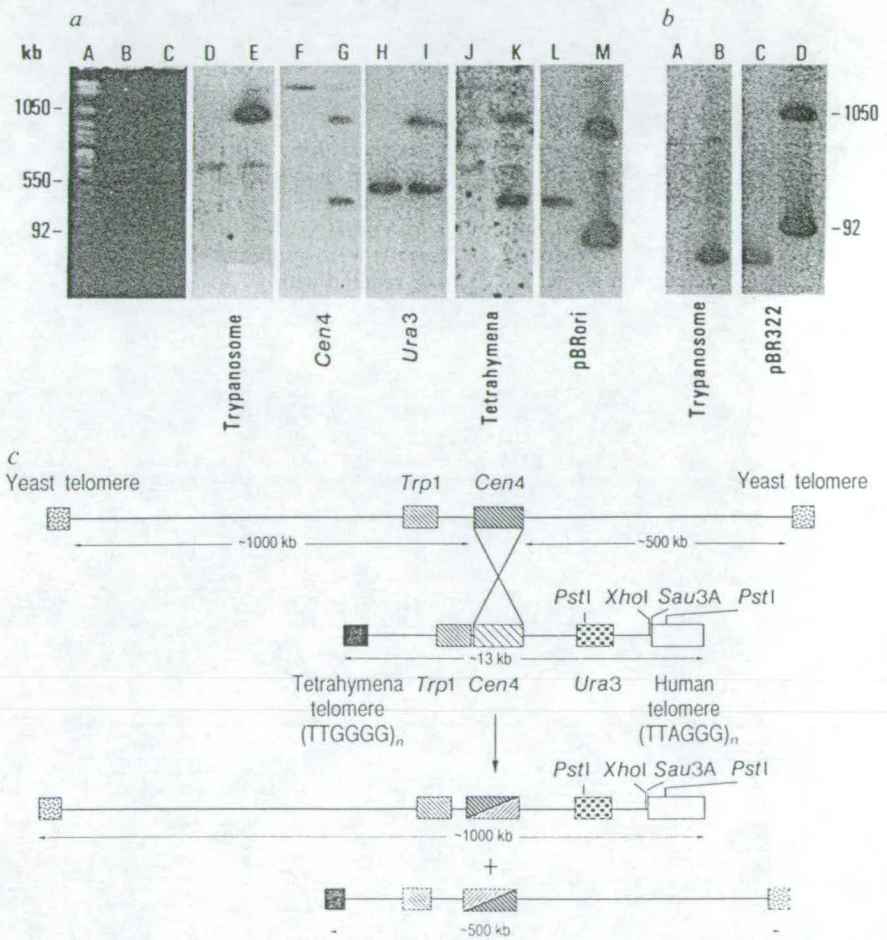


FIG. 2 Novel yeast chromosomes with human telomeres. *a*, Hutel 1 and AB1380 analysed by pulsed field gel electrophoresis. A–C, ethidium bromide stained gel; D and E, probed with $(TTAGGG)_4$; F and G, probed with a yeast *cen4* DNA fragment; H and I, probed with yeast *ura3*; J and K, probed with $(TTGGGG)_4$; L and M, probed with pBR *ori* probe (*Pvu*II–*Pst*I). A and M, yeast marker chromosomes; B, D, F, H and J, AB1380; C, E, G, I, K and L, Hutel 1. The larger of the novel chromosomes hybridized to $(TTAGGG)_4$ and to *ura3*. Both novel chromosomes hybridized to a *cen4* probe and the smaller chromosome hybridized with the *ori* region of pBR322 and the tetrahymena telomere oligonucleotide probe $(TTGGGG)_4$. Restriction analysis of Hutel 1 DNA showed that endogenous fragments were retained unchanged in size for the *trp1* and *ura3* genes, but that vector and endogenous *cen4* fragments were changed in size (data not shown). Homologous recombination between the *cen4* region of the vector and the original chromosome 4 would give rise to these chromosomes as detailed (Fig. 2*c*). *b*, Hutel 22 and AB1380 analysed by pulsed field gel electrophoresis. A and B probed with $(TTAGGG)_4$; C and D probed with pBR322. A, AB1380; B and C, Hutel 22; D, YP148 marker chromosomes. Yeast DNAs were prepared in agarose plugs as described[29]. They were run on a 1% agarose gel overnight using a waltzer apparatus. The gel was made and run in 0.25 × TBE (0.089 M Tris, 0.089 M boric acid, 0.002 M EDTA). Conditions were 7 V cm^{-1} at 10 °C and 90″ between 120° reorientations. The gel was blotted to Hybond-N and hybridized with the ^{32}P-labelled probes as listed above. Hybridization with random primed fragments was overnight at 68 °C in 7% SDS, 0.5 M Na_2HPO_4, pH 7.2, 0.5% dried milk powder. Filters were washed in 0.1 × SSC, 0.1% SDS at 65 °C. Hybridization with oligonucleotides was overnight at 50 °C in 0.1% SDS, 0.1% sodium pyrophosphate, 0.05% bovine serum albumin, 0.05% polyvinylpyrrolidone, 0.05% Ficoll, 5 × SSC. Filters were washed in 4 × SSC, 0.1% SDS at 65 °C. Size markers (kb) are indicated. *c*, Recombination between endogenous and artificial chromosomes. Restriction sites relevant to the chromosome construction and subcloning are indicated. Figure is not to scale.

probe detected *Bal*31-sensitive fragments which were the same size as those detected by a $(TTAGGG)_4$ oligonucleotide probe (Fig. 4*a*). In addition, a number of discrete fragments are detected which do not correspond in size to fragments detected by the $(TTAGGG)_4$ probe. When hybridized to digests of blood and sperm DNA from different individuals, pHutel 1 DNA detected heterogeneous smears which were the same size as those detected by the oligonucleotide $(TTAGGG)_4$ and which were larger in sperm than in blood DNAs (Fig. 4*b*). This size difference between these tissues from a single individual is a characteristic of the human pseudoautosomal telomeres[13]. From the restriction map presented in Fig. 2, we would predict that this probe would not detect heterogenous fragments in a *Pst*1 digest of human DNA. As shown in Fig. 4*b*, only discrete bands are detected in *Pst*1 digests by pHutel 1. We conclude that this sequence must be immediately subtelomeric in the human genome.

Telomeres of all organisms studied consist of short tandemly

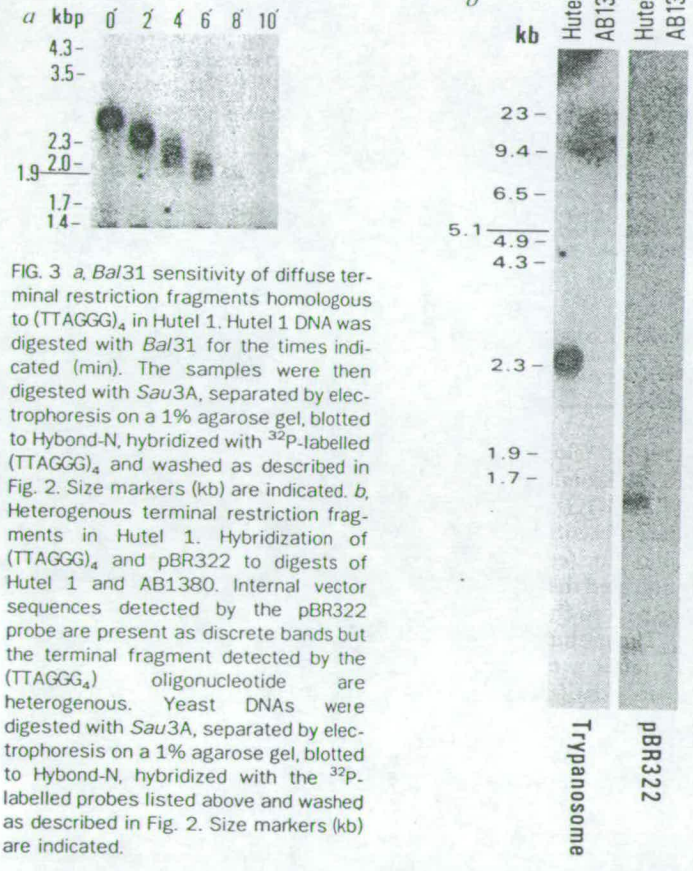


FIG. 3 *a*, *Bal*31 sensitivity of diffuse terminal restriction fragments homologous to $(TTAGGG)_4$ in Hutel 1. Hutel 1 DNA was digested with *Bal*31 for the times indicated (min). The samples were then digested with *Sau*3A, separated by electrophoresis on a 1% agarose gel, blotted to Hybond-N, hybridized with ^{32}P-labelled $(TTAGGG)_4$ and washed as described in Fig. 2. Size markers (kb) are indicated. *b*, Heterogenous terminal restriction fragments in Hutel 1. Hybridization of $(TTAGGG)_4$ and pBR322 to digests of Hutel 1 and AB1380. Internal vector sequences detected by the pBR322 probe are present as discrete bands but the terminal fragment detected by the $(TTAGGG_4)$ oligonucleotide are heterogenous. Yeast DNAs were digested with *Sau*3A, separated by electrophoresis on a 1% agarose gel, blotted to Hybond-N, hybridized with the ^{32}P-labelled probes listed above and washed as described in Fig. 2. Size markers (kb) are indicated.

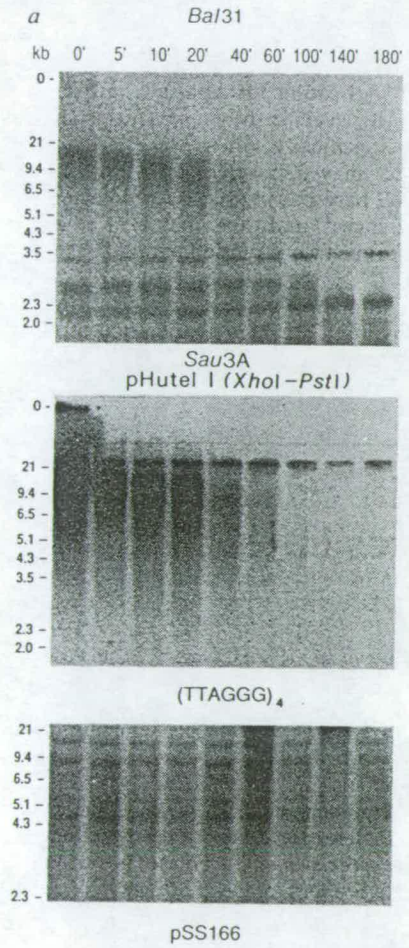
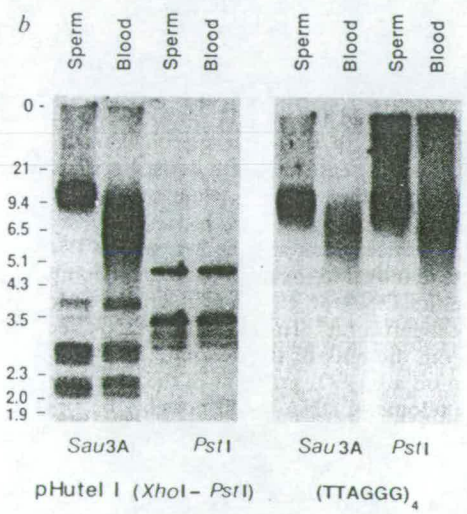


FIG. 4 Terminal location of pHutel 1 human sequences in the human genome. The human sequence in Hutel 1 was restriction mapped and a *Pst*1 site proximal to the terminal repeats was found which, in the recombinant, lies on a *Pst*1 fragment which also contains part of the yeast marker *ura3*. This fragment (pHutel 1) was recovered from a *Pst*1 library of Hutel 1 in pTZ 19 using a *ura3* DNA fragment as a probe. *a*, Human placental DNA was digested with *Bal*31 for the times indicated (min). The samples were then split and digested with *Sau*3A or *Pst*I, separated by electrophoresis on 0.8% agarose gels and blotted to Hybond-N. The filter with the *Bal*31 series digested with *Sau*3A was first hybridized with ^{32}P-labelled *Xho*I–*Pst*I fragment containing human sequences from pHutel 1 and washed as described in Fig. 2. The filter was then stripped as recommended by the manufacturers, hybridized with ^{32}P-labelled $(TTAGGG)_4$ and washed as described in Fig. 2. The filter with the *Bal*31 series digested with *Pst*I was hybridized with ^{32}P-labelled pSS166 (*Hin*dIII fragment from cosmid B4 (ref. 30)). Hybridization was at 60 °C, but otherwise as described in Fig. 2. The filter was washed in 2 × SSC, 0.1% SDS at 65 °C. ○, Gel origin; marker sizes are given in kb. *b*, Blood and sperm DNAs from the same individual were digested with *Sau*3A or *Pst*I, separated by electrophoresis on a 0.8% agarose gel, blotted to Hybond-N and hybridized with the ^{32}P-labelled probes as indicated. Hybridization and washing conditions were as described in Fig. 1. Size markers (kb) are indicated; ○, gel origin.



repeated sequences. Variation in copy number of these repeats results in the heterogeneity observed for terminal restriction fragments[23,24]. In yeast the terminal repeats are the only elements required for telomere function[22], and telomeric repeats from a number of organisms are capable of function in yeast by the addition of yeast sequences. Previous data[12–15] suggested that human telomeres showed many of the structural features of telomeres of lower eukaryotes. We have shown here that a telomeric fragment of human DNA can function as a telomere in yeast, suggesting that the structure of the ends of human chromosomes conforms to the general model of a number of short tandem repeats with a G-rich strand and a C-rich strand: the G-rich strand runs 5′–3′ towards the end of the molecule and ends in a short single-strand extension. This degree of conservation and cross-kingdom function may reflect some similarities of replication mechanisms and protein interaction at the chromatin level. In other organisms there are proteins which interact specifically with the telomeres[24,25]. Alternatively, these sequences may themselves form functionally significant structures. G–G base pairing has been detected in tetrahymena telomeric sequences[26] and there may also be a requirement for the ability to form a four-stranded parallel base-paired structure in meiosis[27].

Cloning of human telomeres in yeast provides a powerful method of obtaining markers at the ends of the physical and genetic maps of the human genome. Linkage of these markers to distal markers on the human linkage map will define the length of this map. □

Received 28 December 1988; accepted 23 March 1989.

1. Blackburn, E. H. & Szostak, J. W. *A. Rev. Biochem.* **53,** 163–194 (1984).
2. Shampay, J., Szostak, J. W. & Blackburn, E. H. *Nature* **310,** 154–157 (1984).
3. Walmsley, R. W., Chan, C. S. M., Tye, B-K. & Petes, T. D. *Nature* **310,** 157–160 (1984).
4. Sugawara, N. & Szostak, J. W. *Yeast* **2** (Suppl.) 373 (1986).
5. Blackburn, E. H. & Gall, J. G. *J. molec. Biol.* **120,** 33–53 (1978).
6. Klobutcher, L. A., Swanton, M. A., Donini, P. & Prestcott, D. M. *Proc. natn. Acad. Sci. U.S.A.* **78,** 3015–3019 (1981).
7. Ponzi, M., Pace, T., Dore, E. & Frontali, C. *EMBO J.* **4,** 2991–2995 (1985).
8. Emery, H. S. & Weiner, A. M. *Cell* **26,** 411–419 (1981).
9. Blackburn, E. H. & Challoner, P. B. *Cell* **36,** 447–457 (1984).
10. Van der Ploeg, L. H. T., Liv, A. Y. C. & Borst, P. *Cell* **36,** 459–468 (1984).
11. Forney, J., Henderson, E. R. & Blackburn, E. H. *Nucleic Acids Res.* **15,** 9143–9152 (1987).
12. Cooke, H. J., Brown, W. R. A. & Rappold, G. A. *Nature* **317,** 687–692 (1985).
13. Cooke, H. J. & Smith, B. A. *Cold Spring Harbor Symp. quant. Biol.* **51,** 213–219 (1986).
14. Allshire, R. C. *et al. Nature* **332,** 656–659 (1988).
15. Moyzis, R. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **85,** 6622–6626 (1988).
16. Richards, E. J. & Ausubel, F. M. *Cell* **53,** 127–136 (1988).
17. Pluta, A. F., Dani, G. M., Spear, B. B. & Zakian, V. A. *Proc. natn. Acad. Sci. U.S.A.* **81,** 1475–1479 (1984).
18. Cooke, H. J. & Cross, S. H. *Nucleic Acids Res.* **16** 11817 (1988).
19. Burke, D. T., Carle, G. F. & Olsen, M. V. *Science* **236,** 806–812 (1987).
20. Burgers, P. M. J. & Percivals, K. *J. analyt. Biochem.* **163,** 391–397 (1987).
21. Murray, A. W., Schultes, N. P. & Szostak, J. W. *Cell* **45,** 529–536 (1986).
22. Szostak, J. W. & Blackburn, E. H. *Cell* **29,** 245–255 (1982).
24. Gottschling, D. E. & Zakian, V. A. *Cell* **47,** 195–205 (1986).
25. Berman, J., Tachibana, C. Y. & Tye, B-K. *Proc. natn. Acad. Sci. U.S.A.* **83,** 3713–3717 (1986).
27. Sen, D. & Gilbert, W. *Nature* **334,** 364–366 (1988).
28. Lathe, R., Kierny, M. P., Skory, S. & Lecocq, J. P. *DNA* **3,** 173–182 (1984).
29. Bellis, M., Pages, M. & Roizes, G. *Nucleic Acids Res.* **15,** 6747 (1987).
30. Nakaseko, Y., Adachi, Y., Funahashi, S., Niwa, O. & Yanagida, M. *EMBO J.* **5,** 1011–1021 (1986).

ACKNOWLEDGEMENTS. We thank those who donated blood and sperm; C. Greider and J. Inglis for oligonucleotides; N. Davidson, S. Bruce and D. Stewart for photographic work and Professor H. J. Evans for encouragement and support. S.H.C. is supported by an SERC studentship.