The structure of alphoid satellite DNA on normal and abnormal human Y chromosomes

by

Rebecca Oakey

Thesis submitted for the degree of Doctor of Philosophy
at the University of Oxford

Linacre College
Michaelmas term 1989

Abstract

Rebecca Oakey
Linacre College

Submitted for the degree
of Doctor of Philosophy
Michaelmas term 1989

The structure of alphoid satellite DNA on normal and abnormal human Y chromosomes

The long-range structure of the Y chromosome alphoid satellite DNA has been determined in the cell lines 3E7 and OXEN. Variation in alphoid DNA block size and restriction enzyme sites were observed. The alphoid block size and restriction enzyme site variations were determined for a collection of 42 normal Y chromosomes. The alphoid DNA polymorphisms observed defined 24 Y chromosome alleles. Unexpectedly, the Y alphoid DNA alleles analysed revealed two distinct groups of Y chromosomes indicating that most of the Caucasian and Asian men analysed were descended from one of two males.

The structure of the alphoid DNA was determined for 25 cell lines expected to contain abnormal Y chromosomes. Six of the cell lines lacked Y chromosomes. Four lacked both alphoid DNA and Y a centromere. 13 out of the remaining 15 Y chromosomes had centromeres and Y alphoid DNA block sizes and restriction enzyme site variation similar to that of normal Y chromosome alphoid DNA. Two of the abnormal cell lines had alphoid DNA blocks significantly different from the normal Y alphoid DNA structure. These results confirm that alphoid DNA is located very close to, or at the centromere and make it a prime candidate for a functional mammalian centromere sequence.

Acknowledgements

I would like to thank Chris Tyler-Smith for all his help and guidance throughout the last three years, and for reading and discussing this thesis. I thank E.M. Southern for his constructive criticism and suggestions and for reading this thesis. My thanks go to Lesley Taylor for tissue culture instruction and punctuation advice; Martin Johnson for maintaining 'Matilda' and my bike and Juliet Honeycombe for her help with chromosome preparation and moral support. I thank John Elder for writing and executing the computer simulation program and for his help in compiling this thesis. I thank Tasnim Alam and Kay Sims for typing the references and Phil Bates, Andy Porter, Alfredo Villasante and Helen Taylor for assorted help and encouragement.

I thank all the people who kindly donated blood samples and Richard Fisher, Thomas Mitinger, Andy Copp and Charles Bangham for taking the blood samples. My thanks also go to the individuals who generously donated cell lines. I thank the Medical Research Council for funding me during the three years and the MRC, Graduate studies and Linacre college for supplying funds for attending conferences.

I thank my mother and all my family for their financial support and encouragement during my three years and Michael Malim for his endless patience and help throughout.

Contents

Chapter 1

Introduction

1.1 The Centromere

Chromosomes carry genetic information from one generation to the next. Their correct segregation from mother to daughter cell is crucial in maintaining the integrity of an organism. At least three types of chromosomal elements are necessary in *cis* to accomplish eukaryotic replication and segregation. One is the origin of replication, a site on the chromosome where synthesis of new DNA begins, of which there are many per chromosome. A second element is the telomere, a repetitive sequence present at each end of a chromosome needed to replicate terminal DNA and protect it from degradation. The third element is the centromere, a sequence of DNA which is responsible for correct chromosome segregation and to which the spindle fibres attach during mitosis.

In mitosis the centromere is required for the segregation of sister chromatids, and in meiosis, for segregation of homologous chromosomes in the first meiotic division and of sister chromatids in the second meiotic division. The centromere is the spindle fibre attachment site of a chromosome. The kinetochore, a complex proteinaceous body attached laterally to the centromeric DNA, is the structure with which the ends of the chromosomal microtubules are associated for chromosome segregation. The centromeric DNA sequences responsible for spindle fibre attachment and for directing sister chromatid segregation in yeasts have been characterised but in mammalian chromosomes they are as yet unidentified.

The unique ability of the centromere to maintain correct, faithful segregation in chromosomes makes its role in cell division an important one. Different approaches have been used to analyse the mammalian centromere. The centromere has been examined directly by cytogenetic analysis using the light and electron microscope [Comings and Okada 1971]. Proteins associated with the centromere and the association between kinetochore proteins and centromere regions have been investigated using anti-kinetochore antibodies from CREST scleroderma patients [Moroi *et al.* 1980, Lica *et al.* 1986]. However, cytogenetic observations and analy-

sis of proteins associated with mammalian centromeres have not identified all important functional elements. Another molecular approach, which has been successful in lower eukaryotes, is to analyse centromere structure at the DNA level. In mammalian systems, such studies are complicated by the large size and complex nature of chromosomes, and the lack of a functional assay for centromeric activity. Thus the first step in the molecular analysis of mammalian centromere function is the elucidation of the detailed structure of DNA sequences closely linked to the centromere.

Yeasts, due to their ease of biochemical and genetic manipulation, small genomes and the ability to assay for centromere function, have provided systems for identification of important centromere components. The next two sections summarise what is known about yeast centromere structure and how this use of structural analysis has led to the identification of sequences responsible for centromere function. Little or nothing is known about mammalian centromere sequences. The choice of a centromere linked sequence and a model chromosome for study is described in the remaining sections.

1.2 Characterization of centromere sequences in the yeast *Saccharomyces cerevisiae*

Saccharomyces cerevisiae is a budding yeast with sixteen chromosomes. In this yeast, centromere elements have been cloned [Clarke and Carbon 1980] first by identifying segments containing flanking centromere linked genes followed by chromosome walking through the centromere. Hybrid plasmids containing DNA segments from around the centromere plus an autonomously replicating sequence (ARS) [Stinchcomb *et al.* 1979, Hsiao and Carbon 1979] were constructed. Unlike other ARS bearing plasmids, these centromere containing plasmids were fairly stably maintained through mitosis and meiosis in yeast. The centromere containing region in chromosome III is known as CEN3 and enables ARS bearing plasmids to behave as minichromosomes in mitosis, meiosis, sporulation and germination. Stability is a predicted property of a replicating unit carrying a functional centromere. The stabilizing element was shown to be confined to a few hundred bp of DNA which was sequenced for CEN3 (chromosome III) and CEN11 (chromosome XI) [Fitzgerald-Hayes *et al.* 1982a] and CEN6 (chromosome VI) [Panzeri and Philippsen 1982], and three conserved regions were identified. Centromere DNA sequence elements (CDE) are: I, a 14bp sequence, II a very (80%) A+T rich region 87-89bp long and CDEIII an 11bp region. The organisation of the centromeric elements are summarised in figure 1.1 (page 9).

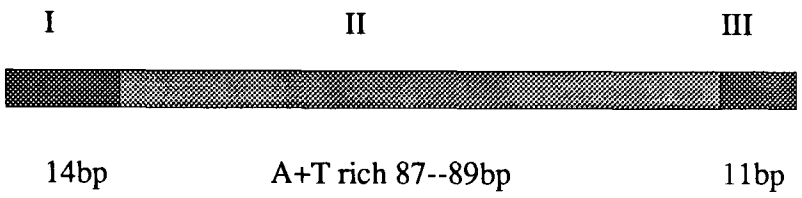
A more direct method for generating large numbers of CEN sequences from chromosomes IV, XI, I, XVI, XIV, X, VI, III, XV and VII [Hieter *et al.* 1985] has been used. In these experiments an ARS plasmid was exploited which is lethal in

high copy number. Due to mother/daughter segregation bias encountered with ARS containing plasmids, the only stable plasmids produced were those which contained a functioning CEN sequence which was able to override the lethality of the plasmid by providing a segregation mechanism and thus controlling copy number. This direct method provided a rapid way of cloning CEN sequences, and 10 out of the 16 CEN sequences were isolated. Sequence alignment revealed the common elements CDEI, an 8bp domain and CDEIII a 25bp domain. These sizes were slightly different to the previous size estimations by [Fitzgerald-Hayes *et al.* 1982b], but CDEII was the same. The right hand boundary of an active centromere has been defined exactly [Hegemann *et al.* 1986] as a triple A sequence present in the CDEIII element of CEN6, as determined by *in vitro* generation and analysis of CDEIII mutants of CEN6 DNA. No sequence further to the right was needed for centromere function.

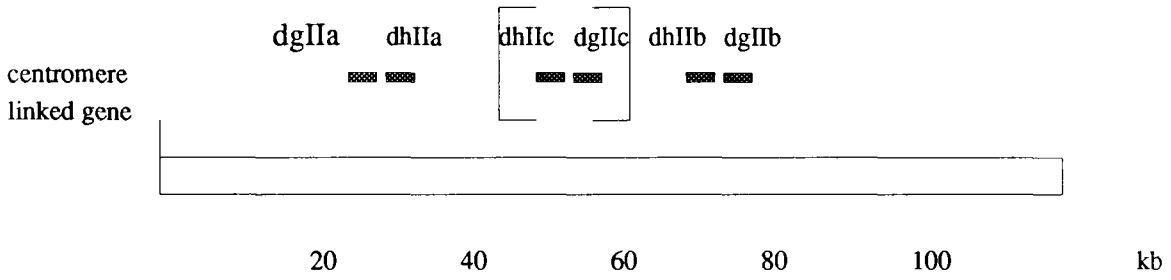
Minichromosomes do not always behave mitotically and meiotically as naturally occurring chromosomes; they tend to be slightly less stable [Clarke and Carbon 1980]. Studies of centromere sequences reintroduced into natural chromosomes provided an alternative approach. Sequence alterations, deletions and substitutions directly into normal yeast chromosomes [Clarke and Carbon 1983] using fragment mediated transformation [Rothstein 1983] allowed the effects of CEN deletions, inversions and replacements to be examined. Deleting the CEN3 sequence in chromosome III led to a dramatic instability of the resulting acentric chromosome, but replacement with CEN11 or inversion of either CEN3 or CEN11 made no difference to stability. [Bloom *et al.* 1984] have also shown that deletions of elements I-III completely inactivates the centromere, but substitution of CEN3 or CEN11 in either direction into chromosome III gives normal centromere function demonstrating the non-specificity of the CEN DNAs. [Huberman *et al.* 1986] have further analysed CEN sequences to determine whether they are conserved between strains of yeast with different degrees of divergence. In a study of two strains of *S.cerevisiae* and one of *Saccharomyces uvarum* where CEN sequences were cloned into a centromere selection vector, tested for centromere function and sequenced, the CEN sequences were found to be similar to the laboratory strain of *S.cerevisiae*. These systematic investigations into structure-function relationships in yeast centromeres have established that there are crucial sequences responsible for correct centromere function in yeast.

Artificially constructed yeast chromosomes have helped investigate the sequence requirements for a correctly segregating DNA molecule. Linear centromeric plasmids have been constructed with CEN regions, ARS elements and functional telomeres [Murray and Szostak 1983]. In their study, the authors found that short (7-15kb) linear plasmids were less stable than circular ones and were present in many copies per cell. Longer (55kb) plasmids, however, attained a much greater stability (although that were not as stable as naturally occurring chromosomes) and were present in one copy per cell, clearly indicating that these conditions of length and sequence components were approaching a more optimal, wild-type state than the shorter linear plasmids.

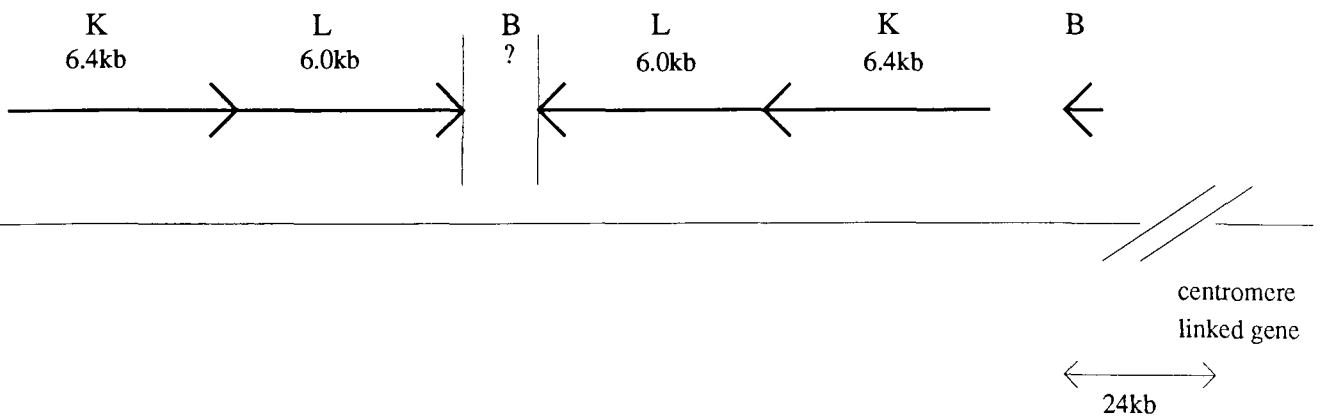
Figure1.1a. Diagramatic representation of the CEN sequences identified in S.cerevisiae (Fitzgerald-Hayes, 1982).



b. Diagramatic representation of the cen sequences identified on chromosome 2 of S.pombe (Chikashige et al, 1989)



c. Diagramatic representation of the cen sequences identified on chromosome 2 of S.pombe (Fishel et al, 1988)



1.3 Characterization of centromere sequences in the yeast *Schizosaccharomyces pombe*

Schizosaccharomyces pombe is evolutionarily distant from *S.cerevisiae* [Kaufer *et al.* 1985]. It is a fission yeast, with three chromosomes which are about six fold larger that those of *S.cerevisiae*, and are condensed during mitosis and separated by kinetochore microtubules in the spindle apparatus in a way more similar to those of higher eukaryotes than *S.cerevisiae*.

Maps have been constructed of the centromeric regions of the *S.pombe* chromosomes by two different groups. The nomenclature used to describe the sequences in these maps differs. [Nakaseko *et al.* 1986] constructed genetic and physical maps of the centromeric regions of I and II by chromosome walking from cloned centromere-linked genes. These maps revealed a 50kb long domain tightly linked to the centromere. Later these maps were found to be inaccurate due to rearrangement of cloned centromeric DNAs in plasmids. However, two classes of repetitive sequences were found. One was chromosome specific and called yn1. The other was designated dg and located in all three centromere regions. The dg elements were 3.8kb long. Another centromeric sequence element was found [Nakaseko *et al.* 1987] which was designated dh. This sequence element was 4–5kb long and was common to all three centromeric domains of the *S.pombe* chromosomes. The organisation of centromeric elements in chromosome 2 are summarised in figure 1.1.

The second group found moderately repetitive DNA elements which they called K, B and L organised into a long inverted repeat or palindrome [Fishel *et al.* 1988]. Centromeric DNA was isolated by chromosome walking. In chromosome 2 the repeat elements were organised into a 35kb inverted repeat with one copy of K and L in each arm of the repeat. The organisation of centromeric elements in chromosome 2 are summarised in figure 1.1.

The use of different nomenclature has complicated the picture of the *S.pombe* centromeric region. Essentially, in cen2, a 3.8kb portion of K was equivalent to dgII. A 4.0kb portion of K and L were equivalent to dhII [Fishel *et al.* 1988].

Direct analysis of centromere DNAs using genomic targeting and homologous recombination techniques [Chikashige *et al.* 1989] showed cen1 to be a 40kb stretch of DNA consisting of two inverted 10kb motifs each containing the centromeric elements dg and dh flanked by a central region. Cen2 was a 70kb stretch of DNA with 3 motifs arranged in inverted and direct orientations with flanking domains (summarised in figure 1.1). Cen3 had about 15 copies of dh-dg constituting a region longer than 100kb. A deletion series was made from a 530kb linear minichromosome containing the centric region of chromosome III [Niwa *et al.* 1989]. Cen3 was localised within a 120kb region comprising 15 centromeric repeated sequences dg-dh arranged in an inverted fashion with a central flanking region. Acentric minichromosomes lacking dg, dh and central flanking regions by deletion had a 100 fold decrease in stability compared with the original minichromosome.

A minichromosome assay for centromere function in *S.pombe* has been developed. The *S.cerevisiae* yeast artificial chromosome (YAC) system [Burke *et al.* 1987] has been used to clone large DNA restriction fragments from the *S.pombe* genome in *S.cerevisiae*. Linear artificial chromosomes with *S.pombe* centromeric DNA regions from chromosomes I and III (cen1 and cen3) were transformed into *S.pombe* and assayed for centromere function (*S.cerevisiae* CEN sequences do not function as centromeres in *S.pombe*). The minichromosomes were mitotically stable and segregated properly through meiosis (as determined by tetrad analysis) which indicated that the cloned fragments indeed contained functional *S.pombe* centromeres [Hahnenberger *et al.* 1989]. The centromere containing *Sal*I fragments of chromosomes I, II and III were sized to 65, 100 and 150kb respectively and had various numbers of copies of the centromere specific sequences K, L and B.

In summary, yeast centromere studies provide a model for the identification of structural and functional elements involved in centromere function in higher eukaryotes. Studies have shown that in both yeast species, mapping and cloning experiments can identify candidate sequences. A functional test can then show which of these are required for centromere activity.

1.4 Centromere sequences in mammals

There has been no work on the structure of mammalian centromeres comparable to the structural analysis of yeast. Little or nothing is known about the sequences involved in mammalian centromere function. The approach used in this study to begin the analysis of centromere-linked sequences was to build molecular maps of centromeric regions and genetic analysis of mutants. This was analogous to the construction of a physical map and physical and genetic analysis of aberrant chromosomes in yeast.

To begin a search for functionally important sequences, it was first necessary to analyse the structure of DNA in the region of the centromere. To do this, it was necessary to identify centromere-linked probes. In yeast, centromere-linked genes were exploited, and molecular maps were constructed by chromosome walking from these reference points. In mammalian genomes, there were no centromere-linked genes available for use. Unique sequences could also be used as centromere-linked probes, but when these studies were begun, no such sequences had been identified and no probes were available. However, probes were suitable for repeated sequences which have been located at the centromeres of human chromosomes by *in situ* hybridisation. The classical satellite sequences are widely distributed at the centromeric C-band regions of many chromosomes [Gosden *et al.* 1975] but not at every chromosome centromere. The 68bp *Sau*3A sequence [Agresti *et al.* 1989], was not at the time well characterised enough to feature as a candidate centromere-linked sequence. However, alphoid DNA has been found closely associated with all human chromosome centromeres. This property makes it a candidate sequence for a component of

the centromere. The next step in the analysis of structure was to construct molecular maps of the centromere-linked sequence of choice. The molecular structure of the alphoid DNA around the centromere of the human Y chromosome was studied by constructing long range restriction maps to determine the normal structure and variation that can be expected in a fully functioning centromere. The structure of the Y alphoid DNA from genetically mutant centromeres that have been inactivated or rearranged were compared to the normal case to see if centromere inactivity had a relationship with an altered Y alphoid DNA structure. In other words, were there any crucial sequence elements that were absent or altered in mutant centromeres compared to the normal functioning centromere? The mutants investigated have been obtained from naturally occurring and artificially made chromosome sources.

The methylation status of DNA can reflect the activity or inactivity of a gene (for a review, see [Bird 1987]). It was therefore decided to ask whether there was a difference in methylation pattern in the alphoid DNA in active and inactive centromeres. In order to observe patterns of methylation in stretches of DNA, restriction enzymes may be used. Most restriction enzymes are unable to cleave at methylated sequences. *Msp*I and *Hpa*II are commonly used for detecting methylC since they are isoschizomers that cleave at CCGG sites. If the internal C is methylated 5′-C^{m}CGG3′ *Hpa*II cannot cleave but *Msp*I can. If the external C is methylated 5′-^{m}CCGG3′ then the *Msp*I cannot cleave. Most methylation occurs at CpG. The distribution of fragments as determined by gel electrophoresis can allow conclusions to be drawn about the general methylation status of a genomic DNA sequence.

1.5 The human Y chromosome

There were a number of factors involved in the choice of the Y chromosome for the study of centromere structure. The Y chromosome is one of the smallest chromosomes in the human genome. The use of flow cytometry to measure the fluorescence of stained metaphase chromosomes, and hence the relative DNA content, showed only chromosomes 22 and 21 were smaller than the Y [Harris *et al.* 1986]. The Y has a relatively small block of alphoid DNA by comparison to other chromosomes. There is only one copy of the Y in a normal male, and this haploid nature makes it easier to study in the absence of a homologue.

The human Y chromosome has relatively few genes associated mainly with sex determination and male gametogenesis. The genes observed so far include the MIC2Y gene, the Y-linked XG locus [Goodfellow *et al.* 1983, Buckle *et al.* 1985] and the H-Y antigen gene [Simpson *et al.* 1987]. A candidate gene for maleness has been identified on the short arm of the Y [Page *et al.* 1987]. Spermatogenesis factors possibly located on Yq in the euchromatic region may be necessary for male fertility [Page 1986]. Naturally occurring mutant chromosomes are viable because of the lack of essential genes on the Y and are therefore available for study. The study of mutants, as discussed earlier, is useful for the systematic structural

analysis of mammalian chromosome elements, such as the centromere. The higher order periodicity of Y alphoid DNA and two Y chromosome alphoid blocks have been characterised previously [Tyler-Smith and Brown 1987], and provide a starting point for the construction of more detailed maps.

In summary, the small number of genes enabling the Y chromosome to be available in mutant forms plus the small size of the haploid chromosome and the presence of a small amount of alphoid DNA at its centromere makes it a good choice for the study of centromere structure.

1.6 Probes and mapping

1.6.1 Repeated sequences

A brief historical view of repeated sequences

A significant but variable component of the eukaryotic genome is comprised of highly repetitive DNA sequences that are often localised in constitutive heterochromatin that is not transcribed. Repeated sequences or satellite DNA are often associated with constitutive heterochromatin at centromeric, telomeric or intercalary heterochromatic sites and are C-band positive [John and Miklos 1979].

The first experiments that revealed the existence of satellite DNAs came from density gradient centrifugation where minor bands of different densities were visibly distinct from main band DNA. The term satellite DNA was coined to describe DNA sequences which banded at a different density from the main band. However, not all the repetitive sequences termed satellite can be separated by density gradients and can remain cryptic, needing the addition of metal ions for visualization or isolation [John and Miklos 1979]. Equilibrium sedimentation experiments in density gradients of mouse DNA led to the observation that the mouse had two bands, a main 1.701g/cm^3 band plus a second satellite band of 1.690g/cm^3 [Kit 1961]. The observation of a second minor component of DNA was also observed in guinea pig liver and spleen DNA, crab *Cancer borelis* and *C. irratorus* testes, calf thymus and mouse spleen and testes [Sueoka 1961]. Mouse satellite DNA was demonstrated by [Waring and Britten 1966]. In addition, the DNA in the satellite fraction reassociated much more quickly than main band DNA. This implied that the minor component consisted of short nucleotide sequences present in many copies. These experiments substantiated the hypothesis that some sequences are frequently repeated in vertebrate DNA. A measurement of reassociation is useful for comparing the relationships between sequences, and Cot analysis was developed to do just that [Britten and Kohne 1968]; DNA from organisms may be characterised by their Cot values.

Density gradient centrifugation and reassociation kinetics studies have been extended by the use of restriction enzyme analysis of repeated sequences. Restriction enzyme analysis can reveal repeat units and tandem organisation of a repeated se-

quence by the presence of integral multiples of the basic repeat unit length, or a ladder of fragments after partial digestion. Together, these methods achieved the isolation and characterisation of numerous classes of repeated sequences in eukaryotic organisms, a few of which are discussed below.

1.6.2 Primate repeated sequences

Two types of repeated sequences are found in primate genomes; highly repeated tandem sequences in long arrays (satellite DNA) and repetitive elements dispersed throughout the genome amid repeated and unique sequences [Singer 1982]. Interspersed repeated sequences are made up of two groups, short and long interspersed repeated sequences (SINES and LINES, [Skowronski and Singer 1986, Paulson *et al.* 1985]). These groups are repeated sequences, but are unrelated to the tandemly repeated sequences which are the subject of this study.

Tandemly repeated sequences

The human genome contains a large number of different tandemly repeated sequence families. These include the four major classical satellites I, II, III and IV, plus the 68bp family and the alphoid satellite DNA. Alphoid DNA is discussed in more detail in a separate section. A summary of the tandemly repeated sequences found in the human genome is shown in table 1.1.

The major classical satellites I, II and III in the human genome contain a collection of simple sequence components which have been defined in terms of their products of digestion with restriction enzymes. These simple repeated sequences are called satellites 1, 2 and 3 to indicate their enrichment of each one in classical satellites which contain other sequences as well [Prosser *et al.* 1986]. Satellite IV is nearly identical to satellite III in being highly enriched in satellite 3. The classical satellites are widely distributed in the centromeric C-band regions of many chromosomes as determined by *in situ* hybridisation. These include chromosomes 1, 5, 7, 9, 10, 12, 13, 14, 15, 16, 17, 20, 21, 22 and the distal C-band region of the long arm of the Y chromosome [Gosden *et al.* 1975].

The *Sau*3A, (68bp) family is a clustered highly repetitive DNA family. The bulk of the 68bp repeats occur in tandem arrays of at least 12 repeats [Meneveri *et al.* 1985] as determined by the presence of a ladder of fragments after *Sau*3A digestion. The main locations of the *Sau*3A family have been shown to be on heterochromatic blocks of acrocentric chromosomes (chromosomes 13, 14, 15, 21 and 22), and on the pericentric region of chromosome 1 [Agresti *et al.* 1989].

1.6.3 Alphoid DNA

The alphoid family of sequences was first seen by reassociation kinetics and modified centrifugation in Cs_2SO_4 in the presence of Ag^+ in the African green monkey (AGM)

Satellite	Repeat unit	Reference
1	AT rich sequences defined by *Rsa*I	[Prosser *et al.* 1986]
2	diverged 5bp ATTCC sequences	[Frommer *et al.* 1988]
3	5bp ATTCC (less diverged than 2)	"
*Sau*3A	68bp repeat, 53% G+C	[Meneveri *et al.* 1985]
*Hae*III	140–169bp repeat (closely related to the *Sau*3A family)	[Agresti *et al.* 1989]
alphoid DNA	171-173bp subunits organised into chromosome specific higher order repeat units	see below

Table 1.1: A summary of the tandemly repeated sequences found in the human genome

Cereopithicus aethiops [Maio 1971]. AGM alpha satellite mainly hybridises to the centromere region of all the chromosomes. Related, but non-identical alpha satellites are present in *Gorilla gorilla, Pan troglodytes, Cereocebus aterrimus, Macaca mulatta, Mandrillus sphinx, Papio cynocephalus, Cerecopithecus pygenthrus, Colobus badicus* and *Cebus capuchinus* [Maio 1971]. Restriction enzyme digests reveal a 172bp repeat in AGM alpha satellite. The repeats are not identical but represent a set of closely related sequences and constitute 13–20% of AGM DNA.

In humans, alphoid DNA is a tandemly repeated 170bp to 172bp sequence similar to the alpha sequence found in the AGM [Wu and Manuelidis 1978]. The 170bp units are arranged in arrays and can also be characterised by longer range periodicities marked by the presence of regularly spaced restriction enzyme sites that define a higher order repeat unit length. In some chromosomes the 170bp subunit is present as a 340bp dimer [Wu and Manuelidis 1980]. The 170bp subunits in a block may be divergent in sequence from one another, but the higher order repeats have been found to be more closely related, and thus probably represent the amplification unit or domain of repeating block responsible for spread of the sequence during evolution [Waye and Willard 1986]. Alphoid DNA has been suggested to constitute up to 5% of the human genome but the precise amount has not been firmly established.

A cloned human alphoid sequence called p82H [Mitchell *et al.* 1985] showed a tandemly repeated *Hae*III unit of length 162–177bp. *In situ* hybridisation of this sequence revealed hybridisation to every human chromosome centromere region. The cross hybridisation of alphoid DNA on different chromosomes indicates that alphoid DNA from different chromosomes have sequence features in common. Indeed, this fact has been exploited for the isolation of chromosome specific alphoid DNA sequences from chromosome libraries. However, many individual human chromosomes may be characterised by distinct alphoid satellite higher order repeat structures defined by restriction enzyme periodicity and sequence [Waye and Willard 1986]. The

Chro-mosome	Higher order repeat	No. of copies per chromosome	Reference
Y	5.7 and 6.0 *Hind*III	variable	[Wolfe *et al.* 1985] [Tyler-Smith and Brown 1987]
X	2.0kb *BamH*I	5000	[Yang *et al.* 1982] [Willard 1985]
17	2.7kb *EcoR*I	500–1000	[Waye and Willard 1986]
22	2.1 and 2.8kb *EcoR*I	-	[McDermid *et al.* 1986]
11	0.85kb *Xba*I	500	[Waye *et al.* 1987a]
7	2.7kb *EcoR*I	10	[Waye *et al.* 1987b]
	1.0kb *EcoR*I	500	
3	2.9kb *Hind*III	about 350	[Waye and Willard 1989]
	2.55kb *Hind*III		
6	3kb *BamH*I	-	[Jabs and Perisco 1987]
18	0.68kb *EcoR*I	-	[Devilee *et al.* 1986]
13, 14,15, 21 and 22	complicated set of subfamilies between these chromosomes	-	[Devilee *et al.* 1986] [Choo *et al.* 1988] [Choo *et al.* 1989] [Jorgensen *et al.* 1988]

Table 1.2: The size and copy number of the alphoid DNA higher order repeat units that have been characterised on some of the human chromosomes.

higher order repeat units have been characterised in a number of chromosomes, some of which are summarised in table 1.2 .

Table 1.2 illustrates the chromosome-specific higher order periodicities of alphoid DNA. In most cases, one periodicity has been reported for each chromosome. In some cases, more than one periodicity has been reported, for example for chromosome 7. An exception to these general cases are the acrocentric chromosomes (or nucleolus organising chromosomes) 13, 14, 15, 21 and 22. These chromosomes have a common subfamily of alphoid DNA as determined by *in situ* hybridisation [Choo *et al.* 1988]. However, more than one subfamily has been reported for these chromosomes. There were subfamilies reported in 13 and 21 [Devilee *et al.* 1986]; 14 and 22 [Jorgensen *et al.* 1988]. A further two subfamilies, one common to 13 and 21 and another common to 14 and 22 have also been reported [Choo *et al.* 1989]. Clearly the higher order structure of the alphoid DNA on the acrocentric chromosomes is more complicated than on other chromosomes.

1.6.4 Y chromosome alphoid DNA

The Y chromosome alphoid DNA has been studied in depth. By screening Y cosmids with the X alphoid DNA clone DXZ1, two Y alphoid DNA cosmids were isolated and localised to the centromeric Y DNA by *in situ* hybridisation. These

cosmids identified a 5.5kb *Eco*RI unit structure present an estimated 100 times on a Y chromosome [Wolfe *et al.* 1985]. Nine Y-specific alphoid DNA cosmids were obtained by screening a cosmid library from the Y only hybrid cell line 3E7 with the alphoid DNA clone p82H [Tyler-Smith and Brown 1987]. These cosmids were characterised by restriction enzyme analysis, and subclones mapped and sequenced to reveal 170bp subunits arranged in a repeat unit of 5.7kb. Some cosmids contained a few clustered 6.0kb units. *Ava*II digestion showed no internal sites for a 5.7kb unit but the presence of a site in the 6.0kb unit. *Hae*III digestion demonstrated the 170bp ladder of the tandem repeat: the 5.7kb unit had 34 170bp subunits and the 6.0kb unit had 36. Sequence analysis showed the 170bp subunits to be between 76% and 86% homologous to one another and homology to a consensus sequence of individual subunits was between 87% and 92%. Genomic analysis of two Y only somatic cell hybrids 3E7 and 853 showed that the 5.7kb unit had one site for *Eco*RI, *Hind*III and *Pst*I, 2 sites for *Xba*I and none for *Bam*HI, *Bcl*I, *Bgl*II and *Pvu*II. 3E7 but not 853 showed a minor band of 6.0kb for *Eco*RI and *Pst*I and a pair of bands at 4.1kb and 1.9kb in *Hind*III digests. This was interpreted as the presence of a subset of variant 6.0kb units amongst the 5.7kb units. Figure 1.2 illustrates the structure of the Y chromosome alphoid DNA.

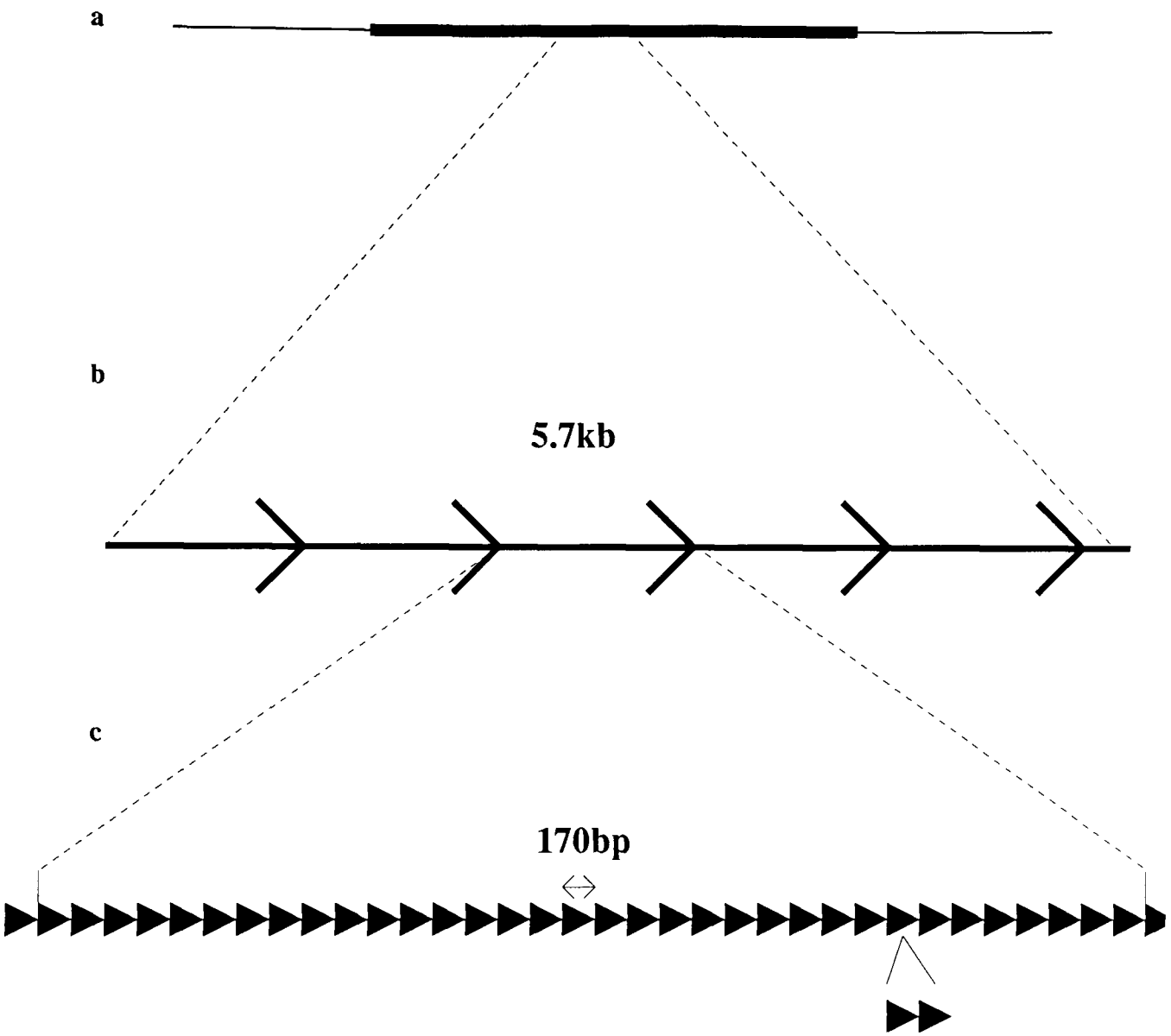


Figure12Structure of the alphoid DNA on the human Y chromosome

The thick line at the top of the figure in a, represents the single block of alphoid DNA on the human Y chromosome and has been found to range in size from 245kb to 1050kb. The block consists of tandemly repeated units most of which are 5.7kb long, as shown by the large arrows in b. Some units are 6.0kb long. The 5.7kb unit is made up of tandemly repeated 170bp subunits as shown in c. The 6.0kb units have two additional 170bp subunits as indicated.

Long range restriction mapping with pulsed field gel electrophoresis (PFGE) using double digests of enzymes which do not cut the units, demonstrated that the units were organised into a single block of alphoid DNA, which was a different size for the two different Y chromosomes. Restriction sites were mapped for the two blocks and it was noted that *Ava*II cuts the Y block at two clusters of sites.

In summary, the Y chromosome alphoid DNA has a higher order repeat structure of 5.7kb units with a few variant 6.0kb units in some chromosomes. Further, more detailed restriction mapping of one of these blocks, the 853 Y alphoid DNA, using double digests showed the arrangement of a much larger group of enzymes at the boundary of the alphoid DNA [Tyler-Smith 1987]. It is noteworthy that the right hand flanking region to the alphoid DNA did not have the restriction enzyme distribution expected for typical human DNA. The closest *Apa*I site was 400kb away and *Kpn*I 200kb away; these should cut every 6.4kb and 8.5kb respectively. This was thought to be a region of simple sequence DNA. The cluster of sites just outside the alphoid DNA were hypothesized to indicate a small stretch of more typical DNA or simple sequence DNA which had these sites frequently. These clusters of sites have since been hypothesized to be present in diverged alphoid DNA units at the edge of the block (Tyler-Smith, personal communication). The left hand flanking region of the alphoid DNA appeared to be typical human DNA with the expected frequency of restriction sites and was postulated to be the boundary between heterochromatin and euchromatin [Tyler-Smith 1987]. Detailed map building is an important step in the analysis of the structure of alphoid DNA, as a prerequisite to the identification of crucial sequences involved in centromere function.

From the comparison of just a few human chromosome alphoid DNAs, it is clear that they cross-hybridise and all have a fundamental 171–173bp repeating subunit structure. However, the alphoid DNA blocks are organised in a chromosome specific way as defined by restriction enzyme analysis of periodicities that define a higher order repeat unit structure. These units probably represent the amplification units which were involved in the expansion of tandem arrays in evolution by unequal crossover. The fact that the fundamental sub-unit structure is conserved at all human centromeres, suggests that it could play a role in correct centromere function, although it is a popular notion that satellite DNAs, such as alphoid DNA play strictly a structural role or are functionless. The fact that there has been divergence of sequence between subunits may suggest that only a part of the sequence is necessary to any role. Nevertheless, it is present in a chromosome specific fashion in all chromosomes and it is important to find out the precise molecular structure of such a sequence in order to ask questions about its possible function.

1.6.5 Possible functions of tandemly repeated DNA sequences

It has been proposed that satellite DNA could play a role in maintenance of chromosome structure, folding, partitioning of chromosomes in the interphase nuclei, homologue recognition [Devilee *et al.* 1986], chromosome pairing and segregation, recombination, and chromosome rearrangements [John and Miklos 1979]. It has been extremely difficult to show conclusive evidence for any of these functions. One series of experiments with *Drosophila* indicated that fitness of *Drosophila* could be associated with the presence or absence of a satellite DNA sequence. The results indicated that deletion of the satellite DNA sequence decreased the fitness of the fly [Wu *et al.* 1989]. Satellite DNA, because of the difficulties in assigning its function, has been labelled selfish DNA. Selfish DNA has two main properties. Firstly it arises when a DNA sequence spreads by forming additional copies of itself within the genome. Secondly it makes no specific contribution to the phenotype [Orgel and Crick 1980]. When a specific function cannot be assigned to repetitive DNA through lack of experimental evidence, it is often given an evolutionary explanation for its existence. For example, facilitating genetic rearrangements which increase evolutionary versatility, or acting as a repository from which new functional sequences can be recruited, or a yet to be eliminated by-product of past chromosomal rearrangements of evolutionary significance [Doolittle and Sapienza 1980].

It is clear that the evidence for specific function assignment to repetitive DNA sequences is weak, and given the abundance of these sequences located throughout the genome, and their specific location at important sites such as centromeres, more investigations are necessary to account for their presence.

1.6.6 Generation and evolution of tandemly repeated sequences

There are a number of theories which attempt to explain the mechanism of the generation, maintenance and spread of repeated sequences in tandem arrays. Perhaps one of the most popular is unequal sister chromatid exchange proposed by [Smith 1976]. The unequal crossovers must occur in the germ line to be evolutionary significant, either in meiosis or at one of the germline mitoses. A crossover can occur between repeats in register producing no change in length, or out of register which will result in one molecule being longer and one shorter. Accumulation of duplications leads to the generation of tandem repeats, and due to their homology, crossovers will occur more frequently in these regions expanding the array. Integral numbers of repeat units will be preferentially deleted or duplicated thus the fundamental repeat pattern does not change. [Smith 1973, Smith 1976]. Mouse satellite DNA had an *EcoR*II repeat unit length of 120bp. However the 240bp dimer was the predominant product of an *EcoR*II digest, [Southern 1975b] lending support to the idea that there has been unequal crossing over in mouse satellite DNA. Highly

repeated DNAs may have evolved by a process of mutation, causing sequence divergence of short repeating elements followed by amplification of the regions generated by mutation [Southern 1970]. Another theory that has been proposed for the evolution of one repeat into another is the rolling helix model [Thomas 1973].

1.6.7 Mapping

Physical analysis of the human genome has been limited by the small size of DNA molecule that it was possible to separate on a conventional gel. The largest useful separation was about 50kb. At the other end of the scale, chromosome maps could be made using the light microscope. These were limited in humans by the visible banding patterns to, at best, 5Mb [Smith and Cantor 1987]. A method has been developed to span this resolution gap and has facilitated physical mapping of the human genome. The pulsed field gel electrophoresis (PFGE) technique was originally developed by [Schwartz and Cantor 1984] and modified by [Carle and Olson 1985]. This technique drastically increased the maximum size of DNA separation possible in an agarose gel. Subsequently, more modifications have been made to the PFG apparatus and one of these, the Waltzer [Southern *et al.* 1987], has been used extensively in this study. The Waltzer is capable of separating up to the largest *S.pombe* chromosome (5.7Mb, [Fan *et al.* 1988]) from larger DNA fragments. This increase in size of fragment separation has made possible the construction of precise long range restriction maps of regions of the genome, and hence allowed detailed analysis of genomic organisation. Determination of the long range structure of the alphoid DNA block at the centromere of the human Y chromosome is an important part of the analysis of the centromere region of a mammalian chromosome.

1.7 Variant chromosomes

1.7.1 Naturally occurring mutant Y chromosomes

The naturally occurring mutant Y chromosomes used in these experiments included individuals with monocentric or dicentric isochromosomes of Yp or Yq; translocations of parts of the Y chromosome onto the X or autosomes, (often seen as XX males or XO males carrying a portion of the Y chromosome elsewhere in the genome); Y chromosomes present as tiny chromosomal fragments and ring chromosomes.

The chromosomes could be used for two purposes. Firstly, for localization of probes by deletion or interval mapping. The presence or absence of Y alphoid DNA was determined for each chromosome. The abnormal Y chromosomes were also used in deletion mapping to localise probes and define intervals along the length of the Y by others. The second use for abnormal Y chromosomes was for the analysis of mutant suppressed centromeres. For example, one Y chromosome portion including a centromere had been translocated onto an autosome. It had been suggested that

the Y centromere was inactive and the autosome centromere active. Analysis of the structure of this type of Y centromere could be important in establishing the relationship between centromere inactivity and the structure of centromere-linked sequences.

It would have been advantageous to have studied many more abnormal Y chromosomes in this way. However, there was a great reluctance on the part of many workers to share cell lines with abnormal Y chromosomes, even after publication.

1.7.2 Chromosome mediated gene transfer and the production of Y chromosome fragments

Chromosome mediated gene transfer (CMGT) can be used to isolate fragments of human chromosomes in human-rodent cell hybrids. Fragments generated in this way provide an alternative to naturally occurring mutant chromosomes. They can be used to determine the arrangement of sequences, in this case on the Y, and provide a source of artificially generated abnormal Y chromosome fragments from 50kb long to a whole chromosome. The stability of fragments does vary however, and an increase in stability was associated with the presence of a centromere [Pritchard and Goodfellow 1987].

CMGT involved the transfection of fragments of human metaphase chromosomes into rodent cells. The size of the transfected chromosomal fragments as a result of CMGT do vary as mentioned above, but have often been found to be in the range of 1–50Mb [Porteous 1987]. However, even though substantial lengths of DNA were thought at the beginning of this project to be transferred intact, it appeared subsequently that molecular rearrangements do take place. *In situ* hybridisations provide evidence for extensive rearrangements of sequence in c-Harvey-ras oncogene transformants [Gosden and Porteous 1987]. Translocation events and duplications were found, so although it was thought that contiguous lengths of DNA could be transferred intact, they were in fact frequently accompanied by intrachromosomal rearrangements. Despite this, using caution, CMGT has been used to map genes with respect to one another [Porteous 1987]. Many CMGT transfectants were found to be unstable. [Klobutcher and Ruddle 1979] suggested that the stabilization process of chromosomal fragments in transfected cells may be viewed as the acquisition of host centromeric activity (integration) for transfected fragments which themselves do not possess a functional centromere.

A modified version of CMGT involved the cotransformation of metaphase chromosomes with a selectable DNA plasmid which enriched for clones containing chromosomal fragments and increased the transfer frequency
[Pritchard and Goodfellow 1987]. For the Y chromosome transfectants used in experiments here, cotransformations were performed by [Pritchard and Goodfellow 1986] using the neo gene which conferred resistance to G418 into mouse LMtk$^-$ cells. The plasmid and transfecting human DNA associates during the transformation process

and preselection with G418 followed by flourescence activated cell sorting (FACS) of 12E7 positive clones, (the 12E7 antigen is present on mouse cells containing X or Y chromosomes) greatly increased the frequency of isolation of Y containing transfectants. Further secondary transfected derivatives were also made by this technique of selected transfectants. Y derived fragments in transfectants were confirmed with Y-linked probes. A quinacrine stained metaphase spread of one of the clones was presented [Pritchard and Goodfellow 1986] which contained a fluorescent fragment which the authors concluded was a portion of the heterochromatic region of the Y chromosome plus an intact short arm. From these data presented, it appeared that these transfectants indeed contained large autonomous portions of the human Y chromosome transfected into a rodent background, an ideal situation in which to study autonomous fragments of the Y chromosome which appeared to contain human centromeres.

Chapter 2

Materials and Methods

2.1 Tissue culture

Cell growth

Cells were stored in 1ml vials at 2×10^6cells per ml in foetal calf serum and 5% DMSO in liquid nitrogen. Cells were thawed quickly and resuspended in 10ml of medium, spun at 1000rpm for 5 minutes before resuspension in a further 10ml of medium and incubation in a 25cm^2 flask. Lymphoblastoid cells were grown in RPMI 1640 (Imperial). To this was added 10% heat inactivated (56°C for 30 minutes) foetal calf serum (Gibco-BRL), 2mM L-Glutamine, 2 IU/ml penicillin and 2μg/ml streptomycin (Gibco-BRL). Monolayer cells were grown in α-modified DMEM or DMEM (Imperial) with the same supplements as RPMI. Some fibroblast cell lines were grown in F-10 medium (Imperial) plus the same supplements as RPMI. For fibroblast cell lines near the end of their lives or for difficult cell lines in general essential and non-essential amino acids and vitamins (Gibco-BRL) supplemented the medium.

All cells were grown at 37°C in 5% CO_2. To split confluent monolayer cells, the medium was removed and the cells washed briefly in PBS (phosphate buffered saline, (Imperial)). 1ml of trypsin (diluted 1:1 with PBS) (Flow) was added to the flasks to detach the cells from the flask surface. The trypsin was inactivated with medium and the cells diluted to about a third of their original density by the addition of medium to the flasks and transferring to more flasks. Full flasks of lymphoblastoid cells were poured from one flask to another and further medium added to adjust the cells to approximately one half of their density. All tissue culture is performed under sterile conditions.

Some of the transfected cell lines required selection to maintain chromosomal fragments. Cells cotransfected with a neo plasmid were grown in 200μg/ml G418 (Geneticin (Gibco)). Cells cotransfected with the tk$^-$ plasmid were grown in HAT medium (100μM Hypoxanthine, 1μM Aminopterin and 10μM Thymidine (Sigma)).

Table 2.1 describes the two cell lines used to build detailed molecular maps of

Cell Line	Description	Selection	Reference
OXEN (L)	49,XYYYY		[Sirota *et al.* 1981]
3E7 (M)	mouse, human Y		[Marcus *et al.* 1976]
853 (M)	hamster, human Y		[Burk *et al.* 1985]
IP2.2 (M)	primary transfectant	HAT	Pritchard and
IP2.6 (M)	primary transfectant	HAT	Goodfellow, 1986
K1P4.2 (M)	primary transfectant		"
JP5.11 (M)	primary transfectant		"
E1P4T10 (M)	primary transfectant	G418	"
E2P3T5 (M)	primary transfectant	G418	"
LMtk$^-$ (M)	Mouse parent cell line		"

Table 2.1: Descriptions of Y cell lines with their selection media where applicable and references. L are lymphoblastoid, M are monolayer cells.

the human Y chromosome alphoid DNA. The cell lines made by chromosome mediated gene transfer containing Y chromosome fragments and their selection media are also shown. In addition, secondary transfectants derived from IP2.6 were used; these are not shown in this table 2.1. All the transfectants were kind gifts from Catrin Pritchard. Most of the normal Y chromosomes involved in the determination of variation of alphoid DNA were obtained from whole blood. Exceptions were OXEN, 3E7, 853 and a few of the samples were obtained from normal male transformed lymphoblastoid cell lines. Lymphoblastoid cell lines from a family of normal individuals (father, mother and 10 sons) were used for determination of the meiotic stability of the Y alphoid DNA. These cells were kind gifts from Dr. Ian Craig.

Table 2.2 lists the abnormal human Y chromosomes used in this study and their sources and references.

Cell harvesting

Monolayer cells were washed with PBS then detached from flasks with trypsin. The trypsin was inactivated with the addition of 10ml of medium. The cells were then poured into 50ml sterile falcon tubes and centrifuged at 1000rpm for 5 minutes. The supernatants were aspirated, the pellet resuspended in 50ml of PBS and a sample taken for a cell count. The rest of the cells were centrifuged and resuspended in the appropriate volume of PBS according to the concentration of cells required. Lymphoblastoid cells were harvested in the same way except that it was not necessary to trypsinise them.

Table 2.2Descriptions and sources of abnormal Y cell lines.

Cell Line	Description*	Source	Reference
870740 (L)	45,X/46,X dicentric iso Yp	PHLS, Porton Down	-
863829 (L)	45,X/46,X dicentric iso Yp	PHLS, Porton Down	-
WSM 184 (F)	70% 45,XO, 25% 46,X, dicentric Yp, 5% 46,XY	A.Chandley	Chandley *et al*, 1986
842968 (L)	46,X monocentric iso Yp	M.A.Ferguson-Smith	Affara *et al*, 1986
861748 (L)	48,XX dic iso Yq,dic iso Yq	M.A.Ferguson-Smith	Affara *et al*, 1986
870377 (L)	46,X iso Yq	PHLS, Porton Down	-
GM06967 (L)	45,X/46,X dicentric Y (p11)/47,X dic Y dic Y	NIGMS cell repository	-
DL-278280R (F)	46,X dicentric iso Y(q11)	M.Fraccaro	Fraccaro, unpublished
GM02668 (F)	45,X/46,X del (Y) (pter >q11.2:)	NIGMS cell repository	-
GM03774 (F)	45,X+fragment/46,X dic iso Y (qter>cen>qter)	NIGMS cell repository	-
863712 (L)	45,X/46,X dicentric Y	PHLS, Porton Down	-
862492 (L)	abnormal Y	PHLS, Porton Down	-
1491/76R (F)	45,X t(Y;18)	M.Fraccaro	Maserati *et al*, 1986
CHoP (L)	45,X t(Y;18)	M.Münke	Münke *et al*, 1988
JOW (L)	45,X/46,X ring Y	A.Chandley	Chandley, unpublished
870374 (L)	45,X/46,X ring Y	PHLS, Porton Down	-
862751 (L)	45,X/46,X ring Y	PHLS, Porton Down	-
862752 (L)	45,X/46,X dicentric Yp	PHLS, Porton Down	-
GM02103 (F)	46,X,t(X;Y)(Xpter> Xq11::Yq11>Yqter)	NIGMS cell repository	Borgaonkar *et al*, 1974
GM02730 (F)	45,X/46,X del (Y) (pter>q11:)	NIGMS cell repository	-
GM03595 (F)	45,X/46,X del (Y) (pter>q11:)	NIGMS cell repository	-
GM02469 (F)	46,X t(X,Y) (Xpter> Xq22::Yq11>Yqter)	NIGMS cell repository	Riddell *et al*, 1986
GM00118 (F)	46,XX,-15,+der15,t(Y;15) (15qter>15p1::Yq11>Yqter)	NIGMS cell repository	Hahnemann *et al*, 1975
GM07970 (F)	47,XX,+der(9),t(Y;9) (9pter>9q13::Yq12)	NIGMS cell repository	-
RW (L)	45,X/46,X +fragment Y	M.A.Ferguson-Smith	Affara *et al*, 1987

* The descriptions of the chromosomes vary in detail according to the information supplied with the cell lines. (L) indicates lymphoblastiod cell line. (F) indicates fibroblast cell line.
The cells from PHLS Porton Down were deposited by M.A.Ferguson-Smith.

In this table, each abnormal Y cell line is described according to the information supplied with the cells from their source, which in some cases is comprehensive and in others, rather vague. A reference is given for cell lines where avaliable. Cells from PHLS and NIGMS were purchased and those from Chandley, Ferguson-Smith, Fraccaro and Münke were kind gifts.

2.1.1 Making plugs

Making plugs from tissue culture cells

Plugs were made according to [Anand 1986]. Cells were harvested as described above. The cells were resuspended to a concentration of 2×10^7 per ml and put to warm briefly at 37°C in a water bath. Meanwhile 1% low gelling temperature (LGT) agarose (Sea Plaque from FMC BioProducts) dissolved in PBS was melted and left to equilibrate to 37°C in a water bath. Equal volumes of cell suspension and 1% agarose were mixed and dispensed quickly into a plug mould. The plug mould was placed on ice until the plugs had set. The final concentration of cells in the plugs was 1×10^7 cells per ml. Individual whole plugs contained $\sim120\mu$l of cell suspension ($\sim1.2\times10^6$ cells). If the DNA content of a diploid cell is 6pg and the cells are diploid, each plug will contain $\sim7.2\mu$g of DNA. The plugs were removed gently from the mould and incubated in NDS (0.5M EDTA, 10mM Tris-HCl, 1% lauroyl sarcosine, pH 9.5) plus ~1mg/ml pronase (Boehringer-Mannheim) at 50°C overnight. The solution was changed and left for a further 48 hours and the plugs carefully rinsed and stored in NDS at 4°C.

Making plugs from blood

50ml of blood was collected and dispensed into 4, 50ml sterile falcon tubes containing 10μl of 50mg/ml of heparin. Three volumes of erythrocyte lysis buffer (155mM NH_4Cl, 10mM $KHCO_3$ and 0.1mM EDTA at pH 7.4) was added to the blood. The mixture was left on ice for 15 minutes and inverted every 5 minutes. The cell suspension was spun at 1000rpm for 10 minutes, the supernatant removed and the pellets pooled. The cells were washed again in lysis buffer, centrifuged and resuspended in 1ml of PBS. A cell count was done and the cells diluted to 2×10^7 to make agarose plugs to a final concentration of 1×10^7 cells per ml.

2.1.2 Making chromosome spreads and C-banding

Preparation of phytohaemagglutinin (PHA) stimulated lymphocytes

Chromosomes were prepared for C-banding in two ways. Firstly 0.4ml of fresh whole blood was added to 5ml of M^{c}Coys medium (Imperial) with PHA (Sigma). The M^{c}Coys medium was supplemented with 2.5μg/ml of PHA, 2.5μg/ml heparin, 20% FCS, 200mM L-glutamine, 100U/ml penicillin and 0.1mg/ml streptomycin. The cultures were left at 37°C for 72 hours. The second set of chromosomes were prepared in a similar way except that a methotrexate (MTX) block was added to the cultures during part of the incubation. This was to prepare chromosomes that were elongated to facilitate analysis of C-bands along the length of the chromosome. 10^{-7}M MTX was added to the M^{c}Coys medium and incubated for 17 hours. The blood culture was then washed twice in medium and resuspended in medium plus

10^{-5}M thymidine and incubated for 5 hours. For non-MTX cultures, 0.1μg/ml colcemid (Gibco) was added to the cultures 90 minutes before harvesting for MTX cultures, 0.1μg/ml of colcemid was added for the final 10 minutes. The cells were centrifuged at 1000rpm for 5 minutes, the supernatant discarded and the pellet resuspended in warm 0.56% KCl (=0.075M). The hypotonic cell suspension was left at 37°C for 10 minutes. The cells were then centrifuged, the supernatant discarded and the cells resuspended in freshly prepared fix (1 part glacial acetic acid to 3 parts methanol). The fix was added dropwise very slowly to start with up to 7ml. This was left at 4°C for 20 minutes. The pellet was centrifuged and washed in fix 3 times. After washing, 2–4 drops of suspension were dropped from about 50cm onto a chilled glass slide. The quality of the chromosomes were checked under the microscope. The preparations were stored at 4°C.

C-banding

The prepared chromosomes were resuspended in fresh fix and slides were made on the same day as C-banding. The slides were treated with 0.2M HCl at room temperature for 30 minutes then rinsed twice in distilled water. The slides were treated with pre-warmed, freshly prepared filtered 0.035M $Ba(OH)_2$ at 37°C for 10 minutes. The slides were rinsed in distilled water 3 times then incubated in 2×SSC at 65°C for 2 hours. The slides were rinsed in distilled water and stained in 5% Giemsa (BDH) for 1–2 hours, then rinsed in water and dried.

The stained slides were observed at a magnification of 1000× using an Olympus system microscope model BHS. Individual metaphase spreads were photographed using 35mm Kodak technical pan film. Hundreds of spreads were scanned for each of 10 individuals, 10 of which were photographed per individual. The photographs were enlarged and printed and the sizes of the long arm C-bands of the human Y chromosome was measured. The visibility or extent of staining of the central C-band was judged by two independent individuals in a blind experiment.

2.2 Restriction enzyme digests in plugs

Each plug was cut into 3 equal sized pieces and one of these was used per digest (about 2.4μg of DNA). Each one third of a plug was placed in an eppendorf tube and soaked overnight in TE (10mM Tris, 1mM EDTA, pH 7.6 with 0.1mM PMSF to inactivate the residual pronase activity) at 4°C. This solution was changed and the plugs soaked for 2 hours with 4 changes of TE on ice. The plugs were then soaked in 500μl of restriction enzyme buffer on ice for one hour. The choice of buffer was dependent on the restriction enzyme and the recipes are shown in table 2.3. [Maniatis *et al.* 1982].

The solution was replaced by 50μl of restriction buffer and between ~10U-50U of enzyme (depending on how well the enzyme digested DNA in plugs) plus

Buffer	NaCl	Tris-HCl pH 7.5	$MgCl_2$
10× High salt buffer	1000mM	500mM	100mM
10× Medium salt buffer	500mM	100mM	100mM
10× Low salt buffer	-	100mM	100mM
10× *HpaI* buffer	200mM	100mM	100mM
10× *EcoO109*I buffer	-	400mM (pH 8.0)	100mM

Table 2.3: Recipes for restriction enzyme buffers.

1mg/ml gelatin for 4 hours. The restriction enzymes were obtained from Boehringer-Mannheim, Northumbria Biologicals Ltd, Amersham and New England Biolabs. Salt concentrations were chosen from table 2.3. Temperatures were chosen according to the manufacturers recommendations, except that 50°C was the maximum temperature for agarose plugs.

Care was taken throughout the soaking and digestion of plugs to avoid DNA degradation by non-specific nucleases.

2.3 Markers for pulsed field gels

λ oligomers were made as described in [Anand 1986]. λ cI857Sam7 DNA was prepared from virions. Oligomer formation was carried out by incubating the DNA at 200μg/ml in 2×SSC (1× SSC is 150mM NaCl, 15mM sodium citrate pH 7.0), 3%Ficoll and orange-G, at 37°C for 30 minutes followed by room temperature for 18–24 hours. Oligomers were stored at 4°C for several months and 5μl (1μg) per track was used.

Yeast DNA markers were prepared according to [Anand 1986]. *S.cerevisiae* X2180-1B chromosomal DNA markers were prepared as follows. Cells were streaked onto a YPD-agar plate (1% yeast extract, 2% bactopeptone and 2% glucose plus 2% agar) and grown at 30°C. One colony was picked and grown in 50ml of YPD (as above without the agar) overnight at 30°C. Two 2L flasks containing 500ml of YPD each were inoculated with 5–10ml of the overnight culture and grown to an O.D. of 0.36 at 600nm giving a total of 4×10^9 cells in the 1L of culture. The cells were harvested by centrifugation at 2000rpm for 10 minutes, the pellets were pooled and resuspended in 1.2M sorbitol, 20mM EDTA and 14mM β-mercaptoethanol on ice. The cells were counted and adjusted to a maximum of 0.5×10^9 cells per ml. 1% LGT agarose dissolved in 1.2M sorbitol, 20mM EDTA, 14mMβ-mercaptoethanol was cooled to 37°C. This was added to the warm yeast cell suspension containing 1mg/ml Zymolase-20T and the mixture allowed to set in a plug mould. The plugs were incubated with the sorbitol solution, 10mM Tris-HCl pH 7.5 and 1mg/ml zymolase for 2 hours. This solution was replaced with sterilized yeast lysis solution (1% lithium dodecyl sulphate, 100mM EDTA and 10mM Tris-HCl pH 8.0) and incubated for 30–60 minutes. An overnight incubation with yeast lysis solution was

performed and the plugs stored at room temperature. [Anand 1986]. This strain of yeast has chromosomes which were assumed to be 245, 280, 360, 450, 600, 690, 770, 800, 840, 940, 970, 1100 and 1370kb (Anand, 1986 and personal communication)

Schizosaccharomyces pombe (972) chromosomal DNA markers were prepared in a similar way except that YPD *S.pombe* was 0.5% yeast extract, 0.5% bactopeptone and 0.5% glucose. Novozyme was added to the cells to a final concentration of 0.5mg/ml and incubated for 1 hour at 37°C. When 50% of cells had spheroplasted, they were mixed with 1%LGT agarose in 1.2M sorbitol, 20mM EDTA and 14mM β-mercaptoethanol and plugs made at 3×10^8 cells per ml. The plugs were incubated in yeast lysis buffer for an hour, then the solution was replaced and incubated overnight [Anand 1986]. The sizes of the *S.pombe* chromosomes have been determined as 5.7Mb, 4.6Mb to 4.7Mb and 3.5Mb [Fan *et al.* 1988].

2.3.1 Size measurement

Measurements were taken from the gel photograph of the distance the size markers had moved from the wells. A standard curve was plotted from these measurements. Measurement of the distances of the radioactive fragments from the wells on the autoradiograph and comparison to the standard curve enabled sizes to be determined with about 5% error.

2.4 Gel electrophoresis

2.4.1 Pulsed field gels

The pulsed field gel apparatus used in these experiments was The Waltzer apparatus as described in [Southern *et al.* 1987]. The Waltzer uses a uniform electric field with a horizontal square gel box and a submerged circular gel rotating at each switching interval. The choice of interval determines the size of separation of DNA and is discussed in [Anand 1986].
A range of pulse times were used during the course of these experiments depending upon the size range of interest. Examples are shown in table 2.4.

1.5% agarose (Sigma low EEO) gels were dissolved in 0.5×TAE (1×TAE is 0.04M Tris-acetate, 2mM EDTA). 0.5% agarose gels were made by pouring a normal 1.5% agarose gel onto the gel plate and cutting away the central portion. The 0.5% gel was poured inside. The wells were made in the gel using a 16 or 18 slot comb with teeth the correct size to allow a third of a plug to slide easily into the well. The gels were loaded using a metal microspatula taking care not to damage the plugs. An appropriate marker (λ concatemers, *S.cerevisiae* or *S.pombe* depending on the pulse time) was loaded alongside the mammalian DNA tracks. The loaded gel wells were sealed with 0.5% LGT agarose. The running buffer was 0.5×TAE. The run time was 30 hours for most gels. For gels used to detect alphoid DNA methylation

Pulse time	Voltage	Temp °C	gel conc.	Size separation
14 s	150	20	1.5%	around 150kb
19 s	150	20	1.5%	up to about 360kb
34 s	150	20	1.5%	around 500kb
44 s	150	20	1.5%	around 600kb
65 s	150	20	1.5%	up to about 800kb
79 s	150	20	1.5%	up to about 900kb
89 s	150	20	1.5%	around 1000kb
99 s*	150	20	1.5%	around 1100kb
10 m	44	4	0.5%	up to 3.5Mb

Table 2.4: Conditions used for PFGE. The table lists the conditions for the apparatus used in most of the experiments described. There was slight variation between gel boxes. *99 s is the longest pulse time useful under these conditions of voltage, temperature and gel concentration.

patterns, the run time was 20 hours to keep around 6.0kb on the gel.

After the run, the gel was stained with ethidium bromide (a few drops of a 10mg/ml stock in 500ml of running buffer) for an hour with shaking and destained in distilled water for 2 hours before photography. The DNA in the gel was visualised under 300nm UV illumination and photographed through an interference filter (λ_{max} = 590nm, half band width = 10mm) on Kodak Technical Pan 4415 film, 5in by 4in.

2.4.2 Conventional gels

Conventional gels were run in vertical glass gel boxes using a 19 well comb large enough to load plugs into. The agarose concentration was 1% in 0.5×TAE and the running buffer was 0.5×TAE. The gels were run overnight at 28V using λ *Hind*III molecular weight markers. The gel was stained for 30 minutes with ethidium bromide and destained briefly. Photography was as described for pulsed field gels.

2.5 DNA Transfer

Gel pre-treatment

DNA transfer was essentially according to [Southern 1975a]. After photography, waltzer gels were depurinated for 20 minutes in 0.125M HCl [Wahl *et al.* 1979]. This was followed by denaturation for 2×30 minutes in 0.5M NaOH, 1.5M NaCl. The gel was then rinsed in distilled water before neutralising in 1.0M Tris (pH 8.0), 1.5M NaCl for 2×30 minutes. Each wash was done with constant shaking. Conventional gels were much thinner than waltzer gels and these were denatured for 30 minutes and neutralised for 30 minutes before blotting.

Component	Volume
Oligolabelling buffer (OLB)	10μl
10mg/ml Bovine serum albumen (BSA)	2μl
DNA in agarose	32.5μl
^{32}P dCTP	4μl
Large fragment of DNA polymerase I	0.5μl

Table 2.5: Oligolabelling mix

Transfer to nitrocellulose

The pre-treated gel was placed on a sheet of Whatman 3mm in contact with a reservoir of 20×SSC. The gel was surrounded by nescofilm to prevent solution from escaping around the gel and causing a short circuit. A sheet of nitrocellulose, (pre-wetted in 2×SSC) was carefully placed on top of the gel followed by a sheet of Whatman 3mm and a packet of tissues. The blot was left overnight and then the filter rinsed briefly in 2×SSC to reduce the salt concentration and remove any gel fragments. The dry filter was baked at 80°C in a vacuum oven.

Transfer to nylon

The transfer procedure for nylon (GeneScreen+) was the same as for nitrocellulose except that the filter was not neutralised before transfer and the transfer solution was 0.4M NaOH. After blotting, the filter was neutralised in 0.2M Tris-HCl pH 7.5 and 2×SSC and then rinsed briefly in 2× SSC and dried.

2.6 Hybridisation

2.6.1 Making the radioactive probe

The Y-derived alphoid probe used in these experiments was pYα1: a 6.0kb Y alphoid DNA unit subcloned into the *EcoRI* site in the vector pEMBL 18$^+$ from cosmid cY77 [Tyler-Smith and Brown 1987]. DNA probes were oligolabelled according to the protocol of [Feinberg and Vogelstein 1983, Feinberg and Vogelstein 1984]. The oligolabelling reaction was set up as indicated in the table 2.5.

The oligolabelling reaction was left for about three hours at room temperature. After incubation, the reaction was made up to 100μl with TE and phenol extracted by mixing with 100μl of phenol followed by 100μl of chloroform. The mixture was spun in a microfuge for 5 minutes and the upper layer transferred to a fresh eppendorf tube. To this was added 400μl of absolute ethanol, 20μl of 10M ammonium acetate and 2μl of 10mg/ml sonicated salmon sperm DNA. This was left on dry ice for 5 minutes to precipitate the DNA. The tube was then spun in a microfuge for 5

minutes, and the pellet dried under vacuum. The pellet was resuspended in 100μl of water, denatured and then added to the hybridisation mix.

2.6.2 The hybridisation solution

For nitrocellulose

Hybridisations were carried out in 6×SET/Denhardts (20×SET is 3M NaCl, 0.4M Tris-HCl pH 7.8 and 20mM EDTA; 20×Denhardts is 0.4% BSA, 0.4% Ficoll and 0.4% polyvinylpyrrolidone), 10% dextran sulphate, 0.1% sodium dodecyl sulphate (SDS), 0.1% sodium pyrophosphate, 0.5mg/ml heparin and 200μg/ml denatured sonicated salmon sperm DNA. 10ml was used per filter. The filters were pre-hybridised in the above solution for at least three hours before addition of the probe. For somatic cell hybrids, the hybridisations with the alphoid probe were carried out at 68°C overnight. For hybridisations to human DNA, the temperature used was 79° overnight to reduce cross hybridisation to alphoid DNA on other chromosomes.

For nylon

The filters were hybridised in 5×Denhardts, 0.5M phosphate buffer pH 7.5 (2.1ml of 2M Na_2HPO_4 plus 0.4ml of 2M NaH_2PO_4 per 10ml of hybridisation solution), 1% SDS, 10% dextran sulphate and 100μg per ml of denatured sonicated salmon sperm DNA. A pre-hybridisation of 3–6 hours preceded the overnight hybridisation at 68 or 79°C.

2.6.3 Washing filters and autoradiography

Filters were washed first in 2×SSC, 1% SDS at 65°C for 30 minutes with shaking. The second wash was in 0.5×SSC at 65°C for 30 minutes. The final wash was in 0.1×SSC at 65°C for 30 minutes. The filters were air-dried and exposed to Fuji Rx X-ray film with a Philips intensifying screen at –70°C.

2.6.4 Determination of intensity of hybridisation from autoradiographs

For dicentric and pseudo-dicentric chromosomes, the number of Y alphoid DNA blocks per Y chromosome was not known. To determine this, the DNA from these cell lines was digested with *EcoR*V, run on a conventional gel (as above) and hybridised with the probes pYα1 and Y190. The intensity of pYα1 hybridisation (which represented the number of Y alphoid DNA blocks per cell) compared to Y190 (which was an internal standard which represented the number of Y chromosomes per cell) was compared. The relative intensity of the two bands in each track

on the autoradiograph were measured with a Molecular Dynamics optical scanner. The Scanner was able to produce a digitised value for the intensity of hybridisation in a given band. For each band scanned, the intensity value was corrected for local background intensity and expressed in arbitrary units.

Chapter 3

Structure of the alphoid DNA on normal Y chromosomes

3.1 Detailed molecular map of the 3E7 Y chromosome alphoid DNA

This chapter describes the structure of the centromere-linked alphoid DNA in normal human Y chromosomes.

Restriction maps have been made of the Y alphoid DNA on two Y chromosomes [Tyler-Smith and Brown 1987] using five enzymes. It was decided to extend the analysis of these regions to create a more detailed view of the organisation of alphoid DNA at the Y centromere.

First, a search was made for additional useful restriction enzymes. Y alphoid DNA cloned into the cosmid cY77 [Tyler-Smith and Brown 1987] was digested with a range of restriction enzymes. Table 3.1 shows which enzymes do not have sites within the alphoid DNA units. These enzymes are useful for mapping the sequences flanking the Y alphoid DNA blocks and have been used here to produce detailed restriction maps of two blocks. The Y alphoid DNA of 853, a human-Y hamster somatic cell hybrid, has been mapped in detail [Tyler-Smith 1987]. At the same time, the Y chromosome alphoid DNA from 3E7 and OXEN were mapped in this study using a selection of restriction enzymes which do not have sites within most of the units.

The digested 3E7 DNA fragments were separated on the Waltzer pulsed field gel electrophoresis (PFGE) apparatus and hybridised with pYα1. Photographs of the autoradiographs of digests which contributed to the maps are shown in figure 3.1. Single digests revealed the fragment sizes, and restriction enzyme sites were placed with respect to one another with double digests. (If the product of a double digest was a smaller fragment than either of the two single digests, then the sites could be placed with respect to one another.)

The restriction map of the 3E7 Y chromosome alphoid DNA constructed from

Restriction enzyme	*Site(s) in alphoid DNA**
*Hpa*I	–
*Sst*I	–
*Bgl*II	–
*Bgl*I	–
*Pvu*II	–
*Acc*I	–
*Ava*II	–
*EcoO109*I	–
*BstX*I	–
*Tth111*I	–
*Sca*I	–
*Bcl*I	–
*BamH*I	–
*BstE*II	–
*Kpn*I	–
*Apa*I	–
*Sfi*I	–
*Nae*I	–
*Nar*I	–
*Sal*I	–
*Eag*I	–
*Cla*I	–
*Pvu*I	–
*Sma*I	–
*Sst*II	–
*BssH*II	–
*Xho*I	–
*Not*I	–
*EcoR*I	+
*EcoR*V	+
*Hind*III	+
*Pst*I	+
*Xba*I	+
*Hae*III	+
*Sty*I	+

Table 3.1: A list of restriction enzymes. – indicates the absence of sites in Y alphoid DNA units. + indicates the presence sites in Y alphoid DNA units. *The restriction enzymes were tested on the cosmid cY77.

the data in figure 3.1 is shown in figure 3.2. The size of the alphoid block was determined from the *Pvu*II single digest. 3E7 was approximately 430kb although the edges of the block were not clearly defined as indicated by the broken edges of the pYα1 homology in figure 3.2. Most of the restriction sites were placed by combined digestion with *Hpa*I. These were *Kpn*I, *BstE*II, *BamH*I, *Bcl*I, *EcoO109*I, *Apa*I, *Sst*I, *Sca*I, *BstX*I, *Pvu*II, *Bgl*I and *Tth111*I (figure 3.1, part a and b). *Ava*II double digests placed *Acc*I, *Bgl*II and *Pvu*II sites (figure 3.1, parts c, d and e). The clustered *Ava*II sites are present in consecutive 6.0kb repeat units. The 6.0kb units constitute a minority of units compared to the 5.7kb units in some but not all of the Y chromosome alphoid DNA blocks. *Ava*II sites have been mapped in cosmid clones [Tyler-Smith and Brown 1987] who showed that the 6.0kb units which contain the extra *Ava*II sites are clustered together. In addition, 3E7 has an *Acc*I cluster of sites within the block. Most of the other restriction sites were close together around the edges of the block. Exceptions included *Apa*I which was approximately 400kb away from the edge of the block; *Kpn*I at 200kb away from the block; next in was *BstE*II about 150kb away followed by *EcoO109*I at 100kb away from the edge of the block (figure 3.2).

Figure 3.1. 3E7 restriction fragments detected by pYα1.
3E7 DNA was digested with restriction enzymes and hybridised with pYα1. The restriction fragments shown here were used to construct a map of the 3E7 Y alphoid DNA (figure 3.2). Each track contains DNA digested with one or two restriction enzymes as indicated. The sizes of the fragments in kb are shown at the side of the tracks.

Restriction enzyme fragment sizes in kb for 3E7.

Restriction enzyme	size in kb
*Hpa*I	540
*Hpa*I *Kpn*I	520
*Kpn*I	750
*Hpa*I *BstE*II	470
*BstE*II	630
*Hpa*I *Bam*HI	475
*Bam*HI	540
*Hpa*I *Bcl*I	465
*Bcl*I	540
*Hpa*I *Eco0109*I	460
*Eco0109*I	570
*Hpa*I *Apa*I	465
*Apa*I	850
*Hpa*I *Sst*I	500
*Sst*I	510
*Hpa*I *Sca*I	480
*Sca*I	515
*Hpa*I *BstX*I	470
*BstX*I	485
*Hpa*I *Pvu*II	450
*Pvu*II	460
*Hpa*I *Bgl*I	475
*Bgl*I	490
*Ava*II	195, 130, 90
*Ava*II *Pvu*II	170, 130, 90
*Ava*II *BstX*I	195, 130, 90
*Ava*II *Bgl*I	195, 130, 90
*Ava*II *Bgl*II	175, 130, 90
*Ava*II *Acc*I	130, 90, 60
*BamH*I	540
*BamH*I *Tth111*I	485
*Tth111*I	540
*Tth111*I *Hpa*I	500
*Hpa*I	540
*Sst*I	510
*Sst*I *Sca*I	470
*Sca*I	515

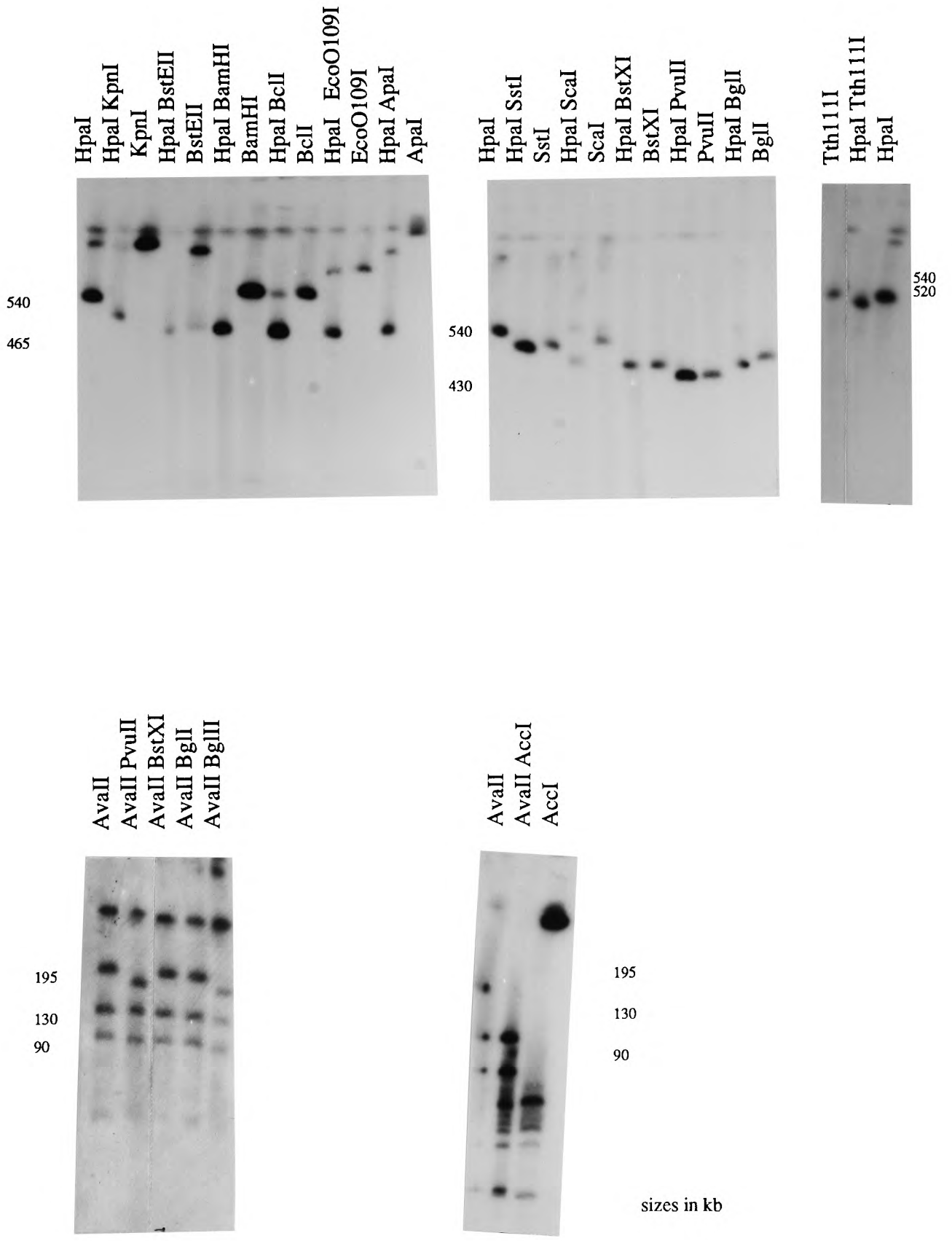
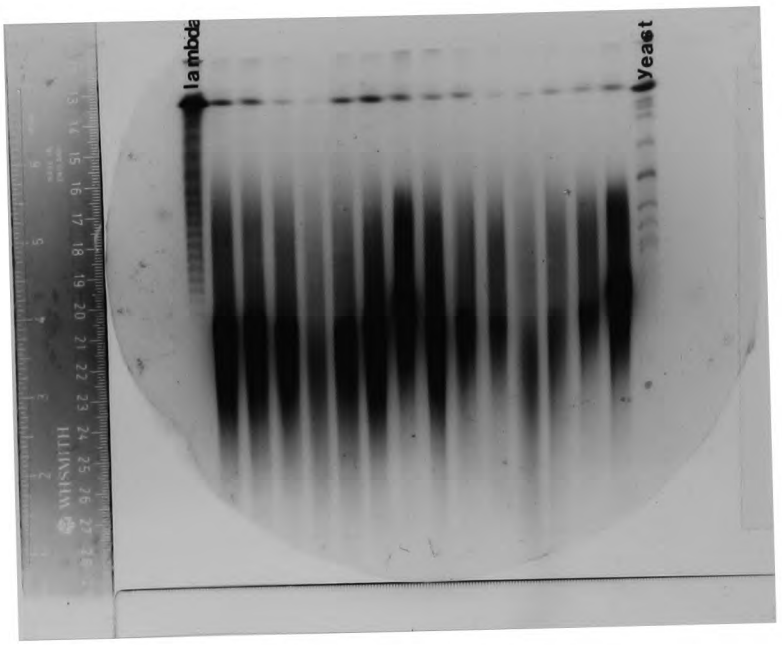

HpaI
HpaI KpnI
KpnI
HpaI BstEII
BstEII
HpaI BamHI
BamHI
HpaI BclI
BclI
HpaI EcoO109I
EcoO109I
HpaI ApaI
ApaI
540
465
HpaI
HpaI SstI
SstI
HpaI ScaI
ScaI
HpaI BstXI
BstXI
HpaI PvuII
PvuII
HpaI BglI
BglI
540
430
Tth111I
HpaI Tth111I
HpaI
540
520
AvaII
AvaII PvuII
AvaII BstXI
AvaII BglI
AvaII BglII
195
130
90
AvaII
AvaII AccI
AccI
195
130
90
sizes in kb

Figure 3.1

Figure 3.2. Restriction site map of the 3E7 Y alphoid DNA.

The horizontal line represents the chromosomal DNA with restriction enzyme sites marked. The two shorter maps beneath show the positions of restriction sites which were not oriented with respect to the longer map and are therefore shown in both possible directions. The *Ava*II sites in the upper map are the same as the two outer most *Ava*II sites on the maps below. The sites with multiple branches indicate restriction enzyme sites which are clustered. The thick line beneath the maps indicates the approximate region of homology to pYα1 and the broken edges show that the ends of the alphoid DNA are not clearly defined. The scale shows the length of the alphoid DNA in kb, which is around 430kb in 3E7.



Gel photograph showing yeast and lambda markers

Fig 3.11

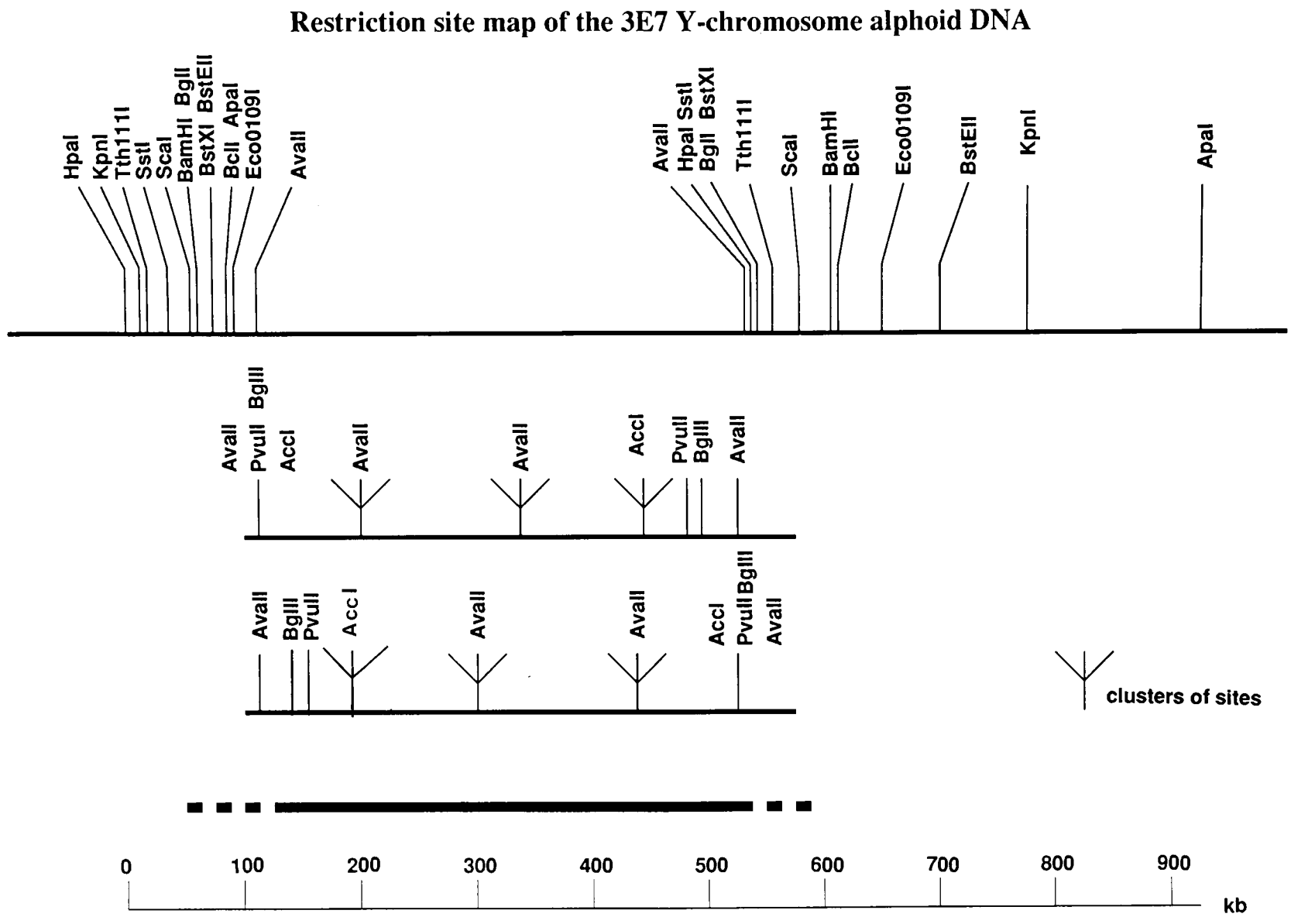

Restriction site map of the 3E7 Y-chromosome alphoid DNA
HpaI
KpnI
Tth111I
SstI
ScaI
BamHI
BglI
BstXI
BstEII
BclI
ApaI
Eco0109I
AvaII
AvaII
HpaI
SstI
BglI
BstXI
Tth111I
ScaI
BamHI
BclI
Eco0109I
BstEII
KpnI
ApaI
AvaII
PvuII
BglII
AccI
AvaII
AvaII
AccI
PvuII
BglII
AvaII
AvaII
BglII
PvuII
AccI
AvaII
AvaII
AccI
PvuII
BglII
AvaII
clusters of sites
0
100
200
300
400
500
600
700
800
900
kb

Figure 3.2

3.2 Detailed molecular map of the Y chromosome alphoid DNA from OXEN

pYα1, although derived from the human Y chromosome, [Tyler-Smith and Brown 1987] also hybridised to alphoid DNA sequences on other chromosomes at 68°C (see figure 3.3, left hand side). The cross-hybridisation produced a complex series of bands making it difficult to identify the Y-specific bands. To observe Y chromosome alphoid DNA in a human background, the temperature of the hybridisation was raised to 79°C. This increased hybridisation stringency reduced the cross-hybridisation to other chromosomes as illustrated on the right hand side of figure 3.3. The Y chromosomes from the OXEN 49,XYYYY lymphoblastoid cell line were mapped in a similar way to the 3E7 chromosome. The restriction enzyme digests which contributed to the sites on the map are shown in figure 3.4. The restriction enzyme sites were again placed using single and double digests, (figure 3.5). *Hpa*I double digests placed *Kpn*I, *BamH*I, *Apa*I and *BstE*II (figure 3.4, part b). The remaining sites for *EcoO109*I, *BstX*I, *Pvu*II, *Bgl*I, *Tth111*I, *Ava*II and *Acc*I were placed with *Sst*I and *Bcl*I double digests (figure 3.4 part c). On the left hand side of the map (figure 3.5), there are two *Sst*I sites which produced a small alphoid fragment of 70kb. The restriction enzyme sites which occurred in this fragment cleaved it into smaller fragments allowing precise mapping in this region. On the right hand side of the map (figure 3.5) there were two *Bcl*I sites. Cleavage of this small (130kb) fragment placed restriction sites precisely on the right hand side of the map.

The size of the alphoid block was 1000kb although where the alphoid sequence ends was not clearly defined. The OXEN Y alphoid DNA block had an *Acc*I site in the middle of the block, and as mentioned above, an extra *Sst*I site on the left hand side and an extra *Bcl*I site on the right hand side of the block. There was also an additional alphoid fragment produced when the DNA was digested with *Ava*II or *EcoO109*I which indicated that these two enzymes had sites within the alphoid block.

The detailed restriction maps of these two Y alphoid DNA blocks and that of 853 [Tyler-Smith 1987], showed variation between the alphoid blocks on different Y chromosomes. The most obvious difference was block size; 3E7 was 430kb; OXEN was 1000kb and 853 was 540kb. In addition there were internal restriction enzyme site polymorphisms. In OXEN and 853, digestion with *EcoO109*I produced two fragments, so there was a site for this enzyme inside the block. 3E7 lacked this internal site. 3E7 and OXEN had internal *Ava*II sites whereas 853 lacked them. It was therefore decided to make a detailed survey of a collection of normal males in order to determine the extent of polymorphism between chromosomes. An assessment of the extent of variation in the population would then allow a comparison to be made between the alphoid DNA of normal centromeres and that on cytogenetically abnormal Y chromosomes.

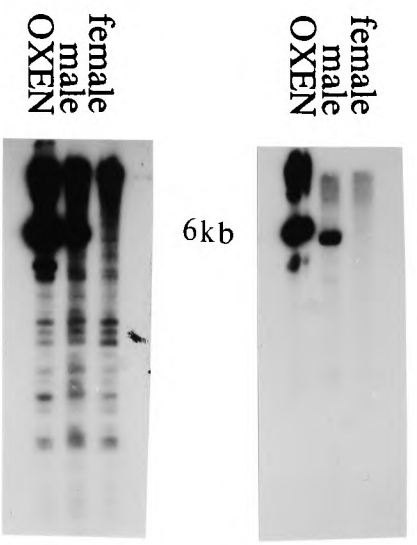


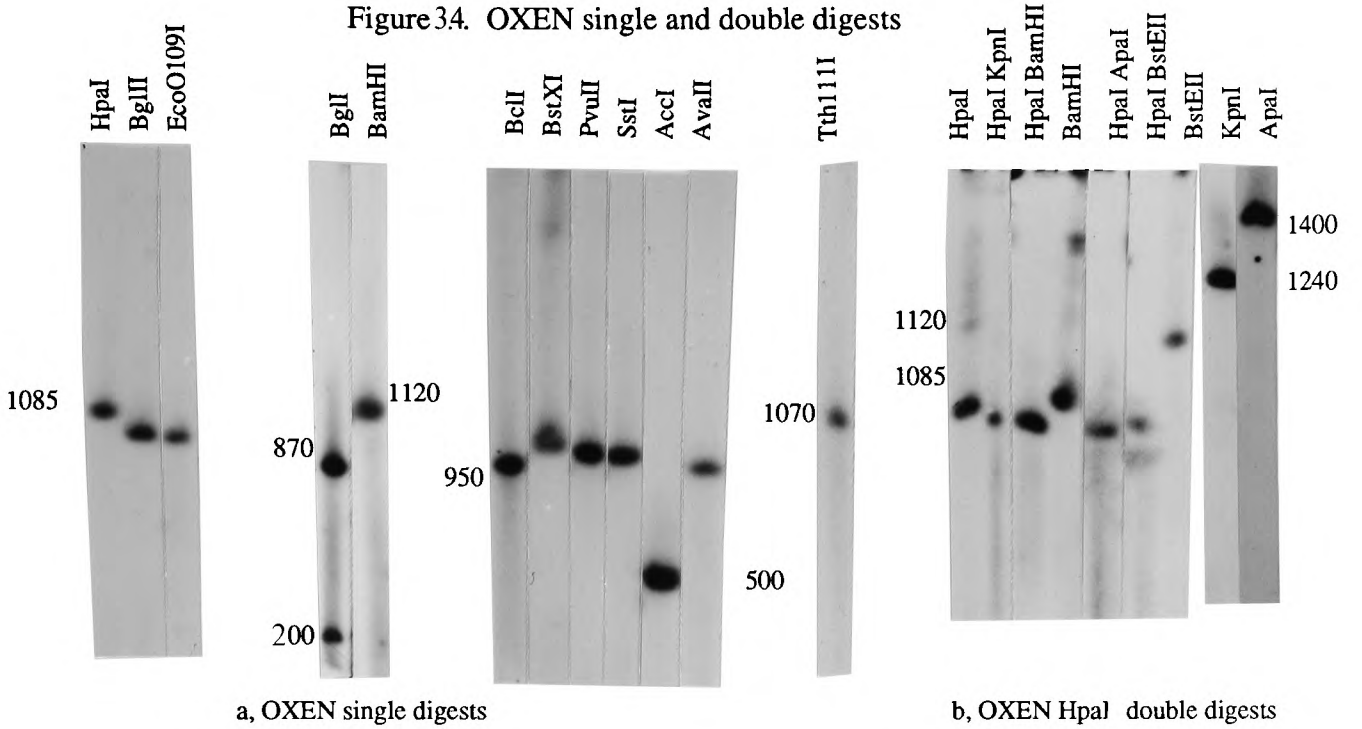
Figure33. Hybridisation of pYalpha1 to OXEN, male and female DNA samples.

The HindIII digests on the left hand side were hybridised at 68oC and those on the right at 79oC.

Figure 3.4. OXEN restriction fragments detected by pYα1

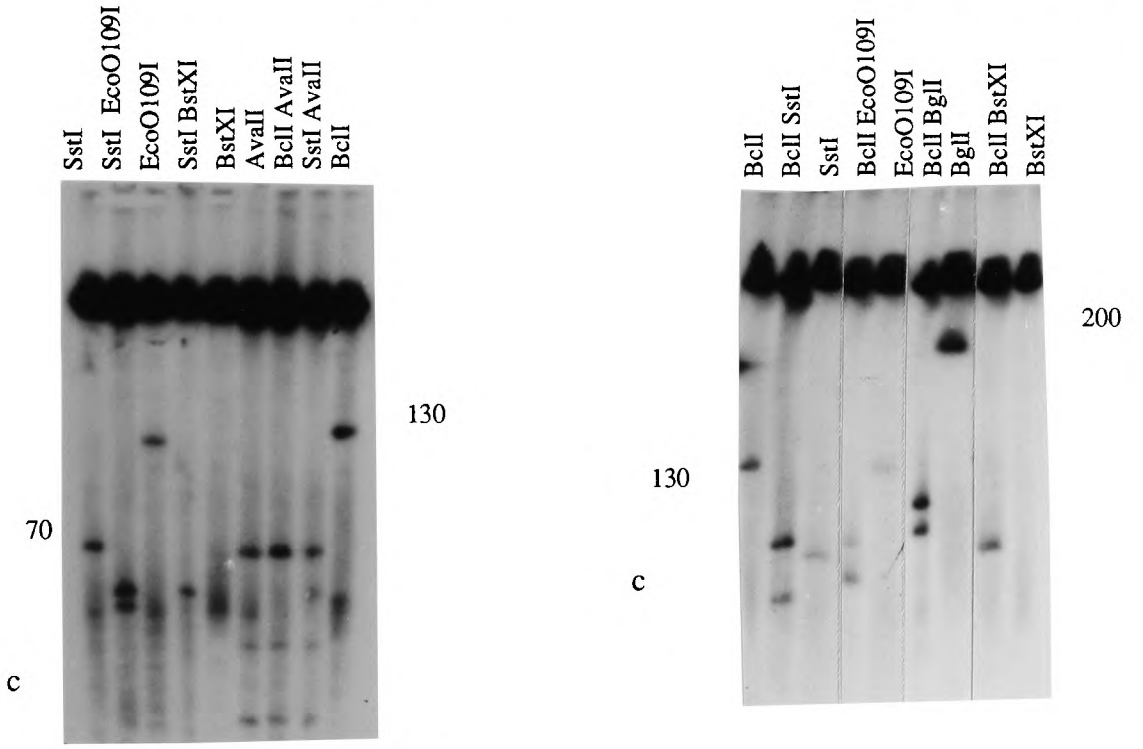
OXEN DNA was digested with restriction enzymes and hybridised with pYα1. The restriction fragments shown here were used to construct a map of the OXEN Y alphoid DNA (figure 3.5). Each track contains DNA digested with one or two restriction enzymes as indicated. The sizes of the fragments in kb are shown at the side of the tracks.

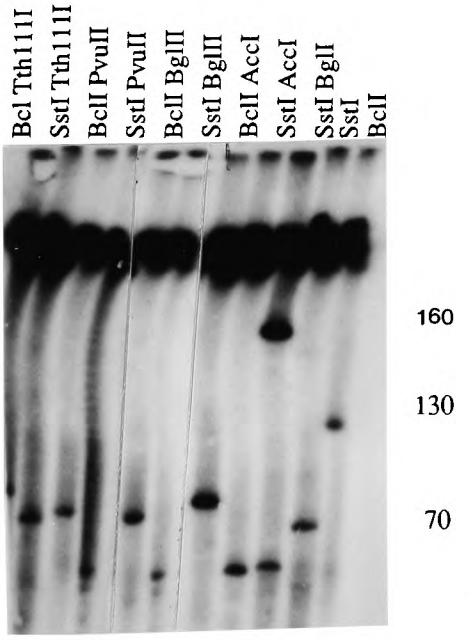
Figure 34. OXEN single and double digests



a, OXEN single digests

b, OXEN HpaI double digests





c, OXEN SstI and BclI double digests

Sizes are in kb

Figure 3.5. Restriction site map of the OXEN Y chromosome alphoid DNA.

The horizontal line represents the chromosomal DNA with restriction enzyme sites marked. The thick line beneath the map indicates the approximate region of homology to pYα1. The broken edges show that the ends of the alphoid DNA are not clearly defined. The scale shows the length of the alphoid DNA in kb, which is around 1000kb in OXEN.

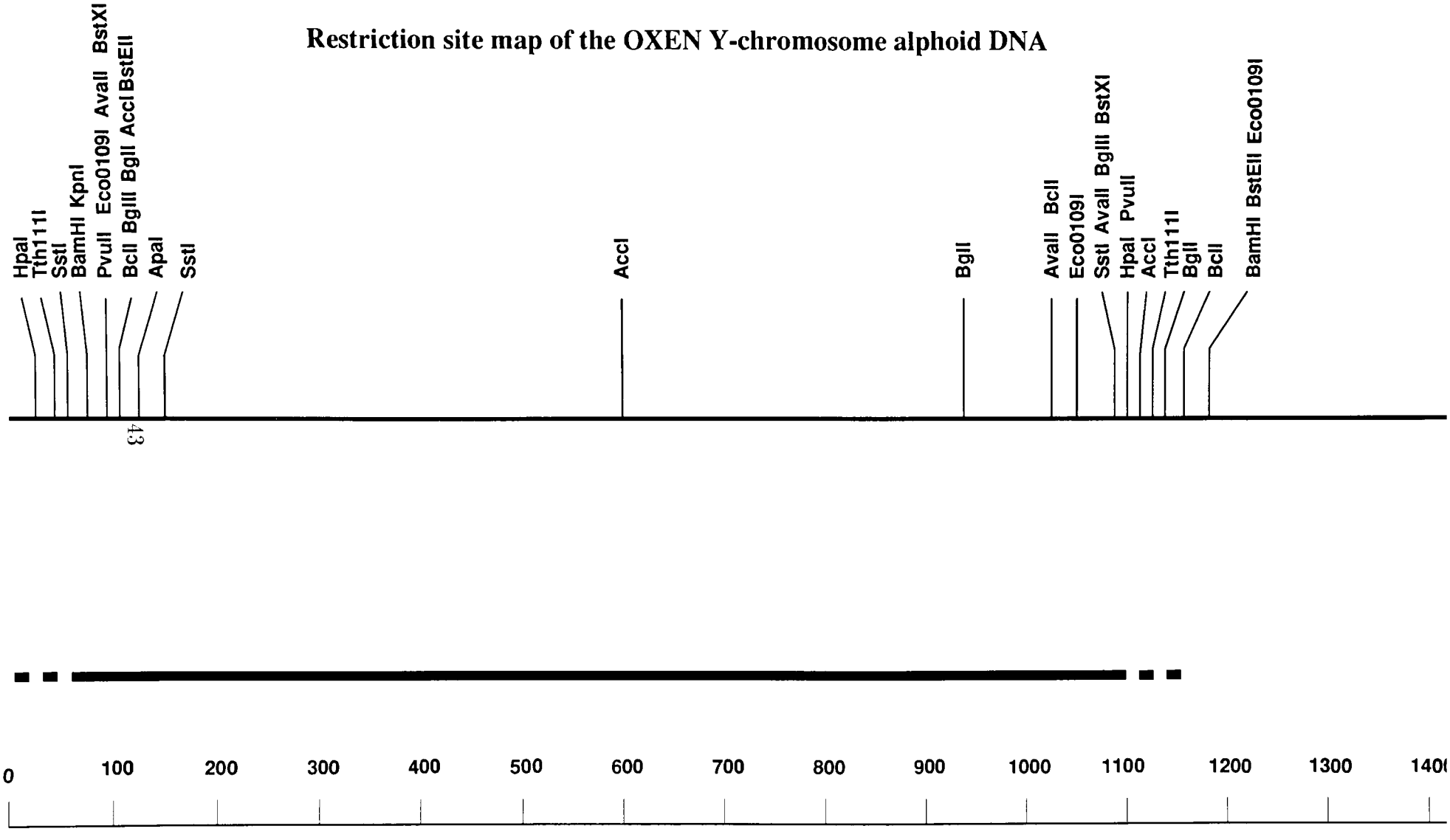

Restriction site map of the OXEN Y-chromosome alphoid DNA
HpaI
Tth111I
SstI
BamHI KpnI
PvuII Eco0109I AvaII BstXI
BclI BglII BglI AccI BstEII
ApaI
SstI
AccI
BglI
AvaII BclI
Eco0109I
SstI AvaII BglII BstXI
HpaI PvuII
AccI
Tth111I
BglI
BclI
BamHI BstEII Eco0109I
43
0
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
140

Figure 3.5

3.3 Y alphoid DNA polymorphisms in 42 normal Y chromosomes

Selection of Y chromosomes and restriction enzymes to detect Y alphoid DNA polymorphisms

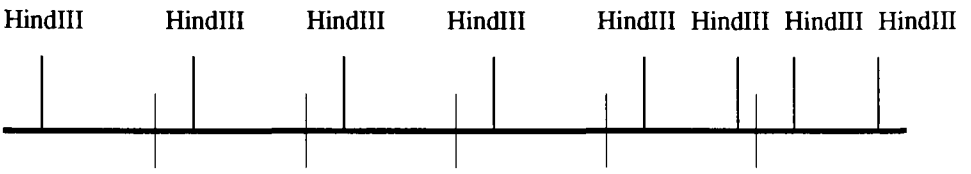
A set of restriction enzymes was selected to survey the Y alphoid DNA polymorphisms in a group of unrelated males. These males were selected from the population according to availability and willingness to give blood. Most of the 42 normal Y chromosomes came from blood samples. A few were from lymphoblastoid cell lines and OXEN, 3E7 and 853 cell lines. The enzymes chosen to detect polymorphisms of alphoid DNA block length were *Bgl*II, *Pvu*II, *Ava*II, *Acc*I, *EcoO109*I, *Bcl*I and *Apa*I. All of these enzymes were expected to detect block size polymorphisms. *Pvu*II and *Bgl*II revealed the block size. *Bcl*I, *EcoO109*I, *Ava*II and *Acc*I were included because they showed additional internal sites in one or more of the three Y chromosomes mapped in detail. *Apa*I was included because it cut approximately 400kb from the block and thus allowed any variation in the 400kb of flanking DNA to be analysed.

*Hind*III has a site in all of the Y alphoid DNA 5.7kb units, so *Hind*III detected a 5.7kb fragment in all individuals. Some blocks showed additional minor 4.1kb and 1.9kb fragments due to an extra site in the 6.0kb units. Figure 3.6 shows the typical unit structure of part of an alphoid block. Structure *a* has some 6.0kb units. This type of structure would produce the extra minor fragments. The second structure illustrated in *b*, has only 5.7kb units as no minor fragments were seen in a *Hind*III digest. *Hind*III polymorphisms were detected on conventional gels and are illustrated in figure 3.3. The first male track has 6.0kb subunits and the second lacks them.

Figure 3.6. The typical unit structure of part of a large and a small Y alphoid DNA block.

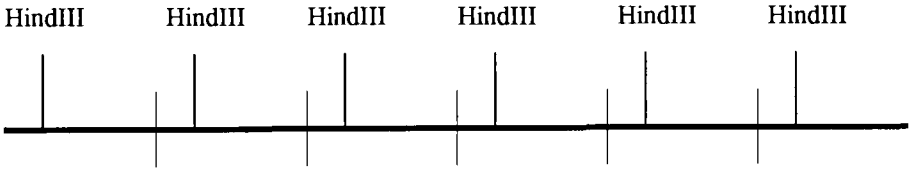
The horizontal line represents a region of six units of Y alphoid DNA. The thin vertical bars indicate the unit structure. In small blocks, as shown in *b*, the 5.7kb units had one *Hind*III site per unit. The *Hind*III sites are represented by thicker vertical bars. A small block therefore had a 5.7kb fragment in a *Hind*III digest. In large blocks, there were mostly 5.7kb units which gave 5.7kb *Hind*III fragments. However there were also some clustered 6.0kb units as shown in case *a*. These had an extra *Hind*III site and additional fragments of 4.1kb and 1.9kb were present in a *Hind*III digest.

a



The position of the HindIII sites in part of a large Y alphoid DNA block

b



The position of the HindIII sites in part of a small Y alphoid DNA block

5.7kb

Figure3.6.The typical unit structure of part of a large and a small Y alphoid DNA block.

Y alphoid DNA polymorphisms

The variation in block size determined by *Pvu*II digestion is illustrated in figure 3.7. The size range for the 42 individuals was from 245kb (male 32) to 1000kb (male 19). This demonstrated the polymorphic nature of Y alphoid DNA. There are variations in the intensity of some of the bands. The very faint bands are as a result of damage to the digested plugs during gel loading.

The fragment sizes for each enzyme for the 42 individuals are summarised in table 3.2. Column 1 lists the individuals; columns 2–8 show the size(s) (in kb) of the alphoid fragment(s) detected after digestion with the enzyme indicated. Column 9 shows the presence of the 4.1kb and 1.9kb alphoid fragments (+) or their absence (–). Column 10 indicates the alphoid allele assignment and column 11 indicates the group to which the chromosome belongs (discussed below).

The *Pvu*II fragment size distribution was plotted on the histogram in figure 3.8. At least 20 different sizes could be distinguished between 42 chromosomes. The distribution of alphoid DNA polymorphisms is clearly not random. The Y alphoid DNA blocks appear to fall into two different groups. The small alphoid DNA group has a block size centred about 300kb and the large group at 900kb. A group number was assigned to each chromosome and listed in column 11 of table 3.2. The *Bgl*II and *Acc*I fragments in most cases were similar in size to the *Pvu*II fragment size. The *Bcl*I fragment was usually approximately 50kb larger than the *Pvu*II size. Thus the internal size polymorphisms detected by *Acc*I and *Bcl*I are not common in the population analysed. The *Apa*I fragment was usually approximately 400kb larger than the *Pvu*II fragment size. The flanking sequences contained in the *Apa*I fragment are not very polymorphic. Block size however, was not the only polymorphic feature associated with each individual. *Ava*II and *EcoO109*I detect both block size differences and frequent point mutations. These polymorphisms along with the *Hind*III polymorphism are represented by symbols inside the squares on the histogram (figure 3.8).

In summary, small alphoid DNA blocks (245–445kb, apart from 3E7), had no internal *Ava*II sites and a *EcoO109*I site approximately 100kb away from the block. These individuals (except 3E7) had no 6.0kb units in their alphoid block and so no extra *Hind*III minor fragments. Large alphoid DNA blocks (650–1050kb) had internal sites for *Ava*II and *EcoO109*I. These individuals had some 6.0kb units which contained extra *Hind*III sites. The *Ava*II and *EcoO109*I polymorphisms are summarised in figure 3.9. All individuals in the small group were of Caucasian or part Caucasian origin. The individuals in the large group were Caucasian or Asian (including Oriental) in origin.

Table 3.2Alphoid DNA restriction enzyme fragment sizes for 42 normal Y chromosomes.

	alphoid									
	BglII	PvuII	AccI	AvaII	EcoO109I	BclI	ApaI	HindIII		
1	2	3	4	5	6	7	8	9	10	11
1	265	275	285	280	380	330	650	-	22	1
2	320	320	330	330	480	380	745	-	21	1
3	310	310	330	310	450	380	695	-	22	1
4	360	350	360	355	430	410	720	-	21	1
5	275	280	275	275	390	355	690	-	22	1
6	840	810	790	680+90	830+175	900	1200	+	11	2
7	330	330	330	330	420	375	720	-	22	1
8	320	320	340	340	420	400	750	-	22	1
9	910	910	950	840+90	950+120	1030	1300	+	7	2
10	315	315	335	335	425	380	680	-	22	1
11	415	415	430	415	500	460	790	-	18	1
12	950	940	990	910+90	950+120	1030	1320	+	6	2
13	415	415	430	415	515	475	840	-	17	1
14	940	920	960	850+100	925+120	1010	1330	+	7	2
15	295	295	315	295	400	350	665	-	22	1
16	970	950	990	920+100	970+120	1060	1340	+	6	2
17	335	335	360	335	425	400	720	-	21	1
18	300	295	305	295	395	350	685	-	22	1
19	1065	1050	1130	1010+90	1070+120	1205	1360	+	1	2
20	270	270	285	285	400	360	690	-	22	1
21	310	310	325	325	425	370	700	-	23	1
22	680	660	620	700+80	640+150	660	1060	+	13	2
23	950	940	950	890+80	910+120	1000	1320	+	6	2
24	350	350	370	360	470	410	740	-	22	1
25	1020	1020	1080	950+90	1020+120	1100	1340	+	2	2
26	335	335	350	345	465	415	725	-	21	1
27	800	800	800	730+70	760+120	820	1200	+	12	2
28	860	860	860	730+70	840+120	900	1180	+	10	2
29	545	545	550	545	510+120	590	920	-	15	-
30	345	345	360	360	465	410	740	-	20	1
31	280	275	290	280	385	335	670	-	22	1
32	245	245	255	250	350	305	640	-	24	1
33	880	870	890	575+175 +60	870+120	940	1300	+	8	2
34	505	505	520	510	490+120	570	900	-	15	-
35	270+115	375	385	375	475	435	760	-	19	1
36	920	920	920	840+100 +50	910+120	960	1260	+	9	2
37	650	650	650	600+70	630+120	725	1030	+	14	2
38	1000	1000	1010	950+70	1000+120	1050	1270	+	5	2
39	650+360	650+360	650+375	650+300 +100	675+350	675+425	1380	+	3	2
OX	970	1000	500+500	930+80	910+120	950+130	1400	+	4	2
3E7	400	430	350+45*	190+140 +100	610	520	890	+	16	-
853	540	540	530	530	530+120	610	930	-	15	-

DNA fragment sizes (in kb) and alphoid haplotypes of 42 normal human Y chromosomes. Column 1 lists the individuals. Columns 2-8 show the size(s) of the alphoid fragment(s) detected after digestion with the enzyme indicated. Column 9 shows the presence of the 4.1kb and 1.9kb alphoid fragments (+) or their absence (-). Column 10 indicates the alphoid allele assignment. Column 11 indicates the group to which the chromosome belongs.

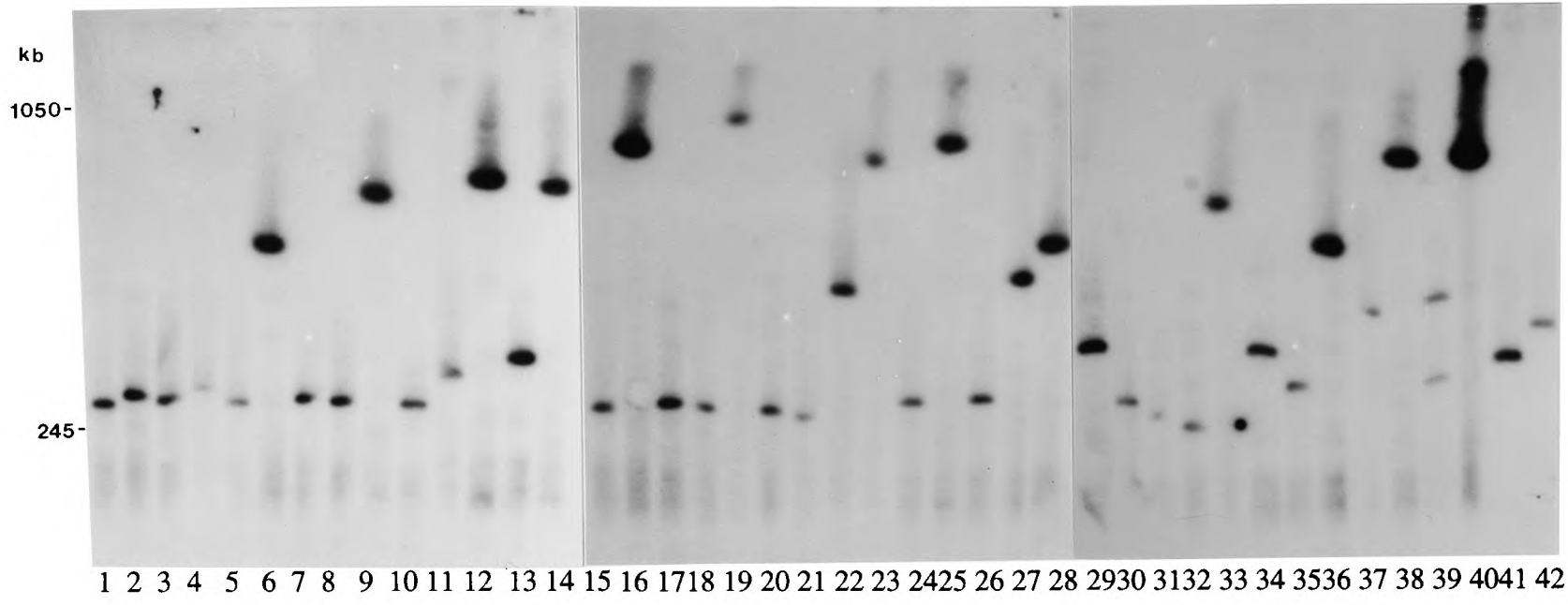


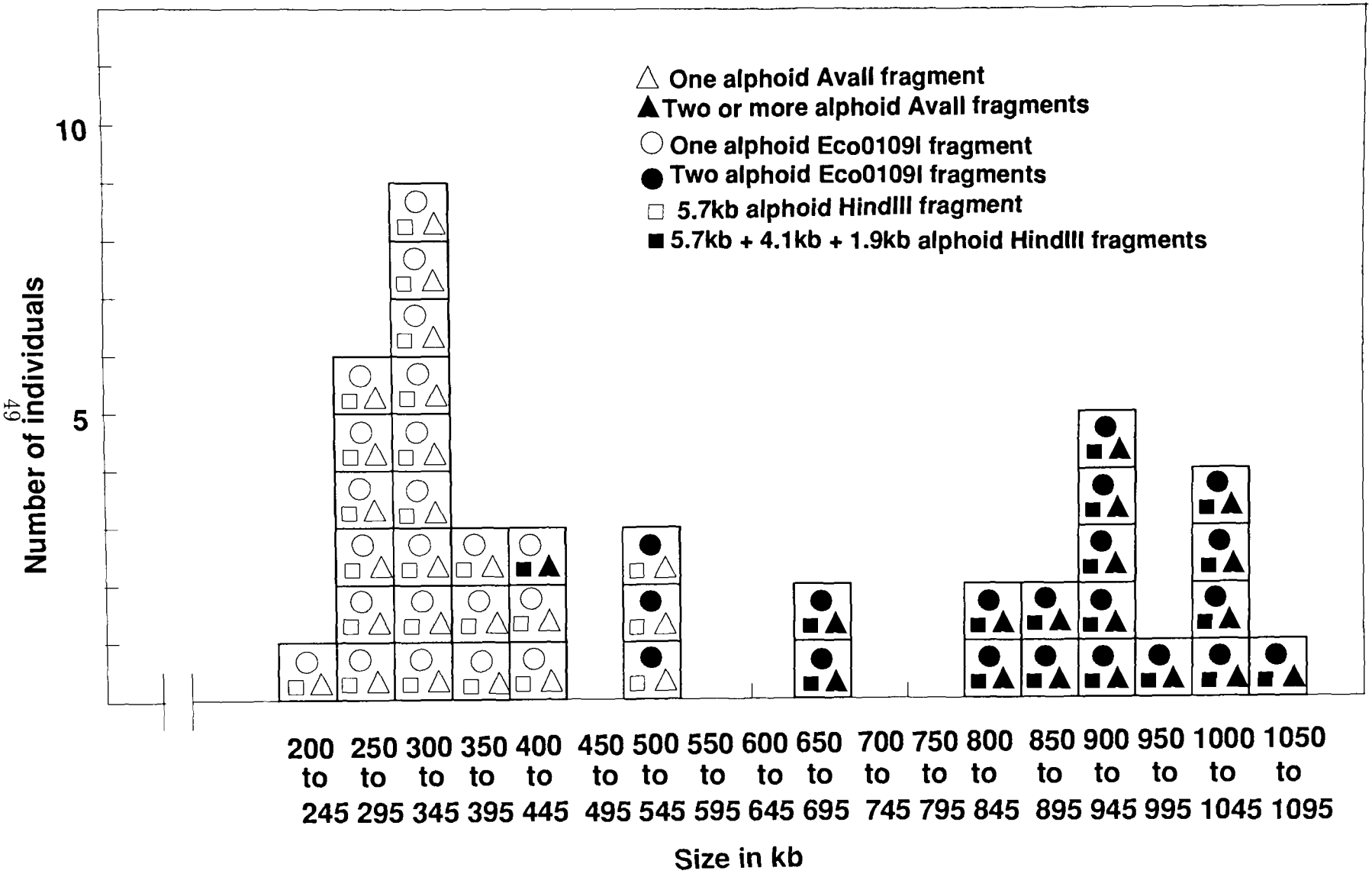
Figure 3.7 Detection of polymorphic Y alphoid DNA fragments.

Tracks 1 to 39 contain DNA from males 1 to 39. Tracks 40 to 42 contain DNA from OXEN, 3E7 and 853 respectively. PvuII digests, pYalpha1 probe.

Figure 3.8. Distribution of alphoid DNA sizes and associated polymorphisms on 42 Y chromosomes.

Each square represents one individual. The size of the *Pvu*II fragment (the sum of the two fragments for individual 39) was used as the measure of block size. Additional polymorphisms found in the individuals are indicated by the symbols inside the square. The Y chromosomes seem to fall into two groups. Most of the small Y alphoid DNA blocks from 245kb to 445kb have one *Ava*II fragment, one *EcoO109*I fragment and 5.7kb alphoid *Hind*III fragments. Large blocks from 650kb to 1095kb are associated with two or more *Ava*II fragments, 5.7kb, 4.1kb and 1.9kb *Hind*III fragments and two alphoid *EcoO109*I fragments. The intermediate group at 500-545kb indicates the possible presence of a third group of Y chromosomes and one of the individuals in the small group (3E7) had an odd combination of characteristics.

Figure3.8. Histogram illustrating the Y alphoid DNA polymorphisms.



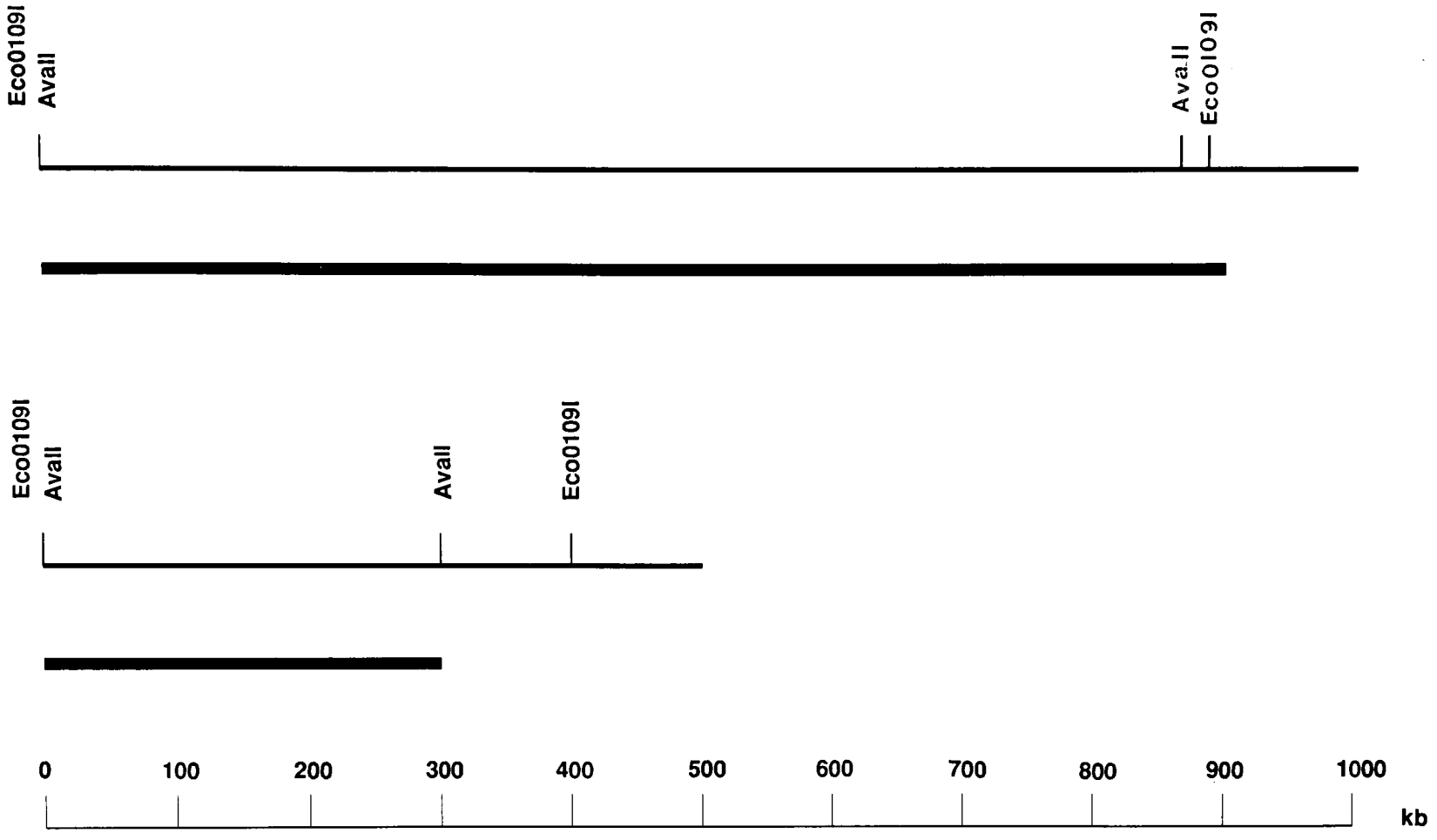


Figure3.9 The typical structure of large and small Y alphoid DNA blocks

The size of the blocks are indicated by the scale. The thick line shows the region of homology to pYalpha1.
In large blocks, the AvaII and EcoO109I sites were inside the alphoid DNA block.
In small blocks, the AvaII site was at the edge of the block and the EcoO109I site was usually 100kb outside.

There were a few individuals who had exceptional features. Male 35 (table 3.2) had an internal *Bgl*II site. Male 39 (table 3.2) had a cluster of such sites for each enzyme except *Apa*I, possibly due to an insertion of non-alphoid DNA into the block. Male 19 had an *Acc*I and a *Bcl*I fragment much larger (approximately 100kb larger) than expected. The *Apa*I fragment on the other hand was about 100kb smaller than expected. This block therefore had a few unusual features compared to the majority of the chromosomes.

The presence of three individuals between 500–545kb (figure 3.8) with intermediate characteristics, indicated the possibility of a third group. This group had a size that was intermediate between the other two groups and the internal restriction enzyme sites were a mixture of those observed in the other two groups (figure 3.8). The chromosomes in this group had two alphoid *EcoO109*I fragments typical of a large block and one alphoid *Ava*II fragment plus 5.7kb only alphoid *Hind*III fragments typical of a small type chromosome.

The highly polymorphic nature of the alphoid DNA has enabled the combination of restriction enzymes used here to distinguish 24 Y alphoid DNA alleles. The alleles are listed in column 10 of table 3.2. Since most of the Y chromosome does not recombine, the Y alleles defined represent distinct male lineages. The results define 24 male lineages, clustered into two main groups (column 11, table 3.2). The error associated with size measurements made on different gels was about ±5%; the error when samples were analysed on adjacent tracks of a gel was about ±1%. These errors were taken into account in assigning an allele. In many cases, apparent but small differences were not considered different beyond doubt and were therefore assigned to the same allele. In a few cases, sizes differed more significantly but nevertheless could not be divided into distinct categories and were assigned to one allele. The 24 allele in table 3.2 are therefore an underestimate of the number of alleles present in this population of Y chromosomes.

3.4 Analysis of meiotic stability of Y alphoid DNA

The Y alphoid DNA blocks of a family with a large number of males were analysed for variation. The restriction enzyme *Bgl*II detected two fragments; 270kb and 115kb due to an internal site for *Bgl*II in this alphoid block. The digests can be seen in figure 3.10. Track 1 shows the father; track 2 shows the mother, who had no Y alphoid DNA fragments. Tracks 3 to 12 are sons. All but one of the sons had identical bands to their father's. However, the son in track 8 had a 285kb band instead of a 270kb band. A difference might occur as a result of non-paternity or of mutation. The individual represented in track 8 was almost certainly a legitimate son of the father because, as determined by the survey of the 42 males, an internal *Bgl*II site is rare in the population but is present in all members of this family

including the non-identical individual (male 35, table 3.2). In addition, further exposure of the autoradiographs revealed cross-hybridisation to autosomal alphoid bands. In the population and in the mother and father these patterns are very variable. However all of the bands in track 8 can be identified in the DNA of either the mother or the father. The increase in size of the alphoid band is therefore likely to be due to an unequal sister chromatid exchange event which took place in male meiosis. The increase in size of the block in track 8 was about 15kb, equivalent to an addition of 2 or 3 Y alphoid DNA units.

There is another explanation for these results. It is possible that the increase in block size observed in this individual resulted from a mitotic event which took place in one or more of the lymphocytes of the individual. The lymphoblastoid cell line was established from one or a few cells, it is possible that this cell had a Y chromosome alphoid DNA block which had undergone a mitotic event and it appeared to be a meiotic event. The result could be clarified by examining a blood sample.

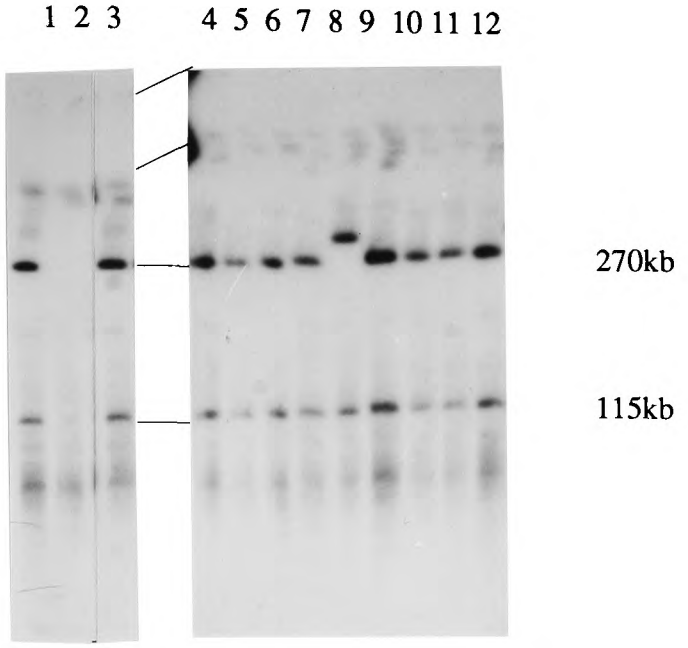


Figure 3.10Analysis of the meiotic stability of Y alphoid DNA.

Track 1 : Father. Track 2 : mother. Tracks 3 to 12 : sons.
All DNA samples were digested with BglII and hybridised with pYalpha1.
Two bands were present because the individuals had an internal BglII
site in the Y alphoid block. Track 8 had a 285kb band, not a 270kb band.

3.5 Human Y chromosome C-band polymorphisms

C-bands are associated with constitutive heterochromatin. There are two C-bands on the human Y chromosome: a small pericentric C-band and a large polymorphic C-band located on the long arm. As shown above, the Y chromosome alphoid DNA block varies in size at least four fold and as alphoid DNA is a satellite sequence and satellite sequences are found in heterochromatin, we may ask whether there is any correlation between the size of the alphoid DNA block and the size of the pericentric C-band, as would be expected if alphoid DNA was a major constituent of the central heterochromatic C-band. In addition, the size of the long arm heterochromatic block was compared to the size of the central C-band.

C-banded metaphase chromosome spreads from 10 of the 42 chromosomes listed in table 3.2 were analysed. 5 were chosen from the large alphoid block group and 5 from the small alphoid block group. The mean sizes of the long arm C-bands and the relative visibility (or the intensity of staining) of the central C-bands were determined from photographic enlargements by two independent people. The relative sizes of the central C-bands were assessed visually. The long arm C-bands were measured with a ruler and so were less subjective. The measurement of long arm C-band length was expressed as a percentage of the length of the whole Y chromosome because different Y chromosomes vary in length and there were differences in the extent of condensation of chromosomes within a preparation.

In a blind experiment, two individuals divided the central C-bands into two groups, one with small (faint and quite clear bands) and one with large (clear and very clear) C-bands. The two individuals agreed with each other in 8/10 cases. The assessment of relative intensity was therefore similar for both individuals. The size of the central C-bands did not however correspond to the amount of alphoid DNA present on each of the chromosomes. The results comparing the intensity of the central C-band, size of alphoid DNA block and the percentage of the Y chromosome consisting of long arm heterochromatin material are shown in table 3.3. An example of a C-banded Y chromosome is shown in figure 3.11.

Figure 3.11. C-banded chromosomes from a normal male

The length of the long arm C-band from the Y chromosome was measured. The intensity or visibility of the pericentric Y chromosome C-band was noted for each individual. In this case the pericentric C-band was very clear.

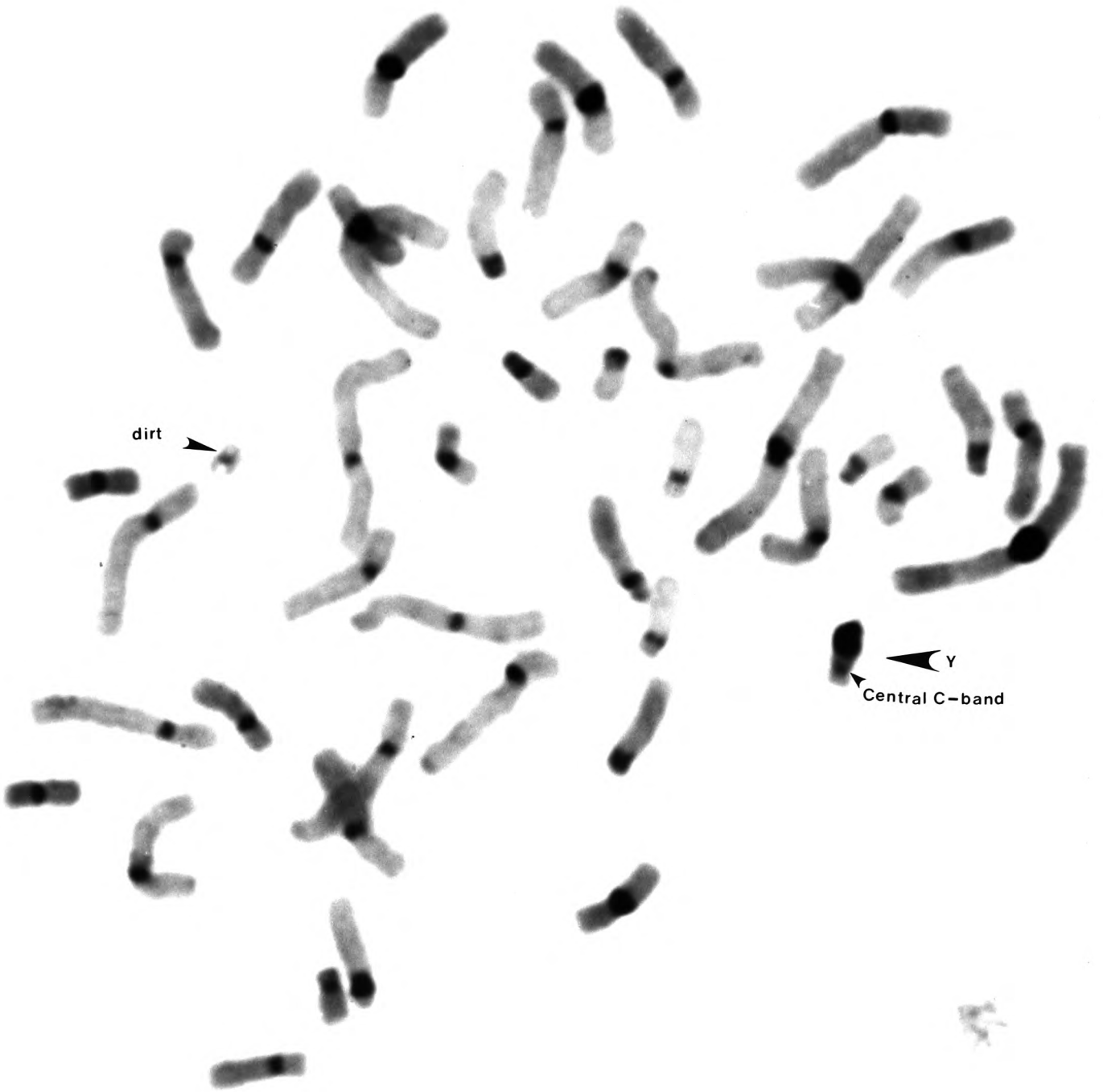


Figure 3.11. C-banded chromosomes.

The Y chromosome is indicated by the arrowhead.

The central C-band intensity is assessed from 1–4 as faint (1), quite clear (2), clear (3) and very clear(4). 1 is the least distinct and 4 is the most distinct.

Size of the alphoid DNA block (large or small) Kb	Central C-band relative intensity	Long arm C-band (as a % of the total Y)
small **245**	quite clear **2**	53
small **350**	faint **1**	53
small **275**	clear **3**	54
small **295**	quite clear **2**	50
small **415**	clear **3**	49
large **940**	clear **3**	54
large **1050**	quite clear **2**	57
large **940**	very clear **4**	59
large **1020**	quite clear **2**	56
large **920**	very clear **4**	55

Table 3.3: Comparison between C-band and alphoid DNA block size

The central C-band was polymorphic in apparent size. The size of the central C-band was not related to the amount of alphoid DNA. In other words large central C-bands did not necessarily accompany large alphoid DNA blocks and vice versa. This suggested that alphoid DNA was not the main component of the heterochromatin present in the centromeric C-band of the human Y chromosome. Some other satellite sequences must be present in this region to account for the variation in intensity of the central C-band. Satellite 3 sequences have been found near the centromere region of the Y chromosome (Tyler-Smith, personal communication). Variation in the amount of this sequence could contribute to the polymorphic nature of this C-band. The size of long arm C-band was also independent of the size of the central C-band and the alphoid DNA block on the human Y chromosome.

Chapter 4

Structure of the alphoid DNA on abnormal Y chromosomes

4.1 Choice of abnormal Y chromosome cell lines

Cell lines with a variety of abnormal Y chromosomes were used in this study. It was difficult to obtain many abnormal cell lines. Therefore, due to the simplicity of the analysis, alphoid DNA structure was examined in all abnormal cell lines available. The first question asked was whether the alphoid DNA mapped into the centromeric interval or whether any of the chromosomal rearrangements separated alphoid DNA from the Y centromere.

4.2 Structural analysis of Y alphoid DNA

Table 2.2 lists the abnormal Y chromosomes used in this study, and table 4.1 shows the structure of the abnormal Y in each of the cell lines. The 'expected' structures (column 2) indicate the approximate regions of the Y chromosome expected to be present in each cell line on the basis of the cytological description which accompanied the cell lines. The 'actual' structure (column 3) indicates the portion of the Y chromosome present in each cell line, taking into account a molecular analysis of the Y chromosome (Tyler-Smith, personal communication). The 'alphoid DNA' (column 4) indicates the presence or absence of the alphoid DNA determined in this study, for each abnormal Y chromosome. The idiograms representing the Y chromosome in table 4.1 are based on the cytological description by [Magenis *et al.* 1985].

Two of the cell lines, 863712 and 862492, had Y chromosomes described only as abnormal. Since these lines did not have a precise cytological description and have not been studied in detail for the presence of sequences other than alphoid DNA, their structures are represented by a question mark. GM03774 had a minority of cells with an iso Yq chromosome which is represented in brackets. In most cells, a chromosome fragment from the short arm of the Y was present as indicated in table

4.1 ([NIGMS 1986 1987] and Tyler-Smith, personal communication).

870740, 863829, WSM 184, 861748, 870377, GM06967 and DL-278280R are lines which have dicentric iso chromosomes. 842968, has a monocentric iso chromosome. All these Y chromosomes have a Y-derived centromere and, as expected, Y alphoid DNA. 862752 was reported to have a dicentric iso Yp chromosome and would have been expected from its structure to have had alphoid DNA. However, it did not have alphoid DNA. Since it lacked other Y chromosome sequences, the Y chromosome from this cell line was either in very low copy number or had been lost entirely during cell culture.

GM02668 has a Y chromosome with a deletion of the long arm; as expected, a Y centromere and alphoid DNA were present. GM02730 and GM03595 have similar deleted chromosomes with Y centromeres according to their cytological descriptions. The Y chromosomes had been lost from these cell lines as determined by analysis of other Y chromosome sequences (Tyler-Smith, personal communication) which accounted for the absence of Y alphoid DNA. This is also the case for 870374 and 862751 ring chromosomes. The ring chromosome from the cell line JOW had not been lost and alphoid DNA was detected. GM03774 has two abnormal Y chromosomes with Y-derived centromeres and Y alphoid DNA.

The two ill-defined chromosomes in the cell lines 863712 and 862492 have alphoid DNA but the lack of a precise description for these chromosomes meant that these lines were not very informative. Two translocation chromosomes 1491/76R and CHoP have Y derived centromeres and Y alphoid DNA. 1491/76R was of particular interest since it had been suggested that the Y centromere is present but inactive [Maserati *et al.* 1986]. If a loss of alphoid DNA accompanied centromere inactivation, then the change could have been revealed by analysis of alphoid DNA on this Y chromosome. GM02103, GM02469, GM00118 and GM07970 were translocations which had as their only detectable Y chromosome material regions from the long arm. These chromosomes do not have a Y derived centromere and as expected, they lack the Y alphoid DNA interval.

RW has an interesting chromosome which has been reported as having a very small Y chromosome with a centromere but no Y alphoid DNA [Affara *et al.* 1987]. The absence of Y alphoid DNA was confirmed, but the cells received also lacked other Y sequences expected to be present on the basis of the description given by [Affara *et al.* 1987] (Chris Tyler-Smith, personal communication). 50 metaphases were examined and found to be 45,X; in none of them was there a small fragment that could have been the RW Y chromosome (data not shown). Thus the Y chromosome fragment has been lost from this cell line.

Each cell line that was positive for the Y alphoid interval (15 out of 25 cell lines) was analysed in more detail with a series of seven restriction enzymes in the same way as the normal Y chromosomes were analysed. Figure 4.1 shows a set of seven digests *Bgl*II, *Pvu*II, *Acc*I, *Ava*II, *Bcl*I and *Apa*I for each cell line. The fragment sizes are listed in table 4.2.

Table 4.1. Abnormal Y chromosome structure

The structure of the abnormal Y chromosome(s) in each of the abnormal cell lines according to the cytological descriptions are listed in the 'expected' structure column. The structure of the abnormal chromosomes taking into account a molecular analysis of the Y chromosome are listed in the 'actual' structure column. The presence (+) or absence (−) of alphoid DNA in each of the cell lines is indicated in the alphoid DNA column. Chromosomes that had no detailed cytological description and have not been studied by molecular analysis do not have a Y chromosome structure represented in the table.

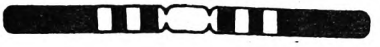
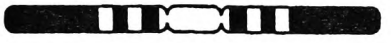
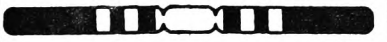
Cell Line	Expected Y chromosome structure	Actual Y chromosome structure	Alphoid DNA
870740			+
863829			+
WSM 184			+
842968			+
861748			+
870377			+
GM06967			+
DL-278280R			+
GM02668			+
GM03774			+
863712	 ?	 ?	+
862492	?	?	+
1491/76R			+
CHoP			+
JOW	ring Y	ring Y	+
870374	ring Y	—	-
862751	ring Y	—	-
862752		—	-
GM02103			-
GM02730		—	-
GM03595		—	-
GM02469			-
GM00118			-
GM07970			-
RW	0	—	-

Table 4.1 . Abnormal Y chromosome structure

Figure 4.1. Alphoid DNA fragments on abnormal Y chromosomes

Alphoid DNA fragments detected in abnormal Y chromosomes with the restriction enzymes *Bgl*II, *Pvu*II, *Acc*I, *Ava*II, *Eco0109*I, *Bcl*I and *Apa*I hybridised with pYα1 are shown. The size to the left of each of the photographs indicates the *Pvu*II fragment. Where there are two sizes, the larger size is the *Apa*I fragment size in kb.

Figure 4.1. Alphoid DNA fragments on abnormal Y chromosomes

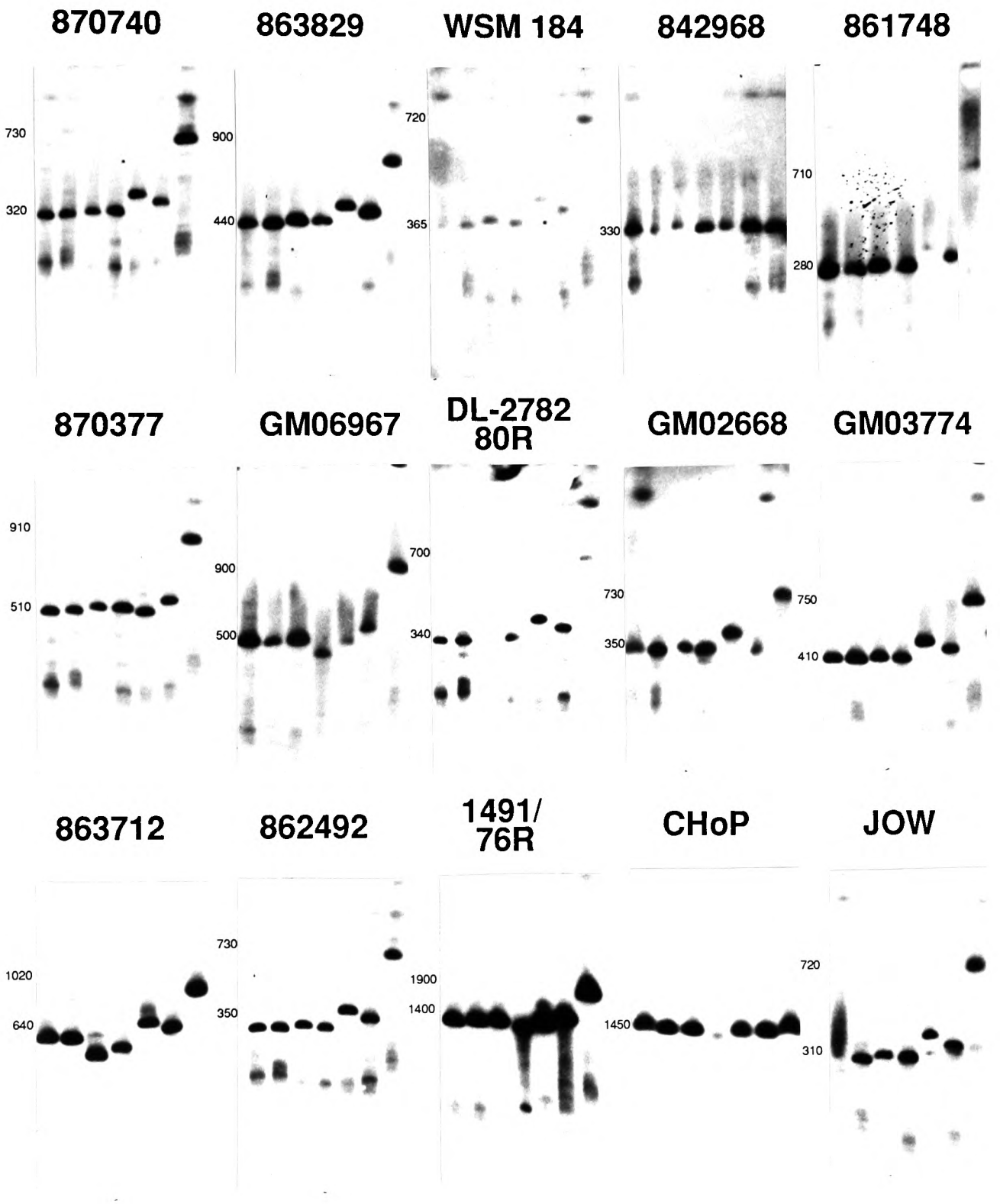


Figure 4.1 and table 4.2, show that there are nine chromosomes which have a typical small block; two chromosomes which have a typical large block; one chromosome which has an intermediate type structure; one chromosome which has a slightly different alphoid DNA structure and two chromosomes which have large differences in alphoid DNA structure. The different categories of alphoid DNA structures are discussed.

870740, 863829, WSM 184, 861748, DL-2782 80R, GM02668, GM03774, 862492 and JOW had chromosomes with typical small blocks, with no internal *Ava*II site and an *Eco0109*I site about 100kb away from the block. These were entirely normal type structures as seen in the study of 42 normal Y chromosomes (table 3.2).

1491/76R has a large block size, an internal *Ava*II site and *Eco0109*I site characteristic of a normal Y alphoid DNA large block. GM06967 has a Y chromosome alphoid DNA with an internal *Ava*II and *EcoO109*I site characteristic of a large block, and an intermediate block size (probably a small size of the large type block).

870377 has an internal *Eco0109*I site which is typical of a large type block, a block size *Ava*II fragment, characteristic of a small type block, and an intermediate block size. The characteristics of this chromosome fitted in with the intermediate group (a third possible group in the population) at 500-545kb on the histogram of normal individuals (figure 3.8).

863712 has a chromosome with an *Acc*I alphoid fragment smaller than expected probably as a result of the presence of an internal *Acc*I site. The Y in 863712 has a combination of alphoid polymorphisms from the two different groups. It has an internal *Ava*II site characteristic of a large block, an *Eco0109*I site about 100kb away from the block characteristic of a small block and an intermediate block size. These small differences are likely to represent naturally occurring polymorphisms.

842968, with a monocentric iso Yp, has an alphoid DNA structure significantly different from that seen in any normal Y chromosome. All the restriction fragments are the same size (330kb). There is a block size *Apa*I fragment, (*Apa*I fragments were 400kb larger than the block in normal chromosomes) and a block size *Bcl*I fragment (*Bcl*I fragments are 50kb larger than the block in normal chromosomes). The block also lacks internal *Ava*II and *Eco0109*I sites. The cell line CHoP also has a significantly different alphoid structure: again, all the alphoid DNA restriction enzyme fragments in the CHoP digests are the same size as each other.

These analyses indicate that most abnormal chromosomes had Y alphoid DNA block structures which were similar to those seen in normal chromosomes or differed only slightly. The two exceptions were 842968 and CHoP which are considered further in the discussion.

In addition, this analysis of abnormal Y chromosomes allowed the alphoid DNA to be localised on the Y chromosome. Figure 4.2 shows the localisation of the Y alphoid DNA interval on the human Y chromosome determined by the analysis of 861748, GM02888, 1479/76R and 842968 Y chromosomes. The boundary of the alphoid DNA interval was defined on the short arm by the Y chromosome in 861748. This chromosome has the most proximal breakpoint on the short arm as

Table 4.2 Alphoid DNA restriction enzyme fragment sizes for the abnormal Y chromosomes.

Cell Line	Description*	*Bgl*II	*Pvu*II	*Acc*I	*Ava*II	*Eco0109*I	*Bcl*I	*Apa*I
870740 (L)	45,X/46,X dicentric iso Yp	300	320	335	335	470	400	730
863829 (L)	45,X/46,X dicentric iso Yp	440	440	475	455	575	530	900
WSM 184 (F)	70% XO, 25% dicentric Yp, 5% XY	365	365	390	375	480	435	720
842968 (L)	46,X monocentric iso Yp	330	330	350	330	330	315	315
861748 (L)	48,XX dicentric iso Yq dicentric iso Yq	280	280	290	290	420	330	710
870377 (L)	46,X dicentric iso Yq	510	510	530	530	495	600	910
GM06967 (L)	45,X/46,X dicentric Y /47,X dic Y dic Y	520	500	520	450	510	580	900
DL-278280R (F)	46,X dicentric iso Yq	340	340	360	360	450	410	700
GM02668 (F)	45,X/46,X del Yq	360	350	360	360	450	360	730
GM03774 (F)	45,X+fragment/46,X dicentric iso Yq	415	410	415	410	485	450	750
863712 (L)	45,X/46,X dicentric Y	650	640	500	560	750	690	1020
862492 (L)	abnormal Y	350	350	370	350	470	400	730
1491/76R (F)	45,X t(Y;18)	1400	1400	1400	1200 +100	1280 +140	1420	1900
CHoP (L)	45,X t(Y;18)	1460	1450	1450	1420	1450	1460	1450
JOW (L)	45,X/46,X ring Y	325	310	320	310	415	360	720
870374 (L)	45,X/46,X ring Y	-	-	-	-	-	-	-
862751 (L)	45,X/46,X ring Y	-	-	-	-	-	-	-
862752 (L)	45,X/46,X dicentric Yp	-	-	-	-	-	-	-
GM02103 (F)	46,X t(X;Y)	-	-	-	-	-	-	-
GM02730 (F)	45,X/46,X del(Y)	-	-	-	-	-	-	-
GM03595 (F)	45,X/46,X del (Y)	-	-	-	-	-	-	-
GM02469 (F)	46,X t(X,Y)	-	-	-	-	-	-	-
GM00118 (F)	46,XX -15 +der15, t(Y;15)	-	-	-	-	-	-	-
GM07970 (F)	47,XX + der9, t(Y;9)	-	-	-	-	-	-	-
RW (L)	45,X/46,X +fragment Y	-	-	-	-	-	-	-

*The descriptions of the chromosomes vary in detail according to the information supplied with the cell lines. More detailed information for some of the cells is found in table 2. (L) indicates lymphoblastoid cell line (F) indicates fibroblast cell line. All fragment sizes are in kb. - indicates that no alphoid DNA was detected.

The DNA fragment sizes (in kb) for each abnormal Y were detected after digestion with the enzymes indicated in the same way as the normal Y chromosomes. In 10 out of the 25 abnormal lines, Y alphoid DNA was not detected.

determined by molecular analysis with other Y chromosome probes (Tyler-Smith, personal communication). The cell lines with chromosomes which have the most proximal breakpoints on the long arm are GM02668, with a deleted Y long arm chromosome and 1491/76R, with a translocated Y chromosome fragment. 842968, the monocentric iso Yp has a significantly different Y alphoid DNA pattern. Together, the cytogenetic and molecular analyses suggest that the breakpoint in this chromosome is likely to be within the alphoid DNA. The alphoid DNA could therefore be localised to the centromere region of this chromosome as shown in figure 4.2.

Figure 4.2. Alphoid DNA localisation on the Y chromosome

The human Y chromosome idiogram (Magenis *et al*, 1985) showing the localisation of alphoid DNA. On the right hand side of the chromosome the alphoid DNA interval is defined by the breakpoints of the Y chromosomes in the cell lines 861748 on the short arm and GM02668 and 1491/76R on the long arm. On the left hand side of the chromosome, the alphoid DNA localisation is defined by the breakpoint in the alphoid DNA block of the Y chromosome in the cell line 842968.

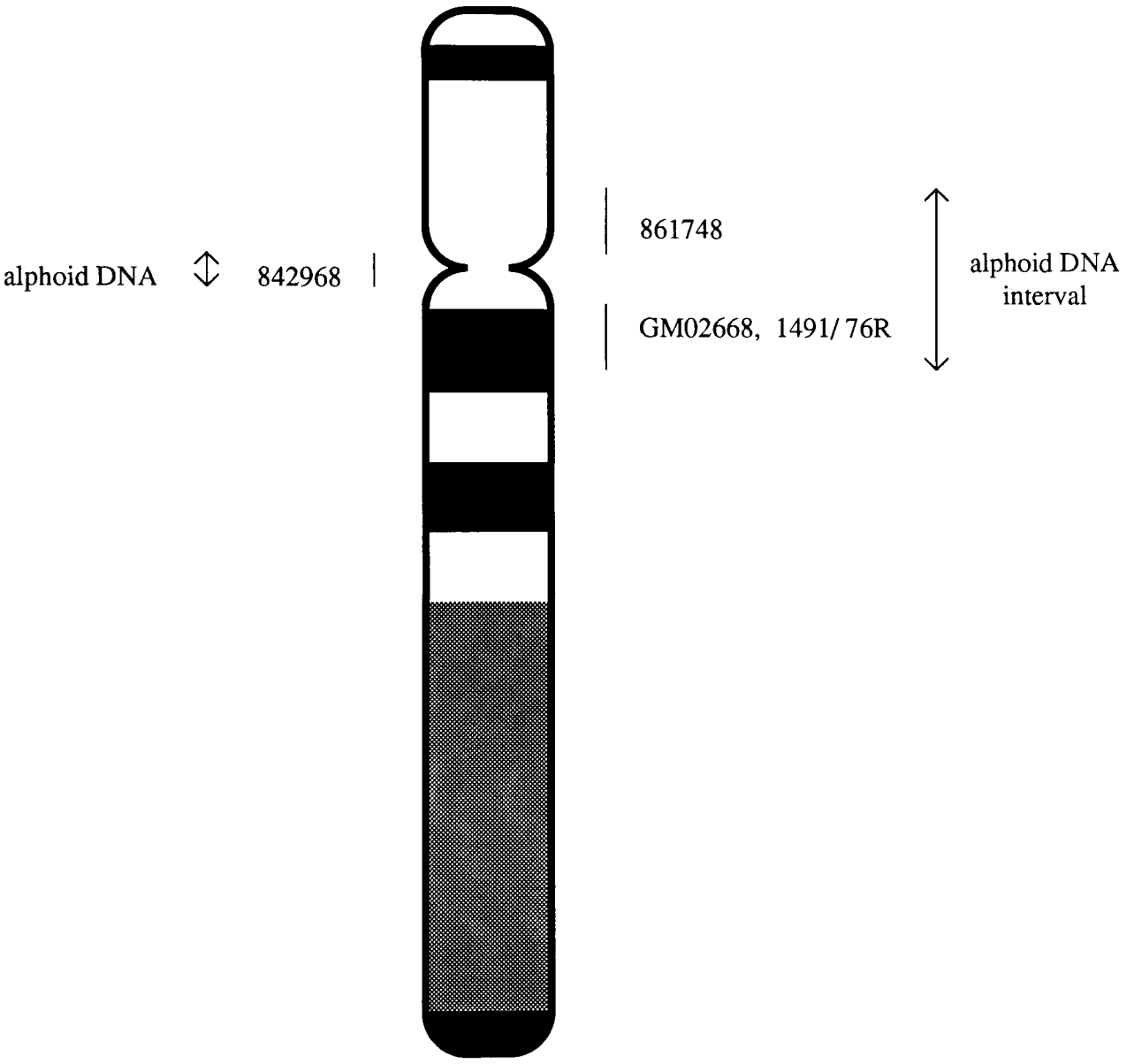


Figure4.2 Alphoid DNA localisation on the Y chromosome

Cell line	Intensity of alphoid (arbitrary)	Intensity of Y190 (arbitrary)	Size of α kb	Size of Y190 kb	Ratio of α:Y190	Assumed no of Y190 blocks	Nearest integral no. of α blocks
normal Y	247027	56610	275	570	1.0	1	1
normal Y	2476951	198344	940	570	0.8	1	1
normal Y	1713529	198344	1050	550	1.3	1	1
normal Y	101376	162741	320	570	1.3	1	1
861748	2563608	526008	280	570	1.1	2	2
842968	500645	183892	330	550	0.5	2	1
870740	762404	168880	320	565	0.9	2	2
863712	2653493	372500	640	640	0.8	2	2
863829	2493639	287489	440	565	1.2	2	2
870377	842557	76745	510	530	1.3	2	2
862492	559136	106749	350	565	0.9	?	?

Table 4.3: Estimation of the number of alphoid DNA blocks in dicentric Y chromosomes

Alphoid DNA and inactive centromeres

Centromeric sequences are often deleted in dicentric yeast chromosomes [Jager and Philippsen 1989]. In mammals, dicentric chromosomes exist and are stable because one centromere is usually inactive. The previous sections have shown that the dicentric Y chromosomes examined have normal Y alphoid structures. This could be because two normal blocks were present per chromosome, or because one normal block was present and one was entirely deleted. Therefore it was decided to measure the number of alphoid blocks per dicentric chromosome.

Hybridisation with pYα1 and comparison of the intensities of the signal between normal and abnormal cell lines would not allow the determination of the number of alphoid DNA blocks because of variation in copy number of Y chromosomes in the cell lines and the differences in cell concentration between plugs. Instead, *EcoR*V digested DNA from abnormal cell lines was hybridised with a mixture of pYα1 and Y190 at the same time. Y190 acted as an internal standard and the intensity of the two bands was compared. The results are shown in table 4.3.

There was an intensity reading for the alphoid band and the Y190 band for each cell line. Due to the differences in specific activity of the two probes, an adjustment factor was necessary to produce comparable intensity figures. It was assumed that the normal Y chromosomes had one alphoid block and one Y190 block. The intensities of the two bands were normalised by dividing the intensity of the alphoid hybridisation (column 2) by the block size (column 4) and Y190 intensity (column 3) by block size (column 5) for the first normal male in table 4.3. A ratio of one alphoid to one Y190 was assumed for this Y (shown in column

6, row 1 as 1.0). Thus, for the first Y in table 4.3; 247047 divided by 275 = 898; 56610 divided by 570 = 99. 898 divided by 99 = 9. This value was used to adjust the intensity values for the other chromosomes for differences due to probe specific activity.

The calculation of the number of blocks of alphoid DNA in an unknown chromosome can be illustrated using the data for 861748. The alphoid intensity 2653608 divided by the alphoid block size 280 = 9477. The Y190 intensity 526008 divided by the block size 570 = 923. Correction for differences due to probe specific activity gives alphoid, 9477 : Y190, 8305 or 1 : 1.1. Since a ratio of 1.0 implies one alphoid : one Y190 block and from the cytological description this chromosome has two Y190 blocks (column 7, from table 4.1), it therefore must have two alphoid DNA blocks (column 8).

When a different normal Y chromosome was used to calculate the adjustment factor for probe specific activity, the absolute values varied slightly, but the nearest integral number of alphoid blocks was the same for each chromosome.

For normal Y chromosomes the ratios implied one alphoid and one Y190 block; for the dicentric iso chromosomes, the ratios implied two alphoid blocks to two Y190 blocks. For the monocentric Yp in 842968, the ratio was 0.5. Since the chromosome has two Y190 blocks (from its cytological structure), it therefore has one alphoid block from the ratio of intensities.

4.2.1 Methylation of Y alphoid DNA

It was decided to compare the extent of methylation of the alphoid DNA in normal and abnormal Y chromosomes to see if active and inactive centromeres were associated with different methylation patterns. The restriction enzymes *Msp*I and *Hpa*II were used to detect methylation patterns in alphoid DNA since they cleave at the same site (which is present once in most units) and *Hpa*II is sensitive to CpG methylation while *Msp*I is not. The enzymes were titrated to determine the extent of maximum digestion of the alphoid DNA. The DNA was digested first with *Pvu*II to cut out the alphoid DNA from the rest of the DNA, then digested with 0, 20, 40 or 60 units of either *Msp*I or *Hpa*II. The results are shown in the first part of figure 4.3. 20 units of the enzymes were sufficient to achieve a limit digest. The two enzymes produced similar patterns indicating that there is little methylation detectable at the *Hpa*II sites.

To take into account the possibility of cell type affecting the extent of methylation of the alphoid DNA, two lymphoblastoid cell lines, OXEN and male 3 (table 3.2), and one fibroblast cell line (5659A) were digested with *Pvu*II plus either *Msp*I or *Hpa*II. The results are shown in the second part of figure 4.3. The male *Pvu*II digest in track 14 was faint due to damage of the digested plug during gel loading. There was no significant variation in the alphoid DNA methylation pattern between lymphoblastoid and fibroblast cell types: no methylation was detected in either.

The single Y centromere from 1491/76R is thought to be inactive

[Maserati *et al.* 1986]. Any difference in methylation pattern resulting from centromere inactivation would therefore have been seen in this analysis. For example, if centromere inactivation was accompanied by alphoid DNA methylation then the *Hpa*II digest would not have been as completely digested as the *Msp*I digest. However, the two digests are similar (figure 4.3). Thus centromere inactivation is not accompanied by methylation at the *Hpa*II site.

GM02668 (a deleted Yq), GM06967 (an iso Yq) and 870740 (an iso Yp) were analysed in the same way and the results are shown in the fourth and fifth parts of figure 4.3. There was no difference in fragment size between the *Msp*I and *Hpa*II digests in GMO2668, so there was no detectable methylation. In GMO2668 and 870470, the average fragment sizes in the *Hpa*II digests appear to be larger than in the *Msp*I digests, therefore there may be some methylation. However, in no case was any block entirely methylated. The significance of this slight difference is unclear.

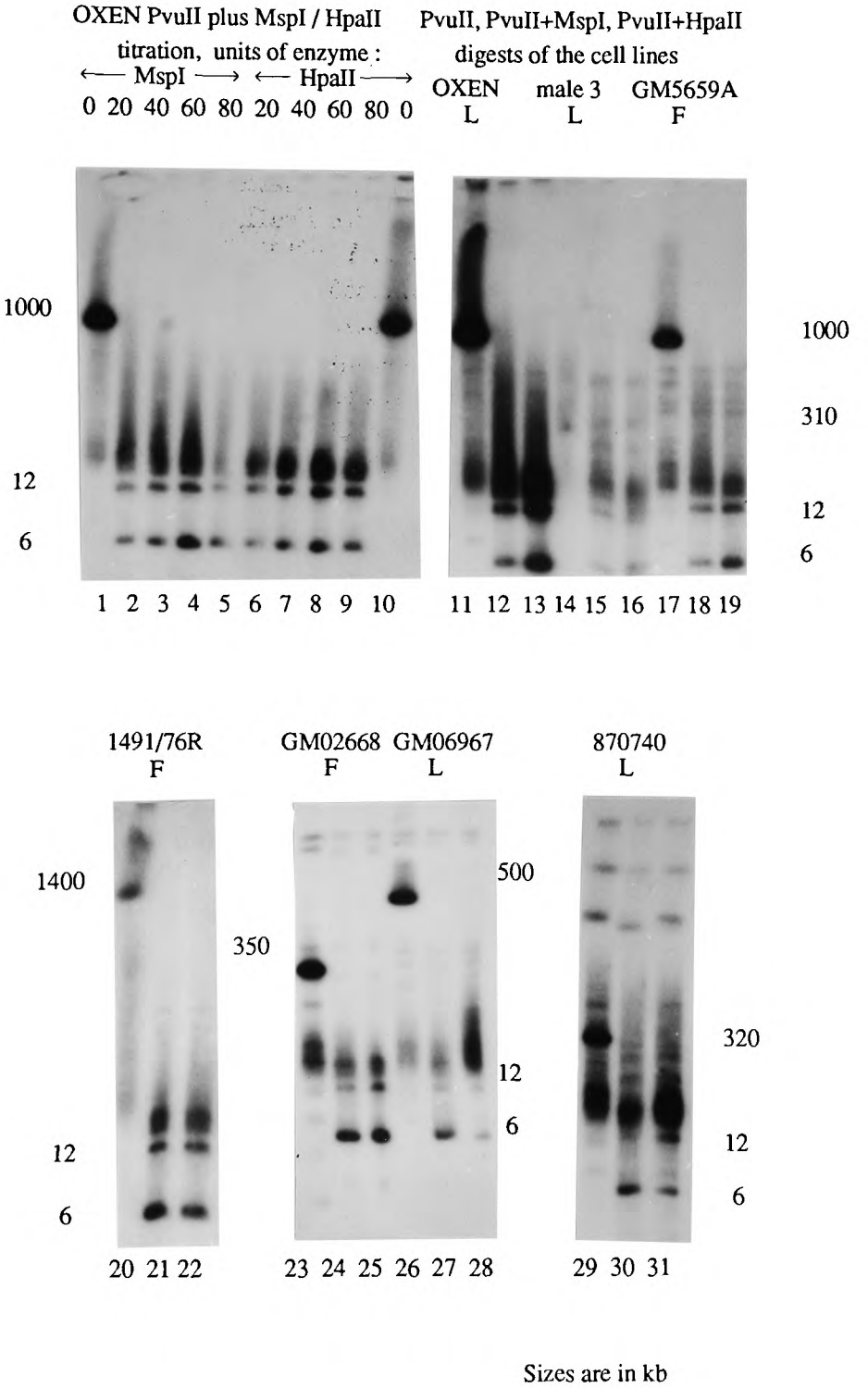
Figure 4.3. The Y alphoid DNA methylation pattern in normal and abnormal cell lines.

Lanes 1 and 10 are *Pvu*II only digests of OXEN DNA. Lanes 2 to 5 are *Pvu*II + *Msp*I double digests, with increasing units of *Msp*I from 20 to 80. Lanes 6 to 9 are *Pvu*II + *Hpa*II double digests, with increasing units of *Hpa*II from 20 to 80.

The abnormal lines used were lymphoblastoid (L) or fibroblast (F). Lanes 11 to 19 show two lymphoblastoid and one fibroblast control cell lines. Digests of DNA with *Pvu*II, *Pvu*II + *Msp*I and *Pvu*II + *Hpa*II for each of the normal Y cell lines: OXEN (tracks 11–13), male 3 (tracks 14–16) and GM05659A (tracks 17–19) are shown hybridised to pYα1, (note that male 3, lane 14 *Pvu*II only digest was damaged during gel loading and appears faint in this picture).

Digests of DNA with *Pvu*II, *Pvu*II + *Msp*I and *Pvu*II + *Hpa*II for each of the abnormal Y cell lines: 1491/76R (tracks 20–22), GM02668 (tracks 23–25), GM06967 (tracks 26–28) and 870740 (tracks 29–31) are shown hybridised to pYα1.

Figure 4.3. The Y alphoid DNA methylation pattern in normal and abnormal cell lines.



4.3 Analysis of Y alphoid DNA in Y chromosome fragments

The alphoid DNA from Y chromosome fragments created by chromosome mediated gene transfer (CMGT) was analysed. At the beginning of this study CMGT was thought to transfer intact fragments of human DNA into rodent cell lines [Pritchard and Goodfellow 1986]. Y chromosome fragments generated in this way would have been useful in the study of the Y centromere. The question to be asked would be whether there was a correspondence between alphoid DNA and an active centromere. Stable autonomous (and therefore centric) Y chromosome fragments in a rodent background would have complemented the study of naturally occurring abnormal chromosomes.

DNA from transfected cell lines was digested with *BamH*I and probed with pYα1. Pulsed field gel electrophoresis (PFGE) analysis of the alphoid DNA in these transfected chromosomes containing fragments of the Y chromosome from OXEN, revealed that these fragments were frequently rearranged and probably integrated into the host chromosomes. Figure 4.4 shows the patterns of fragments produced when some of the transfected chromosomes were digested with *BamH*I, run on a pulsed field gel and hybridised with pYα1 . Tracks 1–6 are the primary transfectants. There is a variety of patterns of Y alphoid DNA fragments. In the cell line IP2.2, one of the alphoid fragments remains in the limit mobility region of the gel, either because it is too large to separate at this pulse time or because the DNA was not completely digested. In addition, several smaller fragments were produced of approximately 270kb–510kb. In the cell line IP2.6, there are two alphoid fragments of 680kb and 460kb. In JP5.11, there was one 725kb fragment. In K1P4.2 there are two fragments of 470kb and 260kb although this track was rather thin due to damage of the digested plug during gel loading. E1P4T10 has a large alphoid fragment at the limit of mobility at this pulse time. E2P3T5 produced one 610kb fragment. The remaining 7 tracks on the gel are secondary transfectants derived from IP2.6. Each track has a different pattern of alphoid fragments, often with multiple bands: figure 4.4.

Since it is seen that most transfectants have several small bands, so transfection resulted in small pieces of alphoid DNA (smaller than the parental size of 1000kb). The presence of multiple bands shows that some lines have fragments that are rearranged, particularly the secondary transfectants.

E1P4T10, has a single fragment located in the limit mobility band. This is a large fragment and it was decided to investigate this line further. Single digests using informative restriction enzymes were performed on E1P4T10 and compared to the parental OXEN Y chromosomes alphoid DNA. *Pvu*II, *Acc*I, *Sst*I, *Bcl*I, *Bgl*II, *Apa*I and *BstX*I were chosen because they had been used to map the OXEN Y alphoid DNA and the digests are shown in figure 4.5. The fragments are the same size in OXEN and E1P4T10. This shows that E1P4T10 contains an intact fragment

of the Y chromosome at least 1400kb long.

Further digests with more infrequent cutting enzymes were used to study the longer range structure of E1P4T10 compared to OXEN, and the results are shown in figure 4.6. These digests were limit digests but were not actually complete since the restriction enzymes had CGs in their recognition sequence which are sensitive to methylation and produce multiple bands which may not reflect the distribution of sequences in the alphoid DNA of the two different cell types. Figure 4.6 shows the large fragments separated on these gels. Size determination at this size range is subject to more error than in the previous shorter pulse time gels (i.e >5%) and the yeast markers are sparse in this size range. For *Nae*I, *Bbe*I and *Mlu*I, the alphoid DNA bands in OXEN and E1P4T10 are similar in size. *BssH*II detects a 200kb difference in size between the alphoid DNA fragments in the two cell lines. The two *Sfi*I bands in the two lines are different sizes and in OXEN, *Sst*II detects two alphoid bands, but only one in E1P4T10. The extent of similarity between OXEN and E1P4T10 outside the 1400kb region is unclear.

The stability of the alphoid DNA patterns in the transfected cell lines was tested. Cells were grown in selective and non-selective medium for 4 weeks and the intensity of the alphoid DNA bands compared between the selective and non-selective conditions. All the primary fragments appeared to be stably maintained in non-selective conditions. The intensity of *BamH*I digested fragments in IP2.6 after 2 and 4 weeks is shown in figure 4.7. The fragments appeared to be stable. This may have been due to their being integrated into host chromosomes or because they had an active Y centromere. *In situ* hybridisation may have been able to distinguish between integrated and autonomous alphoid DNA fragments, but attempts at this technique were unsuccessful.

In conclusion, CMGT turned out to be less useful than anticipated in the analysis of Y chromosome centromere structure.

Figure 4.4. Transfectants digested with *BamH*I and hybridised with pYα1

Mouse cells containing Y chromosome fragments generated by chromosome mediated gene transfer were digested with *BamH*I and hybridised with pYα1. The cell lines IP2.2, IP2.6, JP5.11, K1P4.2, E1P4T10 and E2P3T5 are primary transfectants. A2, B2, D2, E1, I2, L1 and L2 are secondary transfectants derived from IP2.6. A variety of fragments were detected, most of which were smaller than the fragment detected in the parental OXEN chromosome *BamH*I digest. Sizes are indicated in kb.

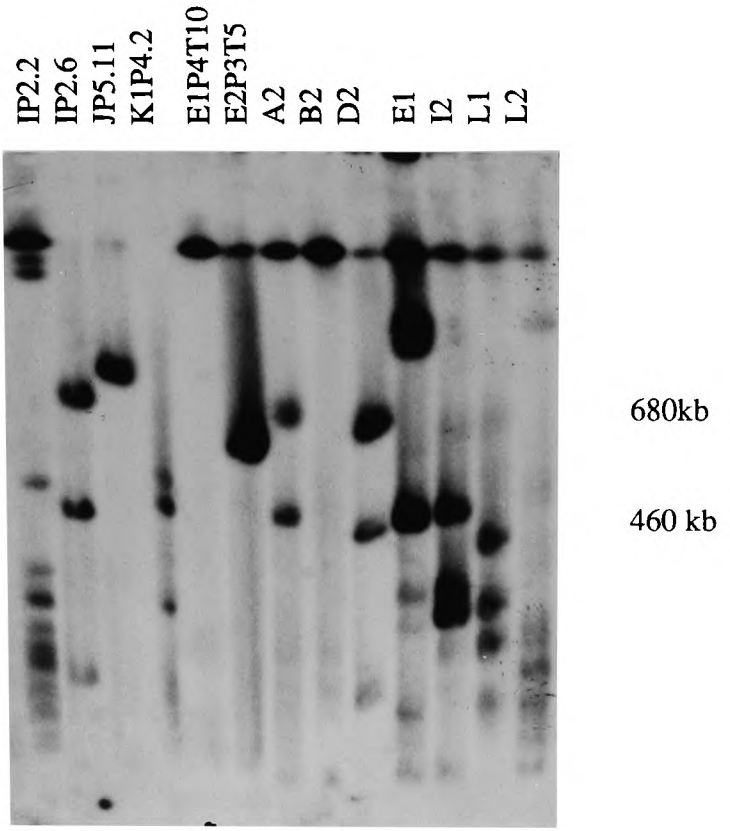


Figure 44. Transfectants digested with BamHI and hybridised with pYalpha1

Figure 4.5. Comparison of the Y alphoid DNA structure in OXEN and E1P4T10.

Restriction enzyme digests of OXEN DNA were compared to digests of DNA from E1P4T10 hybridised with pYα1. E1P4T10 is a transfected cell line containing a large fragment of OXEN alphoid DNA. Each enzyme detected the same size alphoid DNA fragment in both cell lines. The enzymes are listed for each track and the sizes are indicated in kb.

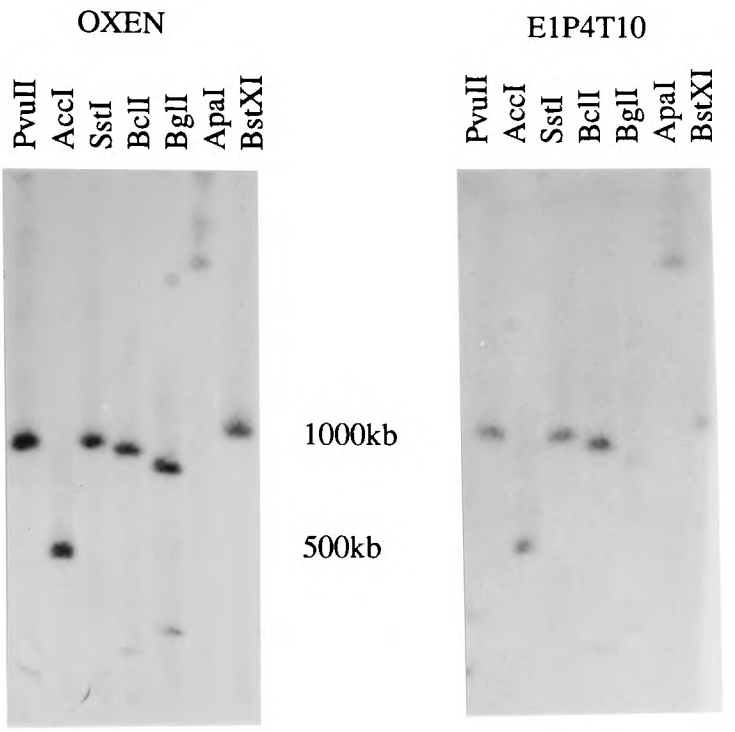


Figure 4.5. Comparison of Y alphoid DNA structure in OXEN and E1P4T10

Figure 4.6. Comparison of Y alphoid DNA structure in OXEN and E1P4T10

Digests with infrequent cutting enzymes of OXEN DNA were compared to E1P4T10 DNA digests hybridised with pYα1. The enzymes detected very large fragments in both cell lines. The fragments were not identical in all cases. *Nae*I, *Bbe*I and *Mlu*I bands were similar in size, but the *BssH*II bands and the *Sfi*I bands were different sizes in the two lines. In OXEN, *Sst*II detected two bands, but only one in E1P4T10. The enzymes are listed for each track and the approximate sizes are indicated in kb.

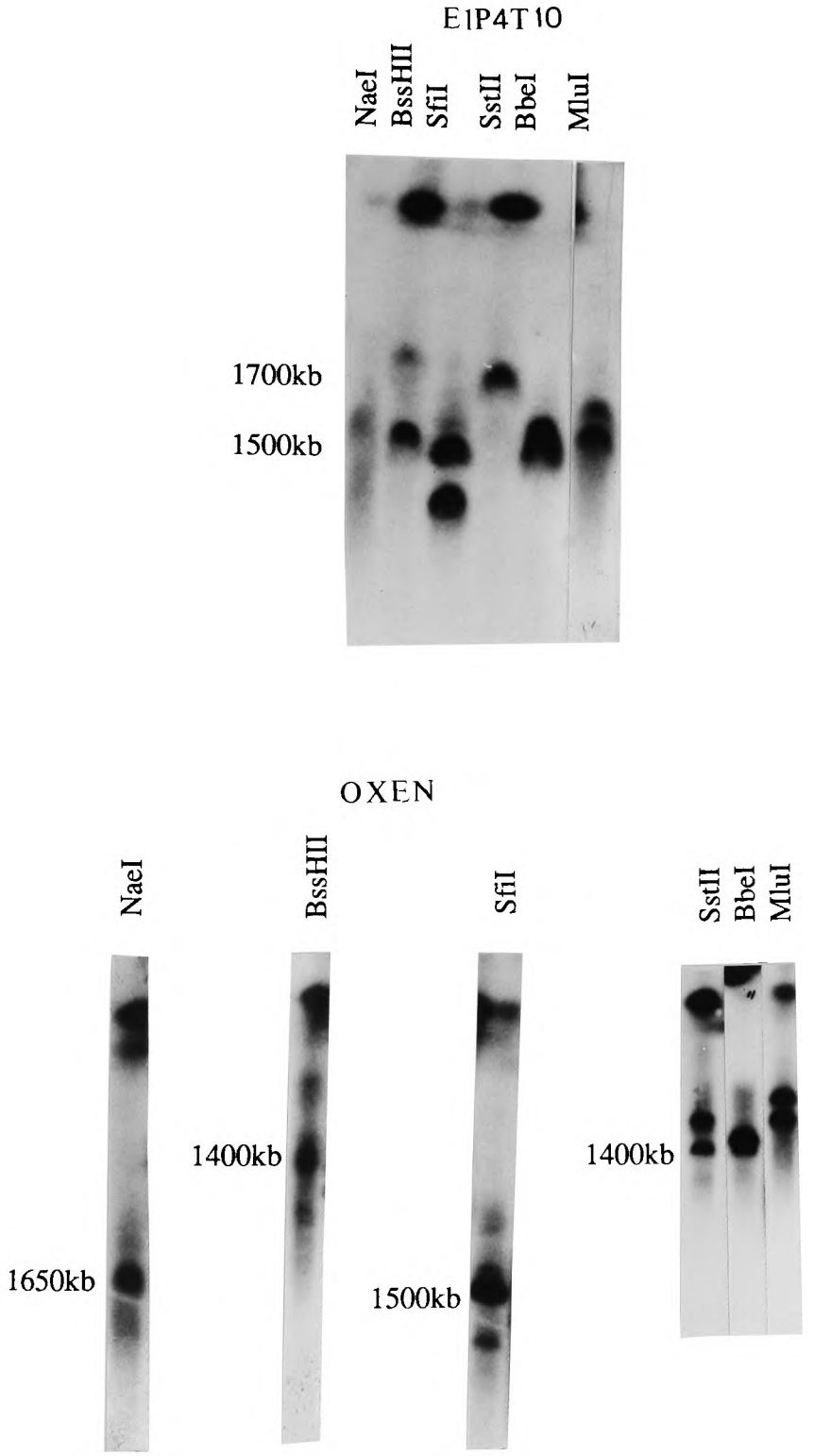


Figure 4.6. Comparison of Y alphoid DNA structure in OXEN and E1P4T10

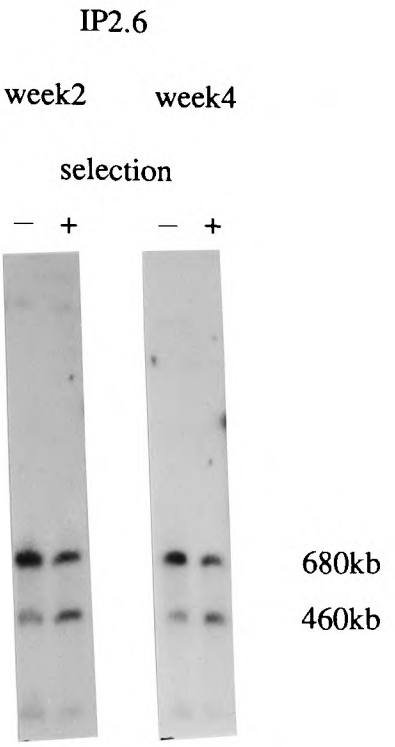


Figure 4.7. Stability of Y alphoid DNA in the transfectant IP2.6

Cells were grown in selective or non-selective medium for 4 weeks. Samples were taken at 2 and 4 weeks. The DNA was digested with BamHI and hybridised with pYalpha1.

Chapter 5

Discussion

5.1 Structure of the Y alphoid DNA in normal Y chromosomes

5.1.1 Detailed structure of the Y alphoid DNA in 3E7 and OXEN Y chromosomes

The aim of this study was to analyse the structure of the DNA in the centromere region of the human Y chromosome. Yeast centromere structure has been elucidated in *S.pombe* by mapping centromere linked sequences followed by comparison of the sequences present in normal and abnormal centromeres. Alphoid DNA is located in the centromere region of human chromosomes. Its structure was therefore examined in normal chromosomes and compared with the structures found in abnormal chromosomes.

The analysis of Y alphoid DNA began in the mouse-human Y hybrid cell line 3E7 because the Y derived alphoid DNA probe pYα1 cross-hybridised with the closely related alphoid DNA sequences on other human chromosomes. This initially made the identification of Y specific bands difficult in human cells. At the same time, the Y alphoid DNA structure was determined in the hamster-human Y hybrid cell line 853 [Tyler-Smith 1987]. These two cell lines have different alphoid DNA block sizes, 3E7 has a 430kb block and 853 has a 540kb block. There are also differences in the presence of some of the restriction enzyme sites, for example 3E7 has extra sites in its block for *Ava*II and *Acc*I (summarised in figure 3.2); 853 has an additional site in its block for *EcoO109*I [Tyler-Smith 1987].

Chromosomes in hybrid cell lines may be rearranged and unrepresentative of the normal chromosome structure: indeed there is evidence at the cytogenetic level that the Y chromosome in 3E7 is rearranged [Marcus *et al.* 1976]. Thus, it is possible that 3E7 or 853 may have different Y alphoid DNA structures to the Y alphoid DNA in normal human Y chromosomes. It was therefore decided to determine the detailed molecular structure of the Y alphoid DNA on a Y chromosome found in

a human cell line. OXEN, a 49,XYYYY lymphoblastoid cell line was chosen. One reason for this choice was that OXEN was the donor cell line for the chromosome mediated gene transfer transfectants studied.

The cross-hybridisation with non-Y chromosome alphoid DNA was reduced by increasing the stringency of the hybridisation conditions. This allowed the Y-specific bands to be distinguished from the non-Y alphoid bands. The detailed structure of the Y alphoid DNA from OXEN was determined in a similar way to that of 3E7. The OXEN Y has a large 1000kb alphoid block which contained additional restriction enzyme sites not found in 3E7 or 853, these were *Bgl*I, *Sst*I, and *Bcl*I. It also has additional *Ava*II and *Acc*I sites found only in 3E7 and an additional *EcoO109*I site found only in 853. (summarised in figure 3.5).

Thus there is variation in the alphoid block size and presence of restriction enzyme sites in the three Y chromosomes. Due to the variation observed between these blocks, to understand fully the structure of this region and make meaningful comparisons between normal and mutant Y centromeres, it was decided to determine the restriction site polymorphisms in a larger collection of normal males.

5.1.2 Y alphoid DNA polymorphisms in 42 normal Y chromosomes

The Y alphoid DNA from a random collection of unrelated Y chromosomes was studied. The Y alphoid block of each of these chromosomes was analysed with seven restriction enzymes chosen from those used to make the detailed maps of the 3E7 and OXEN Y alphoid DNA to detect restriction fragment length polymorphisms (RFLPs). RFLPs can arise from point mutations which result in the loss or gain of restriction enzyme sites, and from differences in the number of units in a tandemly repeated array. Variation in tandem units may result from unequal sister chromatid exchange. The block sizes (number of units in the alphoid DNA block) and distribution of RFLPs in the collection of males was determined. Comparisons between Y chromosomes revealed that the Y alphoid DNA block was very polymorphic in size and restriction enzyme site distribution. In addition, The size of the alphoid DNA units was investigated using *Hind*III. *Hind*III had one site in a 5.7kb unit and two sites in a 6.0kb unit. So blocks with 5.7kb units only could be distinguished from blocks with 5.7kb + 6.0kb units. Some of the alphoid DNA blocks had 5.7kb units only. Other blocks had 6.0kb units as a subpopulation in addition to the 5.7kb units.

The alphoid block size and associated polymorphisms are represented in the form of a histogram (figure 3.8). The Y chromosomes fall into two distinct groups. One of these groups has a small block size, between 200kb and 445kb. Small blocks had an *EcoO109*I fragment approximately 100kb larger than the block one *Ava*II fragment and 5.7kb but no 6.0kb alphoid units. 3E7 had one *EcoO109*I fragment characteristic of a small block, but internal *Ava*II sites and 5.7kb + 6.0kb units

characteristic of large blocks. The small group had 21 members who were all Caucasian or part Caucasian in origin. The second major group had a large block size between 650kb and 1050kb. Large blocks had an internal *Ava*II site, an internal *EcoO109*I site and 5.7kb + 6.0kb units. The large group had 17 members either Caucasian or Asian (including Oriental) in origin. Three members of the group of 42 chromosomes had blocks which were between 500–545kb in size and had an internal *EcoO109*I site (characteristic of a large block) but no internal *Ava*II site and only 5.7kb alphoid units (characteristic of a small block). This may represent a third group in the population. The alphoid DNA polymorphisms defined by these restriction enzymes allow Y chromosome alleles to be defined for each normal male analysed.

5.1.3 Y chromosome alleles and paternal lineages

Mitochondrial DNA (mtDNA), has three properties which make it useful in the study of female lineages. It has a maternal mode of inheritance, it does not recombine and it has a rapid rate of evolution compared to nuclear DNA [Stoneking *et al.* 1986]. Restriction enzyme analysis of polymorphic sites has allowed mtDNA structures to be divided into groups which define maternal lineages and to some extent reflect the geographic or racial origins of the group members. The analysis allows conclusions to be drawn on how, when and where the human gene pool arose. A tree relating the 133 types of human mtDNA and a reference sequence was built, which links modern human populations to a common ancestral female. In the tree, a time scale was derived by assuming that mtDNA sequence divergence accumulates at a constant rate in humans between 2–4% per million years [Cann *et al.* 1987]. The authors concluded that all the mtDNAs tested stem from one woman who is postulated to have lived about 200 000 years ago, probably in Africa [Cann *et al.* 1987].

In an analogous way, the Y chromosome with its haploid mode of inheritance provides a vehicle for the study of male lineages. The human Y chromosome is relatively devoid of polymorphic genetic markers [Casanova *et al.* 1985]. Y-linked RFLPs allow Y-specific alleles to be constructed. The polymorphic nature of the alphoid DNA on the human Y chromosome has revealed 24 alphoid DNA alleles in 42 Y chromosomes. These alleles are shown in table 3.2.

The bimodal distribution of alphoid DNA block sizes was an unexpected finding. A single normal distribution of block sizes would be expected from a random sample of Y chromosomes. It seems unlikely that it would be an advantage to an organism to have a small or a large block of alphoid DNA. This makes it unlikely that the alphoid DNA sequence is under any selective pressure to cluster into two distinct groups. Two explanations are possible to account for the two groups of chromosomes. The Y chromosomes could have become clustered into two groups by a process of convergent evolution. This is a somewhat unlikely possibility since it would require four independent features to become associated with one another. These features are alphoid block size, and *Ava*II, *EcoO109*I and *Hind*III restriction

site polymorphisms. It is more likely that the Y chromosomes in each group descended from a separate male ancestor. The ancestor of group 1 would have had a small alphoid block lacking internal *Ava*II sites, internal *EcoO109*I sites and 6.0kb units with *Hind*III sites. The ancestor of group 2 would have had a large alphoid block with additional sites for *Ava*II, *EcoO109*I and *Hind*III. 38/42 (=90%)of the individuals fall into the two groups; these groups are therefore likely to represent a large proportion of the population.

Descent of present day chromosomes from a single male ancestor is inevitable. It is surprising however, that most of the Y chromosomes examined fall into two main groups indicating ancestry from two males. This does not mean that there were only two males around at the time, but that two male lineages have become dominant in the population and others have died out. The lineages defined here from Y chromosome alleles refer to the Y chromosome lineages not to the inheritance of non-Y sequences. The Y haplotype distribution found in this study suggests that the most common mutations in alphoid DNA are small changes in block size. Large changes in block size and point mutations detected by this method are predicted to be less common. These predictions are supported by the limited data available.

Meiotic stability of Y alphoid DNA

Analysis of mutations in the alphoid DNA block within male members of families to determine the mutation rate and average change in block size per mutation would allow the divergence time of each group to be estimated. The divergence time since the common ancestor of all groups or the Y chromosome 'Adam' could also be estimated.

One family has been investigated. One male member out of 10 showed an increase in alphoid DNA block size of about 15kb or 2–3 units. This small change in block size is probably due to an unequal sister chromatid exchange event which took place during male meiosis. Many more families must be studied in order to produce a statistically significant figure for the rate of mutation of the alphoid DNA. The rate of mutation could then be used to construct estimates for the divergence time of this sequence.

Computers can be used to simulate the process of evolution. Divergence of the alphoid DNA was simulated (J.K.Elder, personal communication). In these simulations, the evolution of one chromosome is followed and the simulation is repeated many times.

The variable parameters are length of DNA sequence to begin with ; the number of crossovers at male meiosis and of these the number that are offset; the number of units that the offset involved; the number of generations and the number of trials. 50 units was chosen as the starting length since this is the average number of units in a small block, this parameter was kept constant. The number of trials was kept constant at 10 000. This parameter affected the smoothness of the curve but not the shape.

The number of crossovers × the probability of a mismatch was varied: 0.001 (figure 5.1), 0.01 (figure 5.2) and 0.1 (figure 5.3). (The single occurrence of a crossover that changed the size of the alphoid block in the family analysed in this study was 1 in 10 a probability of 0.1, but this represents only one set of data). The number of units of offset was assumed to follow a normal distribution with a standard deviation (SD): 1 (bottom rows), 3 (middle rows) and 5 (top rows). The number of generations was varied: 1000 generations ~25 000 years (first columns), 10 000 generations ~250 000 years (second columns) and 100 000 generations ~2 500 000 years (third columns). At each generation, the computer selects chromosomes at random which will contribute to the progeny. The sizes of the blocks in the final population are plotted against the numbers of individuals in the form of a histogram. The results are shown in figures 5.1, 5.2 and 5.3 .

A total of 27 simulations was carried out including all possible combinations of the three variable parameters as described above. Each histogram has a set of parameters, for example, 50 0.001 5 1000 10 000; the variable parameters are, a probability of a change in block size of 0.001; a SD of 5 and 1000 generations.

The histogram obtained when actual alphoid block sizes were sampled is shown in figure 3.8. The spread of sizes is small: a cluster of sizes range from 245kb (41 units) to 445kb (74 units). The simulations show a range of different distributions. Two of them give a curve like the one sampled from the population. One on figure 5.1, bottom row, column 2 has a probability of change in block size of 0.001; a SD of 1 and 10 000 generations. The second histogram which gives a similar curve is in figure 5.2, bottom row, column 1 with a probability of a change in block size of 0.01, a SD of 1 and 1000 generations. These simulations suggest that either the probability of a change in block size is low, 0.001 and the SD (or number of units in an exchange) is 1; or that the lineage is not very old, only about 1000 generations.

With the limited data on meiotic stability in this study, the observed occurrence of a change in alphoid block size was 10%. If this frequency was representative of the actual frequency, then from these simulations, the lineages would have to be less than 1000 generations old. In figure 5.3 the probability of a change in length of the block is 0.1. Even with a SD of 1, the spread is large (bottom row, column 1). All the other histograms in figure 5.3 show a very broad spread of sizes. The ones that appear empty actually have small numbers of individuals scattered around the base line. So, it seems that unless the lineages are indeed very recent (about 2500 years old) which seems unlikely, and the SD is 1, the frequency of a change in block size is unlikely to be as large as 10%.

Although these simulations may not reflect the actual processes going on during unequal sister chromatid exchange, they do provide guidelines for interpreting the data that are available. It would be advantageous to analyse more families to determine more accurately the probability of a crossover changing a block length and by how many units. Knowing the frequency of crossover and SD would allow a time scale (number of generations) to be fixed more accurately to the lineages. This could produce a complementary estimate to the 200 000 year old 'Eve'.

In summary, these simulations suggest that the lineages are not very old, the number of units involved in a crossover is small and the probability of a change in length occurring is likely to be low.

Figure 5.1; 5.2 and 5.3. Computer simulations of alphoid DNA divergence

A computer program produced the 27 histograms in figures 5.1, 5.2 and 5.3. The program simulated the process of evolution. Five parameters were used, of which two were kept constant and three were varied. These parameters are: length of sequence to begin with; the number of crossovers at male meiosis and of these, the number that are offset; the number the offset involved; the number of generations and the number of trials. The starting block length and the number of trials was kept constant. The probability of a change in length was varied, 0.001 in figure 5.1; 0.01 in figure 5.2 and 0.1 in figure 5.3. The average number of units involved in a crossover was varied; 5 in the top rows of each figure, 3 in the middle rows and 1 in the bottom rows. The number of generations was varied; 1000 in the first columns, 10 000 in the second columns and 100 000 in the third columns. At each generation, the computer selects sequences at random which will contribute to the progeny. The sizes of the blocks in the final population are plotted against numbers of individuals in the form of histograms.

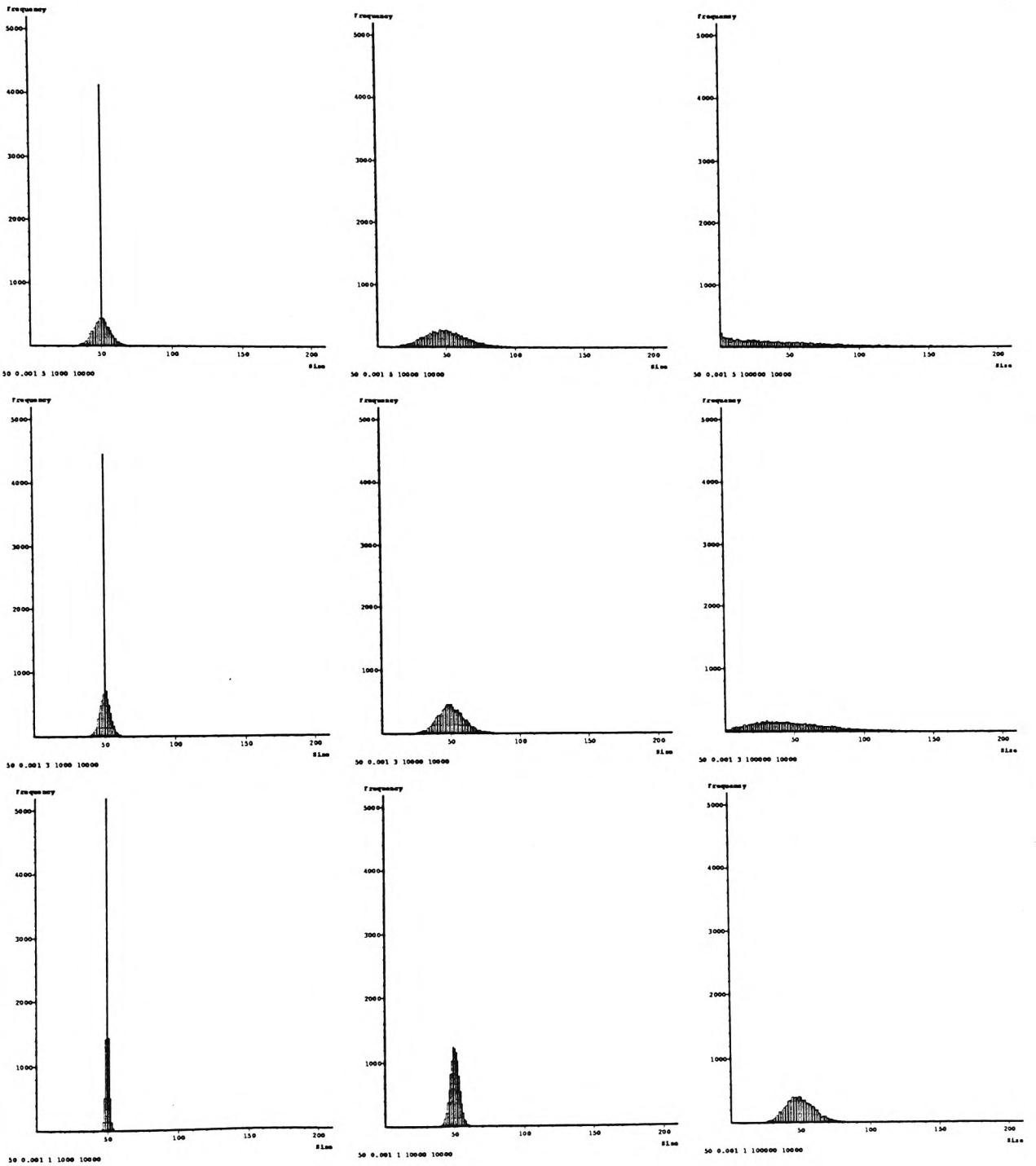


Figure 5.1

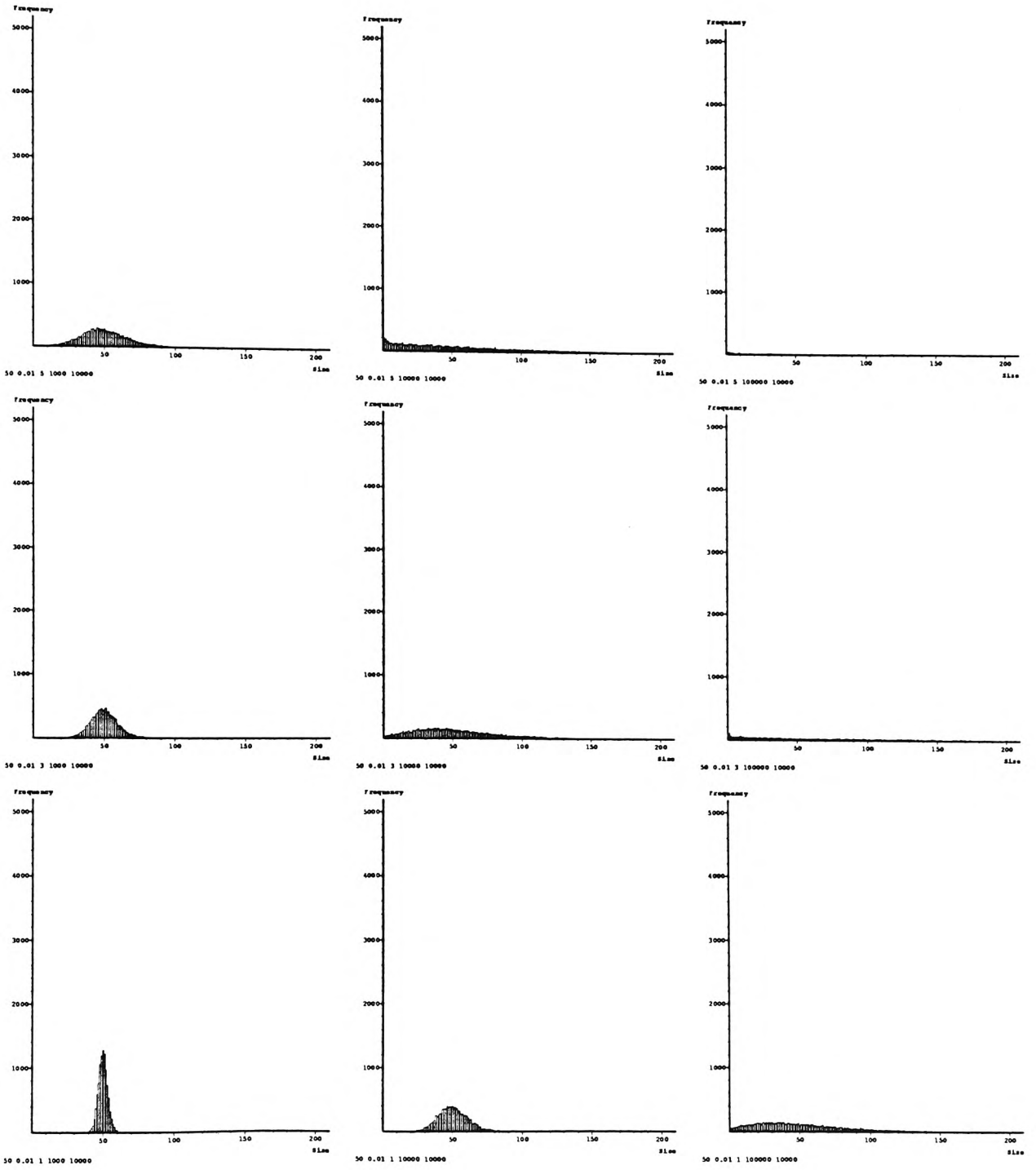

Frequency
Size
50 0.01 5 1000 10000
50 0.01 5 10000 10000
50 0.01 5 100000 10000
50 0.01 3 1000 10000
50 0.01 3 10000 10000
50 0.01 3 100000 10000
50 0.01 1 1000 10000
50 0.01 1 10000 10000
50 0.01 1 100000 10000

Figure 5.2

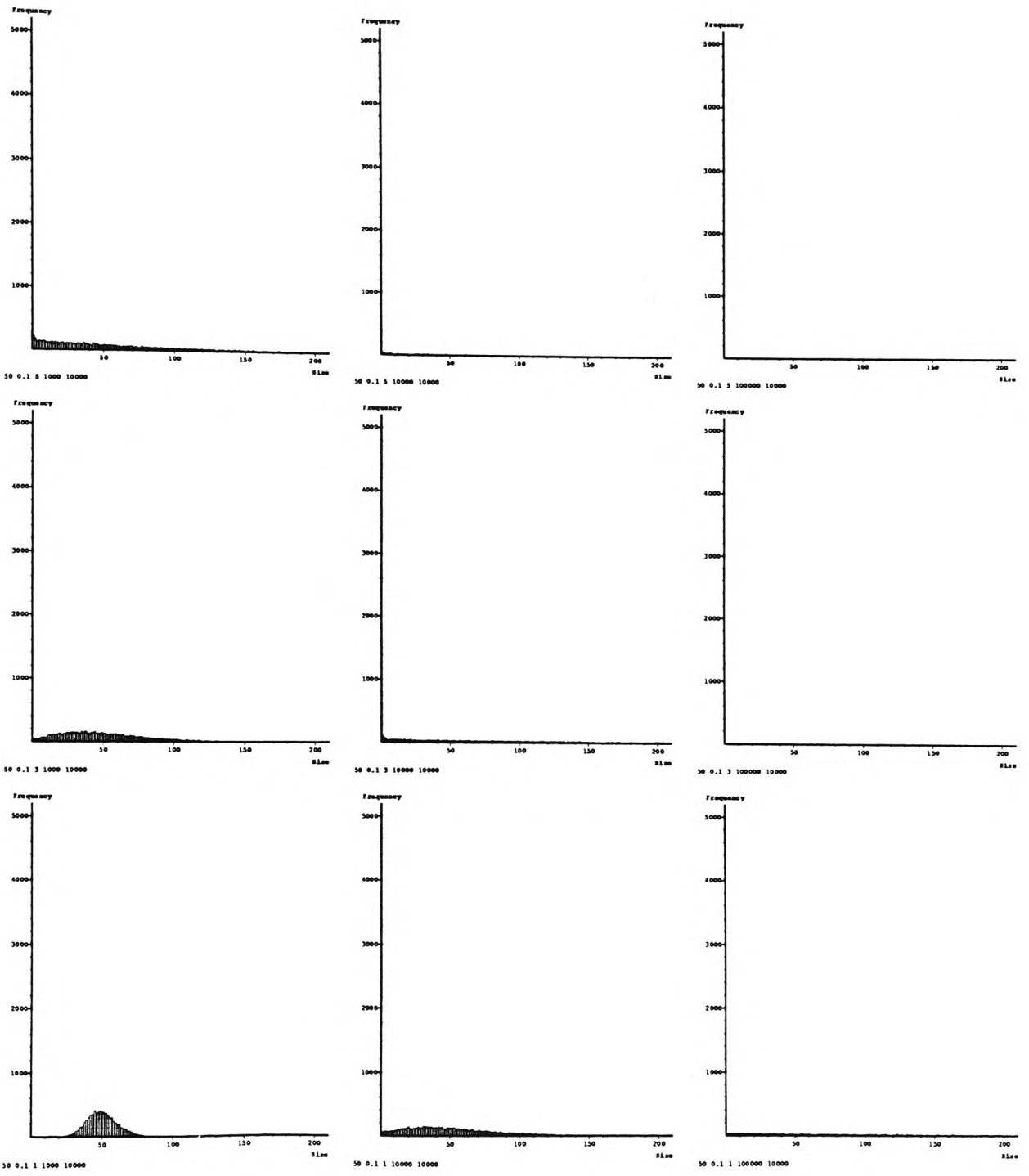

Frequency
5000
4000
3000
2000
1000
50
100
150
200
Size
50 0.1 5 1000 10000
50 0.1 5 10000 10000
50 0.1 5 100000 10000
50 0.1 3 1000 10000
50 0.1 3 10000 10000
50 0.1 3 100000 10000
50 0.1 1 1000 10000
50 0.1 1 10000 10000
50 0.1 1 100000 10000

Figure 5.3

5.1.4 Polymorphic sequences elsewhere on the Y chromosome which define Y chromosome haplotypes

The 42 normal Y chromosomes were analysed with two other polymorphic loci on the human Y chromosome: Y190 and poxY1 (Tyler-Smith, personal communication). Y190 is a moderately abundant tandemly repeated sequence located on the short arm of the Y [Muller *et al.* 1986, Tyler-Smith *et al.* 1988]. poxY1 is detected by a low copy probe adjacent to simple sequence DNA (Tyler-Smith, unpublished data); located on the long arm of the Y chromosome. Variation in block size at the Y190 locus was detected by *Xba*I and poxY1 by *Xba*I and *Bgl*II. The data are presented in table 5.1.

At the Y190 locus, *Xba*I detects block sizes between 360kb (male 18) and 800kb (OXEN); male 6 and 3E7 have additional point mutations creating internal sites (table 5.1, column 10). 12 haplotypes were distinguished (table 5.1 column 14). At the poxY1 locus, *Xba*I detected 7 haplotypes (table 5.1, column 15) from 155kb to 400kb in size (table 5.1, column 11) and *Bgl*II revealed 12 haplotypes (table 5.1, column 16) between 790kb and >3000kb in size (table 5.1 column 14).

Allele is used to describe the polymorphic features of one locus on a chromosome. When more than one locus is involved, the term haplotype has been used. In combination with the alphoid DNA alleles, the 3 loci define Y chromosome haplotypes which distinguish 33 paternal lineages (table 5.1, column 17). The distribution of poxY1 sizes is not random between the two groups defined using the alphoid probe; 21/21 members of group 1 (small group) have a poxY1 *Bgl*II fragment >1300kb while 15/17 members of group 2 (large group) have a fragment $\leq$1300kb. The poxY1 fragment distribution is illustrated in the histogram in figure 5.4 along with the alphoid DNA polymorphism distributions. The distribution of Y190 repeat block sizes in groups 1 and 2 are similar, although particular haplotypes are rarely found in both groups.

The poxY1 fragment distribution supports the alphoid findings that the chromosomes fall into two groups as illustrated by the symbols inside the squares in the histogram in figure 5.4. Since these loci are at different locations on the chromosome, these data also suggest that the two groups have arisen as a result of descent from two ancestral Y's. It is even more unlikely that a convergent evolution process could have been involved in clustering the poxY1 polymorphisms with the alphoid DNA polymorphisms.

Table 5.4DNA restriction fragment sizes and haplotype assignments for 42 normal males.

	alphoid								Y-190	poxY1		allele					
	BglII	PvuII	AccI	AvaII	EcoO109I	BclI	ApaI	HindIII	XbaI	XbaI	BglII						
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	265	275	285	280	380	330	650	-	570	350	1350	22,	9,	3,	2	27	1
2	320	320	330	330	480	380	745	-	570	350	1350	21,	9,	3,	2	25	1
3	310	310	330	310	450	380	695	-	570	350	1340	22,	9,	3,	2	27	1
4	360	350	360	355	430	410	720	-	590	350	1325	21,	8,	3,	3	24	1
5	275	280	275	275	390	355	690	-	590	350	1340	22,	8,	3,	2	28	1
6	840	810	790	680+90	930+175	900	1200	+	625+90	360	1020	11,	3,	2,	8	13	2
7	330	330	330	330	420	375	720	-	570	350	1340	22,	9,	3,	2	27	1
8	320	320	340	340	420	400	750	-	590	350	1355	22,	8,	3,	2	28	1
9	910	910	950	840+90	950+120	1030	1300	+	570	315	1300	7,	9,	5,	4	9	2
10	315	315	335	335	425	380	680	-	590	335	1340	22,	8,	4,	2	29	1
11	415	415	430	415	500	460	790	-	585	345	1350	18,	8,	3,	2	21	1
12	950	940	990	910+90	950+120	1030	1320	+	570	290	1140	6,	9,	6,	7	7	2
13	415	415	430	415	515	475	840	-	590	350	1340	17,	8,	3,	2	20	1
14	940	920	960	850+100	925+120	1010	1330	+	570	285	1120	7,	9,	6,	7	8	2
15	295	295	315	295	400	350	665	-	590	350	1340	22,	8,	3,	2	28	1
16	970	950	990	920+100	970+120	1060	1340	+	570	290	1140	6,	9,	6,	7	7	2
17	335	335	360	335	425	400	720	-	570	350	1350	21,	9,	3,	2	25	1
18	300	295	305	295	395	350	685	-	360	350	1350	22,	12,	3,	2	30	1
19	1065	1050	1130	1010+90	1070+120	1205	1360	+	550	290	1115	1,	10,	6,	7	1	2
20	270	270	285	285	400	360	690	-	615	350	1350	22,	7,	3,	2	31	1
21	310	310	325	325	425	370	700	-	590	350	1350	23,	8,	3,	2	32	1
22	680	660	620	700+80	640+150	660	1060	+	690	350	1345	13,	4,	3,	2	15	2
23	950	940	950	890+80	910+120	1000	1320	+	670	155	1170	6,	5,	7,	6	6	2
24	350	350	370	360	470	410	740	-	590	350	1360	22,	8,	3,	2	28	1
25	1020	1020	1080	950+90	1020+120	1100	1340	+	410	300	1115	2,	11,	6,	7	2	2
26	335	335	350	345	465	415	725	-	610	340	1350	21,	7,	4,	2	26	1
27	800	800	800	730+70	760+120	820	1200	+	570	315	>3000	12,	9,	5,	1	14	2
28	860	860	860	730+70	840+120	900	1180	+	690	360	870	10,	4,	2,	10	12	2
29	545	545	550	545	510+120	590	920	-	545	330	1360	15,	10,	4,	2	17	-
30	345	345	360	360	465	410	740	-	590	350	1360	20,	8,	3,	2	23	1
31	280	275	290	280	385	335	670	-	590	330	1355	22,	8,	4,	2	29	1
32	245	245	255	250	350	305	640	-	590	400	1355	24,	8,	1,	2	33	1
33	880	870	890	575+175 +60	870+120	940	1300	+	650	350	790	8,	6,	3,	12	10	2
34	505	505	520	510	490+120	570	900	-	545	350	1360	15,	10,	3,	2	18	-
35	270+115	375	385	375	475	435	760	-	590	350	1335	19,	8,	3,	3	22	1
36	920	920	920	840+100 +50	910+120	960	1260	+	640	360	1295	9,	6,	2,	4	11	2
37	650	650	650	600+70	630+120	725	1030	+	690	360	840	14,	4,	2,	11	16	2
38	1000	1000	1010	950+70	1000+120	1050	1270	+	665	330	920	5,	5,	4,	9	5	2
39	650+360	650+360	650+375	650+300 +100	675+350	675+425	1380	+	665	310	1180	3,	5,	5,	6	3	2
OX	970	1000	500+500	930+80	910+120	950+130	1400	+	800	340	1125	4,	1,	4,	7	4	2
3E7	400	430	350+45*	190+140 +100	610	520	890	+	660+ 640+90	365	1210	16,	2,	2,	5	19	-
853	540	540	530	530	530+120	610	930	-	540	345	1350	15,	10,	4,	2	17	-

*Each band consisted of several fragments arising from partial cleavage of internal sites.

DNA fragment sizes (in kb) and haplotypes of 42 chromosomes. Column 1 lists the individuals. Columns 2-8 show the size(s) of the alphoid fragment(s) detected after digestion with the enzymes indicated. Column 9 shows the presence of the 4.1kb and 1.9kb alphoid DNA fragments (+) or their absence (-). Column 10 shows the size(s) of the Y-190 repeat DNA fragment(s) detected in an XbaI digest. Columns 11 and 12 show the size of the poxY1 fragment detected after digestion with XbaI or BglII. Columns 13-16 indicate the allele assigned to each chromosome for each locus: column13: alphoid; column 14: Y-190; column 15 poxY1 XbaI; column 16: poxY1 BglII. Column 17 indicates the paternal lineage defined by the compound haplotype. Column 18 indicates the group to which the chromosome belongs.

Figure 5.4. Distribution of alphoid DNA sizes and associated polymorphisms on 42 Y chromosomes.

Each square represents one individual. The size of the *Pvu*II fragment (the sum of the two fragments for individual 39) was used as the measure of block size. Additional polymorphisms found in the individuals are indicated by the symbols inside the square. This histogram is the same as that in figure 3.8, except that the poxY1 polymorphisms are also shown by symbols inside the square.

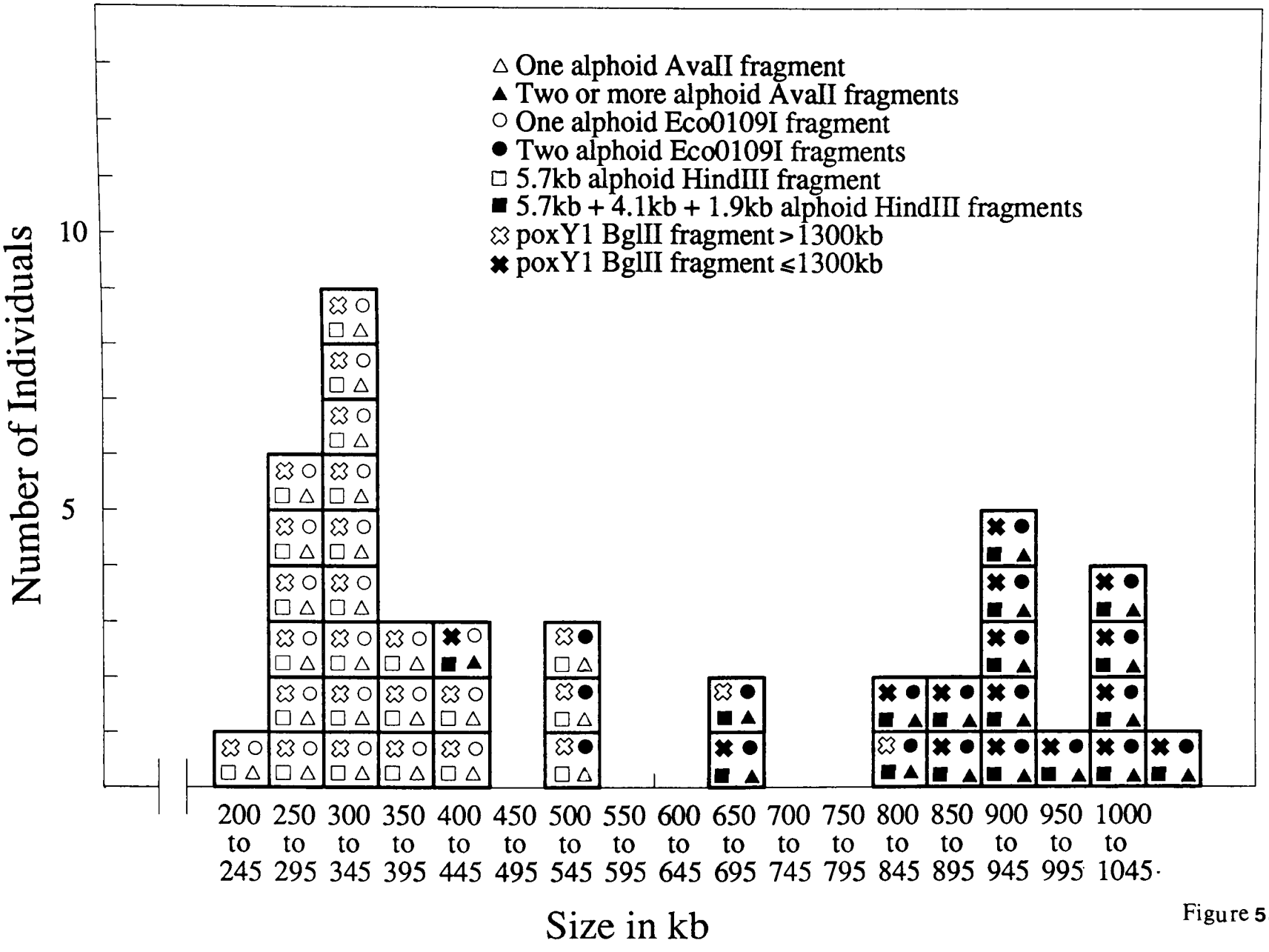

△ One alphoid AvaII fragment
▲ Two or more alphoid AvaII fragments
○ One alphoid Eco0109I fragment
● Two alphoid Eco0109I fragments
□ 5.7kb alphoid HindIII fragment
■ 5.7kb + 4.1kb + 1.9kb alphoid HindIII fragments
⊠ poxY1 BglII fragment > 1300kb
✖ poxY1 BglII fragment ≤ 1300kb
Number of Individuals
10
5
200 to 245
250 to 295
300 to 345
350 to 395
400 to 445
450 to 495
500 to 545
550 to 595
600 to 645
650 to 695
700 to 745
750 to 795
800 to 845
850 to 895
900 to 945
950 to 995
1000 to 1045
Size in kb

Figure 5.4

5.2 Structure of the Y alphoid DNA on abnormal Y chromosomes

The structure of alphoid DNA was analysed on abnormal Y chromosomes in order to address a number of questions. How close is alphoid DNA to the centromere? Is DNA structure altered in a suppressed centromere? Is alphoid DNA the functioning human centromere sequence?

5.2.1 Association of alphoid DNA with the centromere

25 cell lines reported to contain abnormal human Y centromeres were examined. Useful cytogenetic descriptions were provided with 23 of the cell lines. Of these, 18 were expected to have a Y chromosome centromere, it was not clear whether or not RW contained a Y centromere, and four lacked a centromere (figure 4.1). The four translocation chromosomes which lacked Y centromeres (GM02103, GM02969, GM00118 and GM07970) also lacked Y alphoid DNA. Six of the other 18 cell lines lacked Y alphoid DNA as well. Taken in isolation, this result appears to suggest that it is possible for a Y chromosome to have a functioning centromere in the absence of Y alphoid DNA. However, further investigation showed that all six of these cell lines had lost (or reduced to undetectable levels) the Y chromosome (Tyler-Smith, personal communication). Thus no cell line was found which contained a Y centromere which did not contain alphoid DNA.

This finding is in agreement with most of the results reported in the literature. Only one apparent exception has been described. This is RW [Affara *et al.* 1987]. RW was reported to contain a Y-derived dot chromosome which was moderately stable (and thus must have a centromere) but did not have alphoid DNA (probe GMG Y4a). Efforts were made to examine this particularly interesting chromosome in more detail. However, these efforts were unsuccessful because the RW cells received had a 45,X karyotype and had lost the Y-derived chromosome in >99% of cells. There would seem to be three possible explanations for the data of [Affara *et al.* 1987]. 1. The RW Y chromosome has a Y centromere but no Y alphoid DNA because Y alphoid DNA is not required for centromere function. 2. The chromosome has a Y centromere and Y alphoid DNA, but the Y alphoid DNA was not detected by [Affara *et al.* 1987]. 3. The chromosome, which contains at least two non-contiguous Y fragments, actually has a more complex structure and contains a centromere from another chromosome. Further experiments on Y-containing RW cells would be required to distinguish between these explanations.

The long-range structure of the alphoid DNA on 13 of the 15 abnormal Y chromosomes examined was considered to be normal. These structures were either indistinguishable from the structures found in normal chromosomes, or had only slight differences which could be understood as differences in block size and the presence or absence of additional individual restriction enzyme sites. Two Y chro-

mosomes (842968 and CHoP) had alphoid DNA structures which were considered to be significantly different from anything seen in normal individuals. The simplest explanation for these structures is that the rearrangement generating the abnormal chromosome had occurred within the alphoid DNA. The monocentric iso Yp chromosome 842968 is expected to be a giant palindromic DNA structure in which every DNA sequence except the centre of the palindrome is represented twice. Therefore the observation that 842968 has two copies of the Y190 block and only one copy of the alphoid block (table 4.3) is consistent with the alphoid DNA being at the centre of the palindrome. The restriction fragment sizes (figure 4.1 and table 4.2) are also consistent with a palindromic structure formed from two two left-hand halves of a typical map (figure 3.2) but are not consistent with a palindromic structure formed from two right-hand halves of a typical alphoid DNA map. Thus all the molecular and cytogenetic data could be explained by a large palindromic structure centred on the alphoid DNA. Other explanations are possible but are more complex. For example, a single fragment of alphoid DNA flanked by *Bgl*II *Pvu*II, *Ava*II, *Acc*I, *EcoO109*I, *Bcl*I and *Apa*I sites could have been formed in one arm of an asymmetric structure.

The alphoid DNA structure observed in the Y/18 translocation chromosome CHoP can also be understood if the junction between the two chromosomes is within the alphoid DNA. If the 18 DNA at the junction was typical euchromatic DNA, sites for the seven enzymes used in mapping would occur frequently. Replacement of the right-hand side of a typical large Y alphoid map (figure 3.5) with such a piece of DNA could generate the observed fragment sizes. Again, alternative explanations are possible, but are more complex. The suggested interpretations for the structures of these two alphoid blocks both predict an orientation for the alphoid map: that the short arm of the chromosome is on the left-hand side. Thus the two suggested interpretations are consistent with one another. If this interpretation of the 842968 isochromosome structure is correct and the chromosome is truly monocentric, it places the alphoid DNA precisely at the centromere and suggests that alphoid DNA could be the centromeric sequence. However, such a provocative conclusion must be treated with scepticism. It is possible that the structure of the isochromosome is more complex; alternatively, the isochromosome might be dicentric with two centromeres very close together which cannot be distinguished cytogenetically. It would be particularly interesting to examine a monocentric iso Yq chromosome. Unfortunately, no such chromosome has been reliably described: all iso Yq chromosomes examined in detail have been dicentric.

5.2.2 Alphoid DNA structure at inactive centromeres

Dicentric chromosomes are expected to be highly unstable since there is a chance at each cell division that the two centromeres will be pulled in opposite directions by the spindle apparatus. In *S.cerevisiae*, such chromosomes have been artificially constructed [Jager and Philippsen 1989] and are indeed unstable. Stabilisation can

occur by one of two mechanisms: the chromosome may break between the centromeres and new telomeres may be added; alternatively, one centromere may be deleted. In mammalian cells, dicentric or multicentric chromosomes have been observed. Such chromosomes are stable. This is because all centromeres except one have usually been inactivated or suppressed. Suppressed centromeres may appear cytogenetically to have a less marked constriction and may show reduced levels of binding to antikinetochore antisera [Zinkowski *et al.* 1986]. However, molecular analyses of the DNA sequences have not been reported. The results in section 4.2 showed that a normal alphoid DNA structure was present at the suppressed Y centromere of 1491/76R. The results in table 4.3 show that all the dicentric isochromosomes examined have two Y alphoid DNA blocks, and figure 4.1 shows that these are both normal in structure. Furthermore, section 5.2.2 shows that methylation patterns are not significantly altered. These results can be interpreted in two ways. Alphoid DNA could be distant from the centromere; centromere inactivation could then involve deletion of the real centromere sequences (as in *S.cerevisiae*) without affecting the alphoid DNA. Alternatively, alphoid DNA could be very close to or at the centromere, and centromere inactivation could occur by a different mechanism from that in *S.cerevisiae*. For example, protein-protein interactions might be disrupted. At present, these two explanations are difficult to distinguish.

5.2.3 Is alphoid DNA the human centromere?

Alphoid DNA is found by *in situ* hybridisation, close to all human centromeres [Mitchell *et al.* 1985] and thus is a candidate sequence for the functional centromeric DNA. The results presented in this thesis, particularly the observations in the monocentric iso Yp cell line, support the idea that it is very close to the centromere. Since the mammalian kinetochore appears in ultrastructural studies [Comings and Okada 1971] to be a large structure, it might be expected that the functional centromeric DNA sequence would be a moderately repeated sequence, as alphoid DNA is. However, it might also be expected that the centromeric sequence would be highly conserved between different mammals or even over a wider range of eukaryotes. Alphoid DNA is not highly conserved, and is not found in mammals other than primates. For example, mouse cells have no detectable alphoid DNA [Pritchard and Goodfellow 1986]. Human centromeres function quite well in mouse cells, and *vice versa*. Thus if alphoid DNA was the centromere sequence in human cells, it would be necessary to suggest that some feature other than primary DNA sequence (as detected by hybridisation) is responsible for function. Such a feature might, for example, be secondary or tertiary structure. It seems likely that the question of whether alphoid DNA is the functional centromere sequence will not be resolved until a functional assay for centromere activity is developed. Nevertheless, alphoid DNA would be the best candidate sequence so far identified to test such an assay.

Bibliography

[Affara *et al.* 1986] NA AFFARA, N FLORENTIN, N MORRISON, K KWOT, M MITCHELL, A COOK, D JAMIESON, L GLASGOW, L MEREDITH, E BOYD AND MA FERGUSON-SMITH. Regional assignment of Y-linked DNA probes by deletion mapping and their homology with X-chromosome and autosomal sequences. *Nucleic Acids Research* **14**: 5353–5373, 1986.

[Affara *et al.* 1987] NA AFFARA, MA FERGUSON-SMITH, RE MAGENIS, JL TOLMIE, E BOYD, A COOKE, D JAMIESON, K KWOK, M MITCHELL AND L SNADDEN. Mapping the testis determinants by an analysis of Y-specific sequences in males with apparent XX and XO karyotypes and females with XY karyotypes. *Nucleic Acids Research* **15**: 7325–7342, 1987.

[Agresti *et al.* 1989] A AGRESTI, R MENEVERI, AG SICCARDI, A MAROZZI, G CORNEO, S GAUDI AND E GINELLI. Linkage in human heterochromatin between highly divergent Sau3A repeats and a new family of repeated DNA sequences (HaeIII family). *Journal of Molecular Biology* **205**: 625–631, 1989.

[Anand 1986] R ANAND. Pulsed field gel electrophoresis: a technique for fractionating large DNA molecules. *Trends in Genetics* **12**: 278–283, 1986.

[Bird 1987] AP BIRD. CpG islands as gene markers in the vertebrate nucleus. *Trends in Genetics* **3**: 342–347, 1987.

[Bloom *et al.* 1984] KS BLOOM, K AMAYA, J CARBON, L CLARKE, A HILL AND E YEH. Chromatin conformation of yeast centromeres. *The Journal of Cell Biology* **99**: 1559–1568, 1984.

[Borgaonkar *et al.* 1974] DS BORGAONKAR, BM SROKA AND M FLORES. Y-to-X translocation in a girl. *The Lancet* **1**: 68–69, 1974.

[Britten and Kohne 1968] RJ BRITTEN AND DE KOHNE. Repeated sequences in DNA. *Science* **161**: 529–540, 1968.

[Buckle *et al.* 1985] V BUCKLE, C MONDELLO, S DARLING, IW CRAIG AND PN GOODFELLOW. Homologous expressed genes in the human sex chromosome pairing region. *Nature* **317**: 739–741, 1985.

[Burk *et al.* 1985] RD BURK, P MA AND KD SMITH. Characterization and evolution of a single-copy sequence from the human Y chromosome. *Molecular and Cell Biology* **5**: 576–581, 1985.

[Burke *et al.* 1987] DT BURKE, GF CARLE AND MV OLSON. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**: 806–812, 1987.

[Cann *et al.* 1987] RL CANN, M STONEKING AND AC WILSON. Mitochondrial DNA and human evolution. *Nature* **235**: 31–36, 1987.

[Carle and Olson 1985] GF CARLE AND MV OLSON. An electrophoretic karyotype for yeast. *Proceedings of the National Academy of Sciences of the USA* **82**: 3756–3760, 1985.

[Casanova *et al.* 1985] M CASANOVA, P LEROY, C BOUCEKKINE, J WEISSENBACH, C BISHOP, M FELLOUS, M PURRELLO, G FIORI AND M SINISCALCO. A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science* **230**: 1403–1406, 1985.

[Chandley *et al.* 1986] AC CHANDLEY, P AMBROS, S MCBEATH, TB HARGREAVE, F KILANOWSKI AND G SPOWART. Short arm dicentric Y chromosome with associated statural defects in a sterile man. *Human Genetics* **73**: 350–353, 1986.

[Chikashige *et al.* 1989] Y CHIKASHIGE, N KIOSHITA, Y NAKASEKO, T MATSUMOTO, S MURAKAMI, O NIWA AND M YANAGIDA. Composite motifs and repeat symmetry in *S. pombe* centromeres: direct analysis by intergration of NotI restriction sites. *Cell* **57**: 739–751, 1989.

[Choo *et al.* 1988] KH CHOO, B VISSEL, R BROWN, RG FILBY AND E EARLE. Homologous alpha satellite sequences on human acrocentric chromosomes with selectivity for chromosomes 13, 14 and 21: implications for recombination between nonhomologues and Robertsonian translocations. *Nucleic Acids Research* **16**: 1273–1284, 1988.

[Choo *et al.* 1989] KH CHOO, B VISSEL AND E EARLE. Evolution of alpha-satellite DNA on human acrocentric chromosomes. *Genomics* **5**: 332–344, 1989.

[Clarke and Carbon 1980] L CLARKE AND J CARBON. Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* **287**: 504–509, 1980.

[Clarke and Carbon 1983] L CLARKE AND J CARBON. Genomic substitutions of centromeres in *Saccharomyces cerevisiae*. *Nature* **305**: 23–28, 1983.

[Comings and Okada 1971] DE COMINGS AND TA OKADA. Fine structure of kinetochore in Indian muntjac. *Experimental Cell Research* **67**: 97–110, 1971.

[Devilee *et al.* 1986] P DEVILEE, P SLAGBOOM, CJ CORNLISSE AND PL PEARSON. Sequence heterogeneity within the human alphoid repetitive DNA family. *Nucleic Acids Research* **14**: 2059–2073, 1986.

[Doolittle and Sapienza 1980] WF DOOLITTLE AND C SAPIENZA. Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**: 601–603, 1980.

[Fan *et al.* 1988] J-B FAN, Y CHIKASHIGE, CL SMITH, O NIWA, M YANAGIDA AND CR CANTOR. Construction of a Not 1 restriction map of the fission yeast *Schizosaccharomyces pombe* genome. *Nucleic Acids Research* **17**: 2801–2818, 1988.

[Feinberg and Vogelstein 1983] AP FEINBERG AND B VOGELSTEIN. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**: 6–13, 1983.

[Feinberg and Vogelstein 1984] AP FEINBERG AND B VOGELSTEIN. Addendum: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **137**: 266–267, 1984.

[Fishel *et al.* 1988] B FISHEL, H AMSTUTZ, M BAUM, J CARBON AND L CLARKE. Structural organization and functional analysis of centromeric DNA in the fission yeast *Schizosaccharomyces pombe*. *Molecular and Cellular Biology* **8**: 754–763, 1988.

[Fitzgerald-Hayes *et al.* 1982a] M FITZGERALD-HAYES, J-M BUHLER, TG COOPER AND J CARBON. Isolation and subcloning analysis of functional centromere DNA CEN11 from *Saccharomyces cerevisiae* chromosome XI. *Molecular and Cellular Biology* **2**: 82–87, 1982.

[Fitzgerald-Hayes *et al.* 1982b] M FITZGERALD-HAYES, L CLARKE AND J CARBON. Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. *Cell* **29**: 235–244, 1982.

[Frommer *et al.* 1988] M FROMMER, C PAUL AND PC VINCENT. Localisation of satellite DNA sequences on human metaphase chromosomes using bromodeoxyuridine-labelled probes. *Chromosoma* **97**: 11–18, 1988.

[Goodfellow *et al.* 1983] P GOODFELLOW, G BANTING, D SHEER, HH ROPERS, A CAINE, MA FERGUSON-SMITH, S POVEY AND R VOSS. Genetic evidence that a Y-linked gene in man is homologous to a gene on the X chromosome. *Nature* **302**: 346–349, 1983.

[Gosden and Porteous 1987] JR GOSDEN AND DJ PORTEOUS. HRAS-1-selected chromosome mediated gene transfer, *in situ* hybridisation with combined biotin and tritium label localised the oncogene and reveals duplications of the human transgenome. *Cytogenetics and Cell Genetics* **45**: 44–45, 1987.

[Gosden *et al.* 1975] JR GOSDEN, AR MITCHELL, RA BUCKLAND, RP CLAYTON AND HJ EVANS. The location of four human satellite DNAs on human chromosomes. *Experimental Cell Research* **92**: 148–158, 1975.

[Hahnemann *et al.* 1985] N HAHNEMANN, RC MILLER, AE GREENE AND LL CORIELL. A (Y;15) translocation, 46 chromosomes. *Cytogenetics and Cell Genetics* **15**: 408–409, 1985.

[Hahnenberger *et al.* 1989] KM HAHNENBERGER, MP BAUM, CM POLIZZI, J CARBON AND L CLARKE. Construction of functional artificial minichromosomes in the fission yeast *Schizosaccharomyces pombe*. *Proceedings of the National Academy of Sciences of the USA* **86**: 577–581, 1989.

[Harris *et al.* 1986] P HARRIS, E BOYD, BE YOUNG AND MA FERGUSON-SMITH. Determination of the DNA content of human chromosomes by flow cytometry. *Cytogenet Cell Genet* **41**: 14–21, 1986.

[Hegemann *et al.* 1986] JH HEGEMANN, RD PRIDMORE, R SCHNEIDER AND P PHILIPPSEN. Mutations in the right boundary of *Saccharomyces cerevisiae* centromere 6 lead to nonfunctional or partially functional centromeres. *Molecular and General Genetics* **205**: 305–311, 1986.

[Hieter *et al.* 1985] P HIETER, D PRIDMORE, JH HEGEMANN, M THOMAS, RW DAVIS AND P PHILIPPSEN. Functional selection and analysis of yeast centromeric DNA. *Cell* **42**: 913–921, 1985.

[Hsiao and Carbon 1979] C-L HSIAO AND J CARBON. High-frequency transformation of yeast by plasmids containing the cloned yeast ARG4 gene. *Proceedings of the National Academy of Sciences of the USA* **76**: 3829–3833, 1979.

[Huberman *et al.* 1986] JA HUBERMAN, RD PRIDMORE, D JAGER, B ZONNEVELD AND P PHILIPPSEN. Centromeric DNA from *Saccharomyces uvarum* is functional in *Saccharomyces cerevisiae*. *Chromosoma* **94**: 162–168, 1986.

[Jabs and Perisco 1987] EW JABS AND MG PERISCO. Characterization of human centromeric regions of specific chromosomes by means of alphoid DNA sequences. *American Journal of Human Genetics* **41**: 374–390, 1987.

[Jager and Philippsen 1989] D JAGER AND P PHILIPPSEN. Stabilization of dicentric chromosome in *Saccharomyces cerevisiae* by telomere addition to broken ends or by centromere deletion. *EMBO Journal* **8**: 247–254, 1989.

[John and Miklos 1979] B JOHN AND G MIKLOS. Functional aspects of satellite DNA and heterochromatin. *International Review of Cytology* **58**: 1–114, 1979.

[Jorgensen *et al.* 1988] AL JORGENSEN, S KOLVRAA, C JONES AND AL BAK. A subfamily of alphoid repetitive DNA shared by the NOR-bearing human chromosomes 14 and 22. *Genomics* **3**: 100–109, 1988.

[Kaufer *et al.* 1985] NF KAUFER, V SIMANIS AND P NURSE. Fission yeast *Schizosaccharomyces pombe* correctly excises a mammalian RNA transcript intervening sequence. *Nature* **318**: 78–80, 1985.

[Kit 1961] S KIT. Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. *Journal of Molecular Biology* **3**: 711–716, 1961.

[Klobutcher and Ruddle 1979] LA KLOBUTCHER AND FH RUDDLE. Phenotype stabilization and integration of transferred material in chromosome mediated gene transfer. *Nature* **280**: 657–660, 1979.

[Lica *et al.* 1986] LM LICA, S NARAYANSWAMI AND BA HAMKALO. Mouse satellite DNA, centromere structure, and sister chromatid pairing. *The journal of Cell Biology* **103**: 1145–1151, 1986.

[Magenis *et al.* 1985] RE MAGENIS, ME BROWN, T DONLON, S OLSON, R SHEEHY AND TOMAR D. Structural aberrations of the Y chromosomes, including the non-fluorescent Y: cytologic origin and consequences. In SANDBERG, editor, *The Y chromosomes, part A; Basic characteristics of the Y chromosomes*, pages 537–574. New York: Alan R. Liss, 1985.

[Maio 1971] JJ MAIO. DNA strand reassociation and polyribonucleotide binding in the African green monkey, *Cercopithecus aethiops*. *Journal of Molecular Biology* **56**: 579–595, 1971.

[Maniatis *et al.* 1982] T MANIATIS, EF FRITSCH AND J SAMBROOK. A laboratory manual. In *Molecular cloning*. New York: Cold Spring Harbor, 1982.

[Marcus *et al.* 1976] M MARCUS, R TANTRAVAHI, VG DEV, DA MILLER AND MILLER OJ. Human-mouse cell hybrid with human multiple Y chromosomes. *Nature* **262**: 63–65, 1976.

[Maserati *et al.* 1986] E MASERATI, F WAIBEL, B WEBER, M FRACCARO, A GAL, F PASQUALI, W SCHEMPP, G SCHERER, R VACCARO, J WEISSENBACH AND U WOLF. A 45,X male with a Yp/18 translocation. *Human Genetics* **74**: 126–132, 1986.

[McDermid *et al.* 1986] HE MCDERMID, AMV DUNCAN, JL HIGGINS, MJ HAMERTON, E RECTOR, KR BRASCH AND BN WHITE. Isolation and

characterization of an alpha-satellite repeated sequence from human chromosome 22. *Chromosoma* **94**: 228–234, 1986.

[Meneveri *et al.* 1985] R MENEVERI, A AGRESTI, GD VALLE, D TALARICO, AG SICCARDI AND E GINELLI. Identification of a human clustered G+C-rich DNA family of repeats (Sau3A family). *Journal of Molecular Biology* **186**: 483–489, 1985.

[Mitchell *et al.* 1985] AR MITCHELL, JR GOSDEN AND DA MILLER. A cloned sequence, p82H of the alphoid repeated DNA family found at the centromeres of all human chromosomes. *Chromosoma* **92**: 369–377, 1985.

[Moroi *et al.* 1980] Y MOROI, C PEEBLES, MJ FRITZLER, J STEIGERWALD AND EM TAN. Autoantibody to centromere (kinetochore) in scleroderma sera. *Proceedings of the Academy of Sciences of the USA* **77**: 1627–1631, 1980.

[Muller *et al.* 1986] U MULLER, M LALANDE, CM DISTECHE AND SA LATT. Construction, analysis, and application to 46, XY gonadal dysgenesis of a recombinant phage DNA library from flow-sorted human Y chromosome. *Cytometry* **7**: 418–424, 1986.

[Munke *et al.* 1988] M MUNKE, DC PAGE, LG BROWN, BA ARMSON, EH ZACKAI, MT MENNUTI AND BS EMANUEL. Molecular detection of a Yp/18 translocation in a 45,X holoprosencephalic male. *Human Genetics* **80**: 219–223, 1988.

[Murray and Szostak 1983] AW MURRAY AND JW SZOSTAK. Construction of artificial chromosomes in yeast. *Nature* **305**: 189–193, 1983.

[Nakaseko *et al.* 1986] Y NAKASEKO, Y ADACHI, S FUNAHASHI, O NIWA AND M YANAGIDA. Chromosome walking shows a highly homologous repetitive sequence present in all the centromere regions of fission yeast. *EMBO Journal* **5**: 1011–1021, 1986.

[Nakaseko *et al.* 1987] Y NAKASEKO, N KINOSHITA AND YANAGIDA M. A novel sequence common to the centromere regions of *Schizosaccharomyces pombe* chromosomes. *Nucleic Acids Research* **15**: 4705–4716, 1987.

[NIGMS 1986 1987] NIGMS. Catalogue of cell lines. In *Human genetic mutant cell repository.* Coriell Institute for Medical Research: National Institutes of Health, 1986–1987.

[Niwa *et al.* 1989] O NIWA, T MATSUMOTO, Y CHIKASHIGE AND YANAGIDA M. Characterization of *Schizosaccharomyces pombe* minichromosome deletion derivatives and functional allocation of their centromere. *EMBO Journal* **8**: 3045–3052, 1989.

[Orgel and Crick 1980] LE ORGEL AND FHC CRICK. Selfish DNA: the ultimate parasite. *Nature* **284**: 604–607, 1980.

[Page *et al.* 1987] DC PAGE, R MOSHER, EM SIMPSON, EMC FISHER, G MARDON, J POLLACK, B MCGILLIVRAY, DE LA CHAPELLE AND LG BROWN. The sex-determining region of the human Y chromosome encodes a finger protein. *Cell* **51**: 1091–1104, 1987.

[Page 1986] DC PAGE. Sex reversal: deletion mapping the male-determining function of the human y chromosome. *Cold Spring Harbor Symposia on Quantitative Biology* **LI**: 229–235, 1986.

[Panzeri and Philippsen 1982] L PANZERI AND P PHILIPPSEN. Centromeric DNA from chromosome VI in *Saccharomyces cerevisiae* strains. *EMBO journal* **1**: 1605–1611, 1982.

[Paulson *et al.* 1985] KE PAULSON, N DEKA, CW SCHMID, R MISRA, CW SCHINDLER, MG RUSH, L KADYK AND L LEINWAND. A transposon-like element in human DNA. *Nature* **316**: 359–361, 1985.

[Porteous 1987] DJ PORTEOUS. Chromosomes mediated gene transfer: a functional assay for complex loci and an aid to human genome mapping. *Trends in Genetics* **3**: 177–182, 1987.

[Pritchard and Goodfellow 1986] CA PRITCHARD AND PN GOODFELLOW. Development of new methods in human gene mapping, selection for fragments of the human Y chromosome after chromosome-mediated gene transfer. *EMBO Journal* **5**: 979–985, 1986.

[Pritchard and Goodfellow 1987] CA PRITCHARD AND PN GOODFELLOW. Genetic analysis of the human Y chromosome by chromosome-mediated gene transfer. *Development Supplement* **101**: 59–65, 1987.

[Prosser *et al.* 1986] J PROSSER, M FROMMER, C PAUL AND PC VINCENT. Sequence relationships of three human satellite DNAs. *Journal of Molecular Biology* **187**: 145–155, 1986.

[Riddell *et al.* 1986] DC RIDDELL, HS WANG, J BECKETT, A CHAN, JJA HOLDEN, LM MULLIGAN, MA PHILLIPS, NE SIMPSON, K WROGEMANN, JL HAMERTON AND BN WHITE. Regional localization of 18 human X-linked DNA sequences. *Cytogenetics and Cell Genetics* **42**: 123–128, 1986.

[Rothstein 1983] RD ROTHSTEIN. One-step gene disruption in yeast. *Methods in Enzymology* **101**: 202–211, 1983.

[Schwartz and Cantor 1984] DC SCHWARTZ AND CR CANTOR. Separation of yeast chromosomes-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* **37**: 67–75, 1984.

[Simpson *et al.* 1987] E SIMPSON, P CHANDLER, E GOULMY, CM DISTECHE, A FERGUSON-SMITH AND DC PAGE. Separation of the genetic loci for the H-Y antigen and for testis determination of human Y chromosome. *Nature* **326**: 876–878, 1987.

[Singer 1982] MF SINGER. Highly repeated sequences in mammalian genomes. *International Review of Cytology* **76**: 67–112, 1982.

[Sirota *et al.* 1981] L SIROTA, Y ZLOTOGORA, F SHABTAI, I HALBRECHT AND E ELIAN. 49,XYYYY. A case report. *Clinical Genetics* **19**: 87–93, 1981.

[Skowronski and Singer 1986] J SKOWRONSKI AND MF SINGER. The abundant line-1 family of repeated DNA sequence in mammals: genes and pseudogenes. *Cold Spring Harbor, Symposia on Quantitative Biology* **LI**: 457–464, 1986.

[Smith and Cantor 1987] CL SMITH AND CR CANTOR. Preparation and manipulation of large DNA molecules: advances and applications. *Trends in Biological Sciences* **12**: 284–287, 1987.

[Smith 1973] GP SMITH. Unequal crossover and the evolution of multigene families. *Cold Spring Harbor Symposia on Quantitative Biology* **38**: 507–513, 1973.

[Smith 1976] GP SMITH. Evolution of repeated DNA sequences by unequal crossover. *Science* **191**: 528–535, 1976.

[Southern *et al.* 1987] EM SOUTHERN, R ANAND, WRA BROWN AND DS FLETCHER. A model for the separation of large DNA molecules by crossed field gel electrophoresis. *Nucleic Acids Research* **15**: 5925–5943, 1987.

[Southern 1970] EM SOUTHERN. Base sequence and evolution of guinea-pig alpha-satellite DNA. *Nature* **227**: 794–798, 1970.

[Southern 1975a] EM SOUTHERN. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**: 503–518, 1975.

[Southern 1975b] EM SOUTHERN. Long range periodicities in mouse satellite DNA. *Journal of Molecular Biology* **94**: 51–69, 1975.

[Stinchcomb *et al.* 1979] DT STINCHCOMB, K STRUHL AND DAVIS RW. Isolation and characterisation of a yeast chromosomal replicator. *Nature* **282**: 39–43, 1979.

[Stoneking *et al.* 1986] M STONEKING, K BHATIA AND AC WILSON. Rate of sequence divergence estimated from restriction maps of mitochondrial DNAs from papua new guinea. *Cold Spring Harbor Symposia on Quantitative Biology* **LI**: 433–439, 1986.

[Sueoka 1961] N SUEOKA. Variation and heterogeneity of base composition of deoxyribonucleic acids: a compilation of old and new data. *Journal of Molecular Biology* **3**: 31–40, 1961.

[Thomas 1973] CA THOMAS. The rolling helix: a model for the eukaryotic gene? *Cold Spring Harbor Symposia on Quantitative Biology* **38**: 347–352, 1973.

[Tyler-Smith and Brown 1987] C TYLER-SMITH AND WRA BROWN. Structure of the major block of alphoid satellite DNA on the human Y chromosome. *Journal of Molecular Biology* **195**: 457–470, 1987.

[Tyler-Smith *et al.* 1988] C TYLER-SMITH, L TAYLOR AND U MÜLLER. Structure of a hypervariable tandemly repeated DNA sequence on the short arm of the human Y chromosome. *Journal of Molecular Biology* **203**: 837–848, 1988.

[Tyler-Smith 1987] C TYLER-SMITH. Structure of repeated sequences in the centromeric region of the human Y chromosome. *Development Supplement* **101**: 93–100, 1987.

[Wahl *et al.* 1979] GM WAHL, M STERN AND GR STARK. Efficient transfer for large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulfate. *Proceedings of the National Academy of Sciences of the USA* **8**: 3683–3687, 1979.

[Waring and Britten 1966] M WARING AND RJ BRITTEN. Nucleotide sequence repetition: a rapidly reassociating fraction of mouse DNA. *Science* **154**: 791–795, 1966.

[Waye and Willard 1986] JS WAYE AND HF WILLARD. Structure, organization, and sequence of alpha satellite DNA from human chromosome 17: evidence for evolution by unequal crossing-over and an ancestral pentamer repeat shared with the human X chromosome. *Molecular and Cellular Biology* **6**: 3156–3165, 1986.

[Waye and Willard 1989] JS WAYE AND HF WILLARD. Chromosome specificity of satellite DNAs: short-and long-range organization of a diverged dimeric subset of human alpha satellite from chromosome 3. *Chromosoma* **97**: 475–480, 1989.

[Waye *et al.* 1987a] JS WAYE, LA CREEPER AND HF WILLARD. Organization and evolution of alpha satellite DNA from human chromosome 11. *Chromosoma* **95**: 182–188, 1987.

[Waye *et al.* 1987b] JS WAYE, SB ENGLAND AND HF WILLARD. Genomic organization of alpha satellite DNA on human chromosome 7: evidence for two distinct alphoid domains on a single chromosome. *Molecular and Cellular Biology* **7**: 349–356, 1987.

[Willard 1985] HF WILLARD. Chromosome-specific organization of human alpha satellite DNA. *American Journal of Human Genetics* **37**: 524–532, 1985.

[Wolfe *et al.* 1985] J WOLFE, SM DARLING, RP ERICKSON, IW CRAIG, VJ BUCKLE, PWJ RIGBY, HF WILLARD AND PN GOODFELLOW. Isolation and characterization of an alphoid centromeric repeat family from the human Y chromosome. *Journal of Molecular Biology* **182**: 477–485, 1985.

[Wu and Manuelidis 1978] J WU AND L MANUELIDIS. Homology between human and simian repeated DNA. *Nature* **276**: 92–94, 1978.

[Wu and Manuelidis 1980] J WU AND L MANUELIDIS. Sequence definition and organization of a human repeated DNA. *Journal of Molecular Biology* **142**: 363–386, 1980.

[Wu *et al.* 1989] CI WU, JR TRUE AND N JOHNSON. Fitness reduction associated with the deletion of a satellite DNA array. *Nature* **341**: 248–251, 1989.

[Yang *et al.* 1982] TP YANG, SK HANSEN, KK OISHI, OA RYDER AND BA HAMKALO. Characterization of a cloned repetitive DNA sequence concentrated on the human X chromosome. *Proceedings of the National Academy of Sciences of the USA* **79**: 6593–6597, 1982.

[Zinkowski *et al.* 1986] RP ZINKOWSKI, BK VIG AND D BROCCOLI. Characterisation of kinetochores in multicentric chromosomes. *Chromosoma* **94**: 243–248, 1986.