Human Molecular Genetics, 1995, Vol. 4, No. 6 1007–1014

YAC cloning *Mus musculus* telomeric DNA: physical, genetic, *in situ* and STS markers for the distal telomere of chromosome 10

David Kipling*, Helen E.Wilson, Eric J.Thomson and Howard J.Cooke

Medical Research Council, Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK

Received June 27, 1994; Revised and Accepted March 14, 1995

EMBL accession no Z33505

Three *Mus musculus* DBA/2 YAC libraries were constructed using a half-YAC telomere cloning vector. This functional complementation approach yields libraries which include terminal restriction fragments of the mouse genome. Screening all three libraries led to the isolation of 32 independent clones which carry linear YACs containing the mouse terminal repeat sequence, $(TTAGGG)_n$. These YACs provide a resource to isolate regions of the mouse genome close to chromosome termini and excluded from existing conventional YAC libraries. To demonstrate their utility, a hybridization probe was isolated from Mtel-1, the first $(TTAGGG)_n$-containing YAC isolated. This probe detects a ~70 kb *Kpn*I fragment in the mouse genome which is sensitive to pretreatment with BAL31 exonuclease. A PCR-based genetic marker generated from the sequence of this probe maps 4.4 cM from the most distal anchor locus on chromosome 10 in the EUCIB interspecific backcross. STS primers for this locus, *D10Hgu1*, were used to isolate YAC 110F4 from a commercially available mouse YAC library. Fluorescence *in situ* hybridization demonstrates that YAC 110F4 hybridizes to the distal telomere of chromosome 10. Clones in this collection of telomere YACs therefore partially overlap clones in conventional YAC libraries, and thus the previously unavailable terminal regions of the mouse genome can now be linked with the developing mouse STS YAC contig. Genetic markers such as *D10Hgu1* allow the ends of the mouse genetic map to be defined, thus closing the map.

INTRODUCTION

Telomeres are specialized DNA-nucleoprotein structures at the ends of linear chromosomes. Their biological roles include protecting chromosome ends from nuclease attack, distinguishing natural chromosome termini from double strand breaks (thus preventing them from activating cell cycle arrest checkpoints), and providing a substrate for *de novo* DNA synthesis by telomerase to compensate for losses because of incomplete replication (1–3). Small stretches of DNA from human telomeres have been cloned in *Escherichia coli* either directly or utilizing a PCR-based approach, and larger clones have been obtained by telomere cloning in yeast (4). This has led to the telomeres of a number of human chromosomes being placed on the physical and genetic maps (5).

Mouse telomeres are less well characterized than those of humans. *Mus musculus* chromosomes terminate in arrays of $(TTAGGG)_n$ as shown both by *in situ* hybridization and BAL31 exonuclease sensitivity (6–8). Mouse terminal repeat arrays can be five to 10 times larger than those of humans. For example, in the inbred strain DBA/2 the $(TTAGGG)_n$ repeats are present on terminal restriction fragments 20–150 kb in size which can be visualized as discrete restriction fragments using pulsed field gel electrophoresis (PFGE) (7,8). One striking feature of these $(TTAGGG)_n$ arrays is their variability in size between individuals, even within an inbred strain, and new size alleles have been observed frequently in mouse family studies (7,8). The biological mechanism underlying this unusual genetic instability is not understood, and because $(TTAGGG)_n$ is present at all mouse telomeres further analysis of this phenomenon would be greatly facilitated by hybridization probes specific for individual telomeres. However, no such chromosome-specific probes have yet been isolated in the mouse.

One major aim of the mouse genome mapping project is to provide a high resolution genetic map linked with an STS YAC contig (9). Because PCR is used to isolate regions of the mouse genome corresponding to mapped microsatellite markers (10), this approach relies on the availability of genomic libraries (e.g. YAC and P1) that can be screened by PCR. However, because of the nature of the cloning strategy used, the currently available YAC libraries cannot include terminal restriction fragments. To alleviate this problem we have produced three YAC libraries which include these terminal regions of the mouse genome. These libraries allowed the isolation and characterization of 32 $(TTAGGG)_n$-containing clones.

RESULTS

YAC library construction and isolation of $(TTAGGG)_n$-containing clones

Terminal restriction fragments many kilobases in size have been cloned from mammalian genomes by a modified YAC cloning protocol, using the vector pYAC4NEONOT which

*To whom correspondence should be addressed

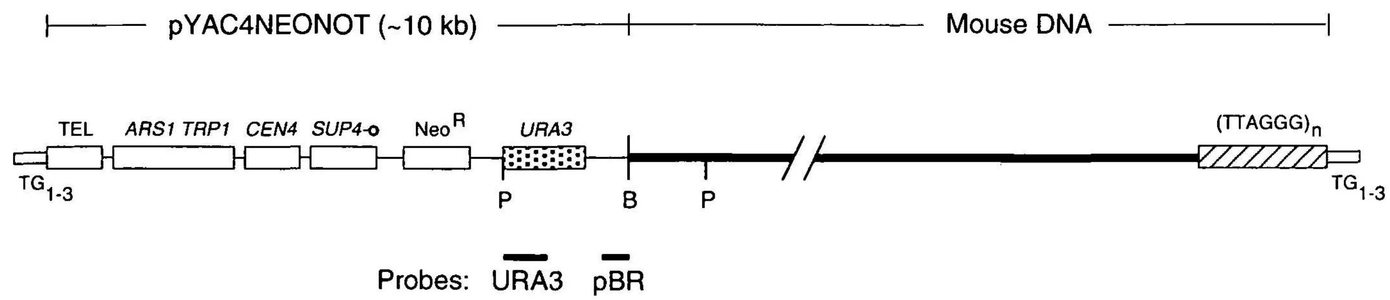

Figure 1. The structure of a typical Mtel YAC in yeast. The $(TTAGGG)_n$ of the mouse insert DNA is shown healed by the addition of yeast TG_{1-3} repeats to provide telomere function. P, *Pst*I sites used for subcloning vector-proximal insert into plasmid. B, *Bam*HI site at vector-insert boundary. Clones in library 3 have an additional *Not*I vector site immediately to the right of the *Bam*HI site. The location of the hybridization probes used for the *URA3* gene and pBR322 sequences are indicated (see Materials and Methods). Not drawn to scale.

provides only a single telomere (4). In order for the YAC to be stably maintained as a linear molecule in yeast the second telomere is provided by the ability of the mammalian telomeric repeat, $(TTAGGG)_n$, to prime the addition of yeast TG_{1-3} repeats, which in turn provide yeast telomere function (see Fig. 1).

Terminal restriction fragments of the mouse genome were enriched by size fractionation of male DBA/2 genomic DNA. Most terminal restriction fragments are >20 kb in size in this strain (7,8) and can be separated from much of the non-terminal fragments. Three YAC libraries were produced, each using a different size selection scheme (Materials and Methods). Genomic DNA was digested to completion with either *Bam*HI or *Not*I, enriched for telomeric restriction fragments by either conventional or PFGE agarose gel electrophoresis, and ligated to pYAC4NEONOT (11). Ura^+ colonies were picked following transformation into *Saccharomyces cerevisiae* and screened for the presence of $(TTAGGG)_n$ by colony hybridization. Thirty-four $(TTAGGG)_n$-positive independent transformants were isolated and designated Mtel-1 to Mtel-34. The fraction of Ura^+ transformants in each library containing $(TTAGGG)_n$ is 1/128 for library 1, 1/58 for library 2 and 1/173 for library 3.

It has previously been shown that in genomic digests with enzymes such as *Bgl*II or *Pvu*II the centromeric minor satellite is on the same terminal restriction fragment as the $(TTAGGG)_n$ array, and these centromere-proximal terminal restriction fragments can be over 1 Mb in size (12). In contrast, the centromere-distal telomeres are <150 kb in a similar digest. Given the large size of the centromere-proximal fragments, together with the possible difficulties of cloning large fragments of satellite DNA into YACs (13), it was likely that mouse telomere YACs made from digests such as *Bgl*II or *Pvu*II would be biased against centromere-proximal telomeres. In an attempt to avoid this problem a *Bam*HI digest was used as this enzyme cleaves between the minor satellite and the terminal repeat arrays, so that almost all the terminal restriction fragments from both proximal and distal telomeres are <150 kb.

Physical organization of mouse telomere YACs

Yeast chromosomes were prepared from all 34 $(TTAGGG)_n$-positive clones, resolved by PFGE, and Southern filters probed with $(TTAGGG)_4$ to determine the size of each YAC (Table 1). Mtel-8 is a false positive. The average insert size for the entire YAC collection is 37 kb, ranging from 21 kb for clones from library 1, 33 kb for clones from library 2, to 58 kb for clones from library 3 (Table 1). All YACs were negative by hybridization with either mouse major or minor satellite (data not shown). This is in agreement with the prediction from the genomic physical map that the minor satellite should be on non-terminal *Bam*HI fragments and thus not present in most clones in this collection (see above).

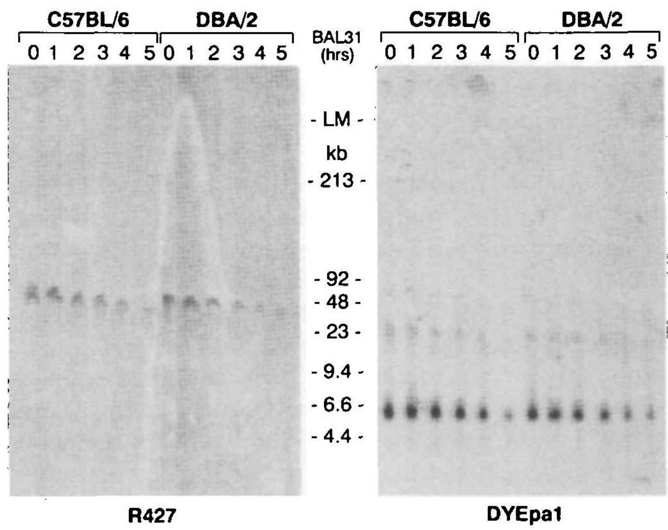

Figure 2. BAL31 sensitivity of R427. Genomic DNA from C57BL/6 and DBA/2 males was digested for increasing times with exonuclease BAL31 before digestion with *Kpn*I and separation by pulsed field gel electrophoresis. The Southern filter was probed with R427 (left) and then DYEpa1 (right) without intervening stripping; some R427 signal is therefore visible in the right panel and facilitates comparison. LM, Limit Mobility.

Southern filters of *Bam*HI digested DNA from each YAC strain were hybridized with a pBR322 probe to detect the pYAC4NEONOT vector arm. As illustrated in Figure 1, if the YAC is linear and unrearranged, such a digest should release a fragment of ~10 kb corresponding to the entire pYA-C4NEONOT vector arm (11). All YACs released a fragment of the predicted size except for Mtel-31 which was therefore

Table 1. Structural analysis of Mtel YAC collection. N/A, not applicable

YAC	Library[a]	Size (kb)[b]	$(TTAGGG)_n$	10 kb vector band in *Bam*HI digest	Maximum $(TTAGGG)_n$ array size (kb)[c]	R427	Locus
Mtel-1	1	46	+	+	0.5	+	Distal telomere of MMU 10
Mtel-2	1	18	+	+	0.6	-	nd
Mtel-3	1	32	+	+	0.6	-	nd
Mtel-4	1	150	+	+	0.5	-	nd
Mtel-5	1	18	+	+	0.5	-	nd
Mtel-6	1	56	+	+	0.5	-	nd
Mtel-7	1	30	+	+	1.1	-	nd
Mtel-8	1	N/A	-	+	N/A	-	nd
Mtel-9	1	18	+	+	0.5	-	nd
Mtel-10	2	40	+	+	0.8	-	nd
Mtel-11	2	30	+	+	1.2	-	nd
Mtel-12	2	60	+	+	1.1	-	nd
Mtel-13	2	60	+	+	0.7	-	nd
Mtel-14	2	48	+	+	1.5	-	Pseudoautosomal telomere
Mtel-15	2	36	+	+	0.5	-	nd
Mtel-16	2	42	+	+	1.0	-	nd
Mtel-17	2	22	+	+	1.9	-	Distal telomere of MMU 9 or MMU 13
Mtel-18	2	65	+	+	0.6	-	nd
Mtel-19	2	52	+	+	0.6	-	nd
Mtel-20	2	25	+	+	0.5	-	nd
Mtel-21	2	48	+	+	1.1	-	Pseudoautosomal telomere
Mtel-22	2	44	+	+	0.6	-	nd
Mtel-23	2	68	+	+	0.5	-	nd
Mtel-24	2	48	+	+	1.3	-	Pseudoautosomal telomere
Mtel-25	2	30	+	+	0.6	-	nd
Mtel-26	2	25	+	+	0.5	-	nd
Mtel-27	2	25	+	+	1.5	-	nd
Mtel-28	2	48	+	+	0.5	-	nd
Mtel-29	2	44	+	+	1.0	-	Pseudoautosomal telomere
Mtel-30	3	20	+	+	0.4	-	nd
Mtel-31	3	See text	+	-	0.8	-	nd
Mtel-32	3	80	+	+	0.6	-	nd
Mtel-33	3	20	+	+	2.6	-	Distal telomere of MMU 9 or MMU 13
Mtel-34	3	150	+	+	0.3	-	nd

[a]Details of the construction of libraries 1, 2 and 3 are given in Materials and Methods. [b]Uncut size as determined by PFGE, and includes 10 kb of vector sequence. [c]Smallest mean size of $(TTAGGG)_n$ signal following *Alu*I, *Hae*III or *Rsa*I digestion. nd, not determined.

excluded from further analysis (Table 1). Mtel-31 also did not migrate successfully using PFGE and may be a circular molecule (data not shown).

DNA from each YAC strain was analysed by Southern hybridization using a $(TTAGGG)_4$ probe following digestion with *Alu*I, *Hae*III, *Rsa*I, *Stu*I, *Eco*RV, *Pst*I, *Dra*I, *Sma*I and *Bam*HI and separation by conventional agarose gel electrophoresis. The distance to the first internal site from the terminal $(TTAGGG)_n$ provides a partial restriction map (Table 1 and data not shown). The maximum size of the $(TTAGGG)_n$ array in each YAC (Table 1) was based on the largest fragment following *Alu*I, *Hae*III or *Rsa*I digestion since these enzymes do not cleave within $(TTAGGG)_n$. In all cases the $(TTAGGG)_n$ array is less than 2.7 kb.

In summary, all of the remaining 32 YACs migrate successfully into PFGE gels, release the correct sized vector fragment in a *Bam*HI digest, and many show a heterogeneous $(TTAGGG)_n$ signal in some digests (data not shown). This would be consistent with each YAC being linear with the $(TTAGGG)_n$ array close to one telomere (see Fig. 1).

Isolation of chromosome-specific probes from Mtel-1

The utility of this collection of mouse telomere YACs was shown by deriving genetic, *in situ* hybridization, and STS markers from one telomere YAC and demonstrating that it can be linked with clones in conventional YAC libraries. At human telomeres subtelomeric repetitive sequences exist close to the $(TTAGGG)_n$ arrays and can be present on more than one chromosome (4). The most internal region of Mtel-1, the first telomere YAC isolated, was subcloned into a plasmid vector. This region was chosen in an attempt to minimize potential problems in isolating single-copy probes. The region cloned was a 1.0 kb vector-adjacent fragment from the *Bam*HI cloning site to the first *Pst*I site in the insert (Fig. 1) and is designated R427. This is released as a 2.6 kb *Pst*I fragment from Mtel-1 which also contains 1.6 kb of vector sequence.

A *Pst*I plasmid library of Mtel-1 genomic DNA was constructed and the 2.6 kb Mtel-1 *Pst*I subfragment which includes R427 was isolated by colony hybridization using a vector probe (Materials and Methods). R427 detected a single *Pst*I fragment by Southern analysis of *M.musculus* genomic DNA, and data from Southern analysis using three other enzymes was consistent with the corresponding locus being present at single copy in the genome (Fig. 2 and data not shown). Southern analysis indicated that R427 hybridized to Mtel-1 but not to any other YAC in this collection (Table 1). These data when taken together argue that the locus corresponding to R427 is single-copy and is not a subtelomeric sequence shared by many telomeres.

BAL31 analysis

There is no *Kpn*I site in Mtel-1 between the $(TTAGGG)_n$ array and R427. If Mtel-1 is a telomeric clone then R427 would be predicted to be on a terminal *Kpn*I fragment in the mouse genome. R427 detects a ~70 kb *Kpn*I fragment in both DBA/2 and C57BL/6 genomic DNA (Fig. 2, left) which reduces in size with progressive pretreatment with exonuclease BAL31. As a control the filter was re-probed with DYEpa1, a repetitive sequence located at non-telomeric sites on the mouse Y chromosome (14). In contrast to the R427 signal, the two DYEpa1 bands are not sensitive to BAL31 (Fig. 2, right). The filter was not stripped before re-probing with DYEpa1, and some R427 signal remains in the right panel of Fig. 2, facilitating comparison of the signals. These data demonstrate that R427 detects a terminal restriction fragment of the mouse genome.

STS and genetic markers from Mtel-1 map to the distal telomere of chromosome 10

Partial sequence (366 bp) of R427 was obtained by single pass sequencing using standard forward and reverse primers. The sequence was compared with the EMBL DNA database using the BLASTN program (15). It was also conceptually translated in all six possible reading frames and compared with the NBRF protein database using the program BLASTX (16). No significant nucleotide or amino acid similarity was detected. PCR primers (E585 and E588; Materials and Methods) were designed based on this sequence using the PRIMER program (10) to provide an STS marker for this locus.

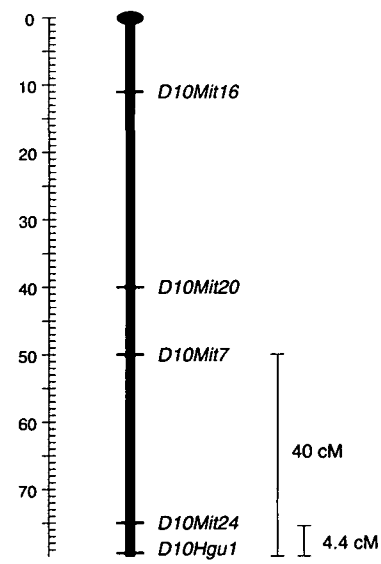

Figure 3. Genetic location of *D10Hgu1*. The location of *D10Hgu1* on chromosome 10 based on the EUCIB backcross is shown, based on data in (17).

There are no $(CA)_n$ or other simple repeats in the sequence obtained which might have formed the basis of a genetic marker. The STS PCR primers E585 and E588 amplify the expected 162 bp product in DBA/2 and an identical sized product in C57BL/6. However, they fortuitously amplify a single band in *M.spretus* which is several base pairs larger in size, sufficient for the C57BL/6 and *M.spretus* alleles to be distinguished by electrophoresis through 4% agarose (data not shown). This polymorphism was used to type 47 animals from the European Collaborative Interspecific Backcross (EUCIB). Twenty-five animals were from the (C57BL/6×*M.spretus*)-F_1×C57BL/6 panel, the remainder being from the backcross to *M.spretus*. This marker segregated as a single Mendelian locus and showed 2/46 recombinants (4.4 cM, LOD = 10.3) with a chromosome 10 marker previously typed in this backcross, *D10Mit24*. As shown in Figure 3, this is the most distal chromosome 10 marker typed in this backcross (17) and maps very close to the distal telomere of this chromosome (10,18). We designate the locus detected by oligonucleotides E585/E588 as *D10Hgu1*.

As *D10Hgu1* is located <70 kb from the end of the chromosome, it is unlikely that it maps centromeric to *D10Mit24*. Because at least eight microsatellite markers map between *D10Mit24* and the distal telomere (10,18), the unlikely scenario of the orientation being *D10Hgu1–D10Mit24*–telomere would require both a large genetic distance (>4.4 cM) and a high density of microsatellites for a small (<70 kb) region of the mouse genome. Mtel-1 was negative for *D10Mit24* and the eight microsatellites which map between *D10Mit24* and the telomere (*D10Mit25*, *D10Mit35*, *D10Mit102*, *D10Mit103*, *D10Mit145*, *D10Mit151*, *D10Mit164*, *D10Mit165*; data not shown). The STS content of Mtel-1 therefore is consistent with the expected mapping of *D10Hgu1* telomeric to *D10Mit24*.

The distance to the next anchor locus, *D10Mit7*, did not allow the orientation of *D10Hgu1* with respect to *D10Mit24* to be accurately determined. However, as shown in Fig. 3,

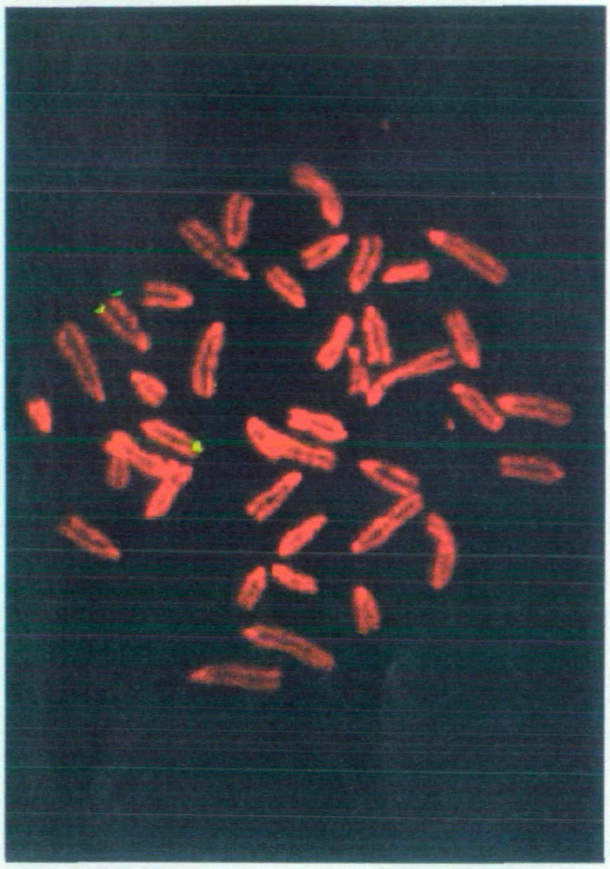

Figure 4. *In situ* hybridization of YAC 110F4 to mouse metaphase chromosomes. YAC 110F4, which contains *D10Hgu1*, was used for *in situ* hybridization to metaphase chromosomes from *M.musculus* ES cell line CGR-8 in the absence of competitor mouse Cot-1 DNA. Chromosomes were counterstained with propidium iodide and the image recorded digitally. In separate experiments with pre-banded chromosomes the single site of hybridization was identified as the distal telomere of chromosome 10.

D10Hgu1 showed 17/42 recombinants with *D10Mit7* (40 cM, LOD = 0.33). *D10Mit7* maps 25 cM from *D10Mit24* in this cross and therefore, if the order were centromere–*D10Mit7*–*D10Hgu1*–*D10Mit24*–telomere, a distance of approximately 20 cM would be predicted between *D10Mit7* and *D10Hgu1*. Instead a distance of ~40 cM between these two loci was observed. Therefore the linkage data are again consistent with *D10Hgu1* being located telomeric to *D10Mit24*. From a combination of physical, STS content and genetic analysis we conclude that *D10Hgu1* maps to the distal telomere of chromosome 10, distal to *D10Mit24*.

In situ hybridization to mouse metaphase chromosomes

The 1.0 kb R427 fragment derived from Mtel-1 proved too small to function as a probe for *in situ* hybridization, so it was necessary to isolate larger clones. Based on the organization of human telomeres no attempt was made to use the entire Mtel-1 YAC as an *in situ* hybridization probe because of the possible existence of subtelomeric repetitive elements which might cross-hybridize to more than one telomere and thus obscure analysis. Instead the STS primers for *D10Hgu1* were used to screen the commercially available Whitehead conventional *M.musculus* YAC library. One positive contained a YAC approximately 450 kb in size visible by ethidium bromide staining following separation by PFGE. This YAC, here designated 110F4, was excised from a preparative low gelling temperature PFGE gel and subjected to a catch-linker PCR protocol to increase the total amount of DNA (Materials and Methods). The PCR products were then used directly as an *in situ* probe against metaphase *M.musculus* chromosomes (Fig. 4). The distal telomeres of a single homologue pair are the only sites labelled. G-banding chromosomes prior to *in situ* hybridization identified this homologue pair as chromosome 10 (data not shown). Thus the YAC isolated using the *D10Hgu1* STS primers hybridizes *in situ* to the distal telomere of chromosome 10, confirming the genetic localization of this locus (see above).

DISCUSSION

The terminal regions of the mouse genome are not represented in conventional YAC, P1, BAC, phage or cosmid libraries, a result of cloning strategies designed to clone only those fragments terminating at both ends with a restriction site. Subtelomeric regions may also be under-represented in large-insert libraries created by partial digestion protocols because many such fragments will be terminal. There is no a priori reason to believe that the terminal regions of the mouse genome will be devoid of genes. Indeed, it is known from telomere fragmentation experiments that mammalian genes can be expressed when located within 1.4 kb of a functional telomere (19) and analysis of the human genome suggests an increased gene density in the terminal regions (20). For this reason it is vital that these regions of the genome are represented in genomic libraries and here we have implemented a targeted approach to clone these regions of the mouse genome. Furthermore, human telomeres have been implicated in the processes of ageing and tumorigenesis (21). To address these questions with a mouse model will require a more detailed understanding of mouse telomere structure and the YAC clones isolated here will aid such studies. Finally, these clones will permit a more detailed analysis of the striking yet poorly understood genetic instability of mouse telomeres (7,8).

How many other mouse telomeres are in this collection? A further six Mtel clones have been analysed and all map to telomeric locations (Kipling *et al.*, manuscript in preparation). Loci mapped include the distal telomeres of chromosomes 9 and 13 and the pseudoautosomal telomere (Table 1). It is likely, therefore, that most clones in this collection correspond to mouse telomeres. A 2n = 40 male inbred strain such as that used here has 41 different telomeres: two telomeres from each of the 19 autosomes, the proximal telomeres of X and Y, and the pseudoautosomal telomere. The restriction maps of most YACs appear distinct from each other without a common pattern of restriction sites (Table 1 and data not shown) and do not suggest that a subset of the telomeres have been preferentially cloned. This is supported by hybridization studies which indicate that the sequence subcloned from Mtel-1 is present on no other YAC in this collection (Table 1). If all 32 clones do indeed contain telomeric YACs, then this collection may contain representatives of more than half of the telomeres in the mouse genome, thus providing a significant resource for further analysis of these regions.

The partial restriction maps do not provide any evidence to support the presence of large amounts of any sequence element shared by different telomeres. This is in contrast to the situation at human telomeres where the presence of subtelomeric sequences shared by a number of telomeres leads to similarities in terminal restriction maps (22). However, we cannot exclude the possibility that such repetitive sequences have been deleted from these YACs.

Genomic digests of DBA/2 DNA with the 4 bp recognition site enzymes *Alu*I, *Hae*III and *Rsa*I release $(TTAGGG)_n$ fragments in the range 20–150 kb (7,8). However, without exception the size of the $(TTAGGG)_n$ fragment of each Mtel YAC in similar digests is less than 2.7 kb (Table 1). Although the possibility cannot be excluded that these YACs have undergone deletion events, there is no evidence for substantial amounts of any other sequence in the YACs between the $(TTAGGG)_n$ arrays and the first adjacent *Alu*I, *Hae*III or *Rsa*I site. The simplest explanation for the difference in genomic and YAC restriction maps is that most genomic $(TTAGGG)_n$-hybridizing *Alu*I, *Hae*III and *Rsa*I fragments are composed predominantly of $(TTAGGG)_n$ and that this sequence has been largely deleted upon cloning in yeast. A similar dramatic decrease in the size of $(TTAGGG)_n$ arrays upon cloning in yeast has been observed for human telomeres (11,22), although its biological basis is not clear.

To demonstrate the utility of clones in this collection a 1.0 kb hybridization probe was isolated from Mtel-1, the first YAC studied. This detects a ~70 kb BAL31-sensitive *Kpn*I fragment in both DBA/2 and C57BL/6 genomes. A PCR-based genetic marker was generated from the sequence of R427. The corresponding locus, *D10Hgu1*, maps to a position consistent with the distal telomere of chromosome 10. STS primers for this locus were used to isolate a 450 kb YAC containing *D10Hgu1* from a conventional YAC library, and this YAC hybridizes *in situ* to the distal telomere of chromosome 10.

R427 provides the first chromosome-specific probe for a mouse telomere. All the previously mapped telomeric sequences have been based on repetitive sequences present on many or all chromosomes. A mouse subtelomeric sequence isolated using a PCR-based approach hybridizes *in situ* to most if not all telomeres (23). Another subtelomeric sequence is the centromeric minor satellite, arrays of which are close to each proximal telomere (12). The arrays found on chromosomes 1, 2 and 14 have been genetically mapped using oligonucleotide probes which detect variants of this satellite (24,25). The distal telomeres of chromosomes 4, 9, 13, X, Y and the proximal telomeres of chromosomes 7 and Y have been placed on the genetic map using RFLPs detected with $(TTAGGG)_n$ probes (12,26,27). Because of the repetitive nature of these probes the corresponding loci are difficult to integrate with the emerging physical map of the mouse genome. In contrast, we have demonstrated that this collection of mouse telomere YACs is capable of yielding chromosome-specific PCR or hybridization probes.

Chromosome-specific telomere probes will be of limited use to analyse the terminal regions of the mouse genome unless they can be integrated with other libraries, thus enabling analysis to be extended into the surrounding region. It is therefore important that any telomeric clones are sufficiently large that they overlap clones in conventional genomic libraries which do not contain telomeres. This requires that such telomere probes be derived from regions at some distance from the $(TTAGGG)_n$ arrays. For example, YAC libraries are usually constructed using and selecting for two vector arms, and therefore only fragments which have a restriction site (e.g. *Eco*RI) at both ends will be cloned. Because of this, terminal *Eco*RI fragments will not be present in such a library, and thus if a telomeric probe is to be present it must be sufficiently internal that there is at least one *Eco*RI site between it and the chromosome terminus. For this reason we chose to clone large fragments of telomeric sequence in yeast, as opposed to using PCR or plasmid cloning, and derived probes from the YACs as distant from the $(TTAGGG)_n$ array as possible. The success of this approach is shown by our ability to isolate YACs containing the corresponding locus from a conventional mouse YAC library, thus enabling this probe to be integrated with the developing STS YAC contig.

This collection of YACs contains regions of the mouse genome not present in other YAC libraries. Our future goal is to assign a chromosomal location to each YAC by *in situ* hybridization prior to their distribution to the research community. Furthermore, DNA from these clones suitable for screening by PCR is immediately available upon request. These YACs will provide a source of chromosome-specific probes and STS markers for individual mouse telomeres for mapping the mouse genome, investigating the unusual genetic behaviour of mouse telomeres, and analysing the role of mammalian telomeres in the processes of ageing and carcinogenesis.

MATERIALS AND METHODS

Manipulation of high molecular weight DNA

Genomic DNA embedded in agarose was prepared from the liver of a male *M.musculus* DBA/2 and digested with *Bam*HI or *Not*I as previously described (7). It was separated by conventional or PFGE through 1% low gelling temperature agarose in 0.5×TBE. The position of limit mobility DNA, as determined by ethidium bromide staining of the adjacent tracks, was excised. It was washed extensively in 1×HI (100 mM NaCl, 10 mM $MgCl_2$, 50 mM Tris–HCl pH 7.5), melted for 10 min at 65°C and cooled briefly to 37°C before adding ATP to 1 mM, dithiothreitol to 1 mM, T4 DNA ligase (Boehringer Mannheim) to 50 U/ml, and pYAC4NEONOT vector arms to give an approximately two- to threefold excess (w/w). The ligation mix was immediately and gently pipetted into PFGE plug formers and allowed to set. Following ligation overnight at 14°C the plugs were electrophoresed through a second conventional or PFGE 1% low gelling temperature agarose gel, using the same conditions as before, to remove unligated vector arms. The limit mobility DNA was again excised, washed extensively in 1×HI, and melted at 65°C for 10 min. It was then liquefied by the addition of 10 U β-agarase (New England Biolabs) per ml of gel slice and overnight incubation at 37°C. The sample was used without further treatment to transform *S.cerevisiae* strain AB1380 spheroplasts.

Three size fractionation schemes were used. For library 1, pYAC4NEONOT (11) was digested with *Not*I and *Bam*HI and dephosphorylated with calf intestinal phosphatase. DBA/2 genomic DNA was digested to completion with *Bam*HI and separated by conventional electrophoresis through 1% low gelling temperature agarose for 20 h at 50 V in 0.5×TBE. The limit mobility DNA (>15 kb) was excised and ligated to the *Bam*HI site of pYAC4NEONOT. This yielded clones Mtel-1 to Mtel-9.

For library 2, pYAC4NEONOT was again digested with *Not*I and *Bam*HI and dephosphorylated with calf intestinal phosphatase. DBA/2 genomic DNA was digested to completion with *Bam*HI and separated by PFGE through 1% low gelling temperature agarose in 0.5×TBE using a BioRad CHEF-DR II apparatus for 48 h at 100 V with a switch time of 0.3 s. The limit mobility DNA, estimated to be >30 kb in this scheme, was excised and ligated to the *Bam*HI site of pYAC4NEONOT. This yielded clones Mtel-10 to Mtel-29.

For library 3, pYAC4NEONOT was digested to completion with *Not*I and dephosphorylated with calf intestinal phosphatase. DBA/2 genomic DNA was digested to completion with *Not*I and separated by PFGE through 1% low gelling temperature agarose in 0.5×TBE using a BioRad CHEF-DR II apparatus for 20 h at 200 V with a switch time ramped linearly from 5 to 15 s. The limit mobility DNA, estimated to be >100 kb, was ligated to the *Not*I site of pYAC4NEONOT. This yielded clones Mtel-30 to Mtel-34.

Yeast transformation

S.cerevisiae strain AB1380 (28) was transformed to uracil prototrophy by the following modification of the procedure of Burgers and Percival (29). AB1380 was grown to an A_{600} of 2.0 in YEPD (1% yeast extract, 2% glucose, 2% peptone) supplemented with 50 mg/l adenine sulphate. Cells were pelleted, washed once with 20 ml sterile distilled water, and once with 20 ml of 1 M sorbitol. The pellet was resuspended in 15 ml of SPE (1 M sorbitol, 10 mM sodium phosphate pH 7.5, 10 mM EDTA), 75 µl of 2 M dithiothreitol, and Yeast Lytic Enzyme (ICN) and incubated for 30 min at 30°C with gentle rotation. The amount of Yeast Lytic Enzyme used was titrated in preliminary experiments to give approximately 75% spheroplasting in 30 min, the degree of spheroplasting being determined by the relative A_{660} of 1:10 dilutions of the cell suspension in 1 M sorbitol and distilled water. The spheroplasts were pelleted and washed twice with 15 ml CaS (1 M sorbitol, 10 mM Tris–HCl pH 7.5, 10 mM $CaCl_2$) and resuspended in 2 ml CaS. 200 µl aliquots of the spheroplasts were added to 10 µl of DNA together with 5 µg of sonicated salmon sperm DNA and incubated at room temperature for 10 min. One millilitre of PEG (20% polyethylene glycol 6000 (BDH), 10 mM Tris–HCl pH 7.5, 10 mM $CaCl_2$) was added and the spheroplasts incubated for a further 10 min at room temperature. The spheroplasts were gently pelleted, resuspended in 150 µl of SOS (1 M sorbitol, 25% YEPD, 10 mM $CaCl_2$) and incubated at 30°C for 40 min before being mixed with top agar and plated on –URA selective sorbitol plates.

Library screening by yeast colony hybridization

Ura^+ transformants were gridded on to 20×20 cm –URA plates in three 8×12 arrays (microtitre dish format) per plate and grown to colonies. Using a 96-prong tool, colonies were replica plated on to Hybond-N membranes placed on the surface of fresh –URA plates. The transferred cells were grown to colonies before being lysed and cross-linked as described (30). The membranes were probed with a radiolabelled $(TTAGGG)_4$ oligonucleotide as described (12). Positive colonies were recovered from the master plates and stored.

Hybridization probes and subcloning Mtel-1

The minor satellite probe was the 'cloned trimer' (25). The major satellite probe was the insert of pMR196 (31). The probe for the *URA3* gene was the 456 bp internal *Stu*I/*Pst*I fragment of the pYAC4 *URA3* gene. The pBR322 probe, designated 'pBR' in Fig. 1, was the 639 bp *Pvu*II/*Ava*I fragment of pBR322 between the *URA3* gene and the *Not*I/*Bam*HI cloning sites of pYAC4NEONOT (11).

To subclone vector-adjacent sequences from Mtel-1, genomic DNA from this yeast strain was digested with *Pst*I and cloned into pBluescript II SK- (Stratagene) using the DH5α *E.coli* strain. The 2.6 kb *Pst*I fragment which includes insert-adjacent vector sequences was isolated by *E.coli* colony hybridization using the *URA3* probe. This approach also isolates the AB1380 genomic *ura3* locus but this can be distinguished by its failure to hybridize to a pBR322 probe, as shown in Figure 1.

BAL31 analysis

High molecular weight DNA embedded in agarose was prepared from male C57BL/6 and DBA/2 mice and digested with BAL31 (Boehringer Mannheim) as described previously (7). Briefly, six 50 µl plugs for each animal were digested with 15 U BAL31 in 500 µl of BAL31 buffer (7). Plugs were removed after 0, 1, 2, 3, 4 and 5 h incubations at 30°C and washed extensively in 10 mM Tris pH 7.5, 1 mM EDTA before digestion with *Kpn*I. The samples were then separated by PFGE using a Biometra Rotaphor. A 0.9% agarose gel and 0.25×TBE buffer at 13°C was used with a switch time ramped logarithmically from 15 to 5 s, voltage ramped logarithmically from 200 to 150 V, an electrode angle of 120° and a total run time of 36 h. Markers used were *S.cerevisiae* chromosomes and Low-Range Marker (New England Biolabs). The gel was neutral blotted by standard techniques onto Hybond N^+ (Amersham). R427 and DYEpa1 were labelled by the random-priming protocol following the manufacturer's instructions (Boehringer Mannheim) and hybridized as previously described (7). Final washes were 0.1×SSC, 65°C, 45 min for R427 and 0.5×SSC, 65°C, 30 min for DYEpa1. Figure 2 shows unedited images obtained using a Molecular Dynamics PhosphorImager.

PCR amplification

D10Hgu1 is defined by oligonucleotides E585 (5'-CTCTGGTTGGAGCTA-CGACC-3') and E588 (5'-GGGCACTCCCTAAATATCAGC-3'). These were designed from partial sequence data of R427 (EMBL database accession number: Z33505). PCR amplification was performed in 25 mM Taps-HCl pH 9.3, 50 mM KCl, 2 mM $MgCl_2$, 0.1% Tween 20, 200 µM each dNTP, 1 µM each oligonucleotide, 25 ng genomic DNA in 50 µl, and 2.5 U Amplitaq *Taq* polymerase. Thermal cycling was performed using a Hybaid OmniGene with a program of 95°C for 5 min (at which point the *Taq* polymerase was added) followed by 35 cycles of 94°C for 30 s, 60°C for 45 s and 72°C for 30 s, with a single final cycle of 10 min at 72°C. Products were separated through a 4% agarose gel (3% NuSieve (FMC), 1% Sigma Type II-A) and stained with ethidium bromide.

Primers to amplify the *D10Mit* markers (10) were purchased from Research Genetics, Alabama. PCR amplification was as above except that the primer concentration was 0.66 µM.

Isolation of YAC 110F4 and localization by *in situ* hybridization to mouse metaphase chromosomes

The Whitehead pYAC4-based C57BL/6J female mouse YAC library produced in AB1380 was purchased from Research Genetics, Alabama. It was screened by PCR using oligonucleotides E585/E588 as described above. Two clones (96G12 and 110F4) positive for *D10Hgu1* were analysed further. Yeast chromosomes were prepared and separated by PFGE (11). 96G12 did not contain a YAC visible by ethidium bromide staining and was not analysed further.

110F4 contains a visible YAC approximately 450 kb in size. Chromosomes from this strain were separated by PFGE through 1% low gelling temperature agarose in 0.25×TBE and visualized by ethidium bromide staining. The YAC was excised and the gel slice washed overnight in 15 ml of 1×M (10 mM Tris–HCl pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$) and three times in 1×M. An equal volume of 1×M buffer was added to the gel slice with 20 U of *Sau*3A. The mixture was incubated on ice for 30 min prior to incubation at 37°C for 3 h. Excess buffer was removed and the enzyme inactivated at 70°C for 15 min. The molten agarose was cooled to 42°C and 10× ligase buffer added to a final concentration of 66 mM Tris–HCl pH 7.5, 50 mM $MgCl_2$, 1 mM dithioerythritol, and 1 mM ATP. 1/20 volume of T4 DNA ligase (5 U/µl) and 1/100 volume of oligonucleotides D921 and D922 (100 ng/µl) were added. Oligonucleotides D921 (5'-GTCAAGAATTCTGTACCGTGGAC-3') and D922 (5'-GATCGTGCACGGTAGCGAATTCT-3') anneal to give a GATC overhang, which can ligate to *Sau*3A digested YAC DNA.

Following overnight ligation at 15°C 0.5 µl of the ligation mixture was PCR amplified using oligonucleotide D921 in a 50 µl reaction containing 250 ng of D921 and the same buffer conditions as for the *D10Hgu1* PCR assay. The cycling program used was 94°C for 3 min, after which 1.5 U Amplitaq *Taq* polymerase was added, followed by 30 cycles of 94°C for 45 s, 60°C for 1 min and 72°C for 1 min (with a 1 s increment each cycle), followed by a final single cycle of 72°C for 10 min. The PCR products were diluted with two volumes of distilled water prior to ethanol precipitation and labelled by nick translation in the presence of biotin-16-dUTP without further processing. *In situ* hybridization to metaphase chromosomes from the *M.musculus* ES cell line CGR-8 was as described (32) with some slides being G-banded, photographed and destained prior to *in situ* hybridization. The hybridized probe was detected with avidin-FITC and chromosomes were counterstained with propidium iodide, fluorescence images being recorded digitally using a BioRad laser-scanning-confocal microscope.

DNA database accession number

The partial sequence of R427 has been submitted to the EMBL database and assigned the accession number Z33505.

ACKNOWLEDGEMENTS

We thank Emma Tarttelin at the Medical Research Council's Human Genome Mapping Project Resource Centre for providing EUCIB DNAs and performing the *MBx* analysis, Ian Jackson for DNA pools from the Research Genetics library, and Charles Patek for DYEpa1. We are grateful to Cathy Abbott, Sarah Bowen and members of the Chromosome Biology Section for their comments on the manuscript. This work was supported by a grant from the Medical Research Council's Human Genome Mapping Project. D.K. is a Beit Memorial Fellow.

REFERENCES

1. Blackburn, E.H. (1991) Structure and function of telomeres. *Nature* **350,** 569–573.
2. Biessmann, H. and Mason, J.M. (1992) Genetics and molecular biology of telomeres. *Adv. Genet.* **30,** 185–249.

3. Kipling, D. (1995) *The Telomere.* Oxford University Press, Oxford.
4. Kipling, D. and Cooke, H.J. (1992) Beginning or end? Telomere structure, genetics and biology. *Hum. Mol. Genet.* **1**, 3–6.
5. NIH/CEPH Collaborative Mapping Group (1992) A comprehensive genetic linkage map of the human genome. *Science* **258**, 67–86.
6. Meyne, J., Ratliff, R.L. and Moyzis, R.K. (1989) Conservation of the human telomere sequence $(TTAGGG)_n$ among vertebrates. *Proc. Natl Acad. Sci. USA* **86**, 7049–7053.
7. Kipling, D. and Cooke, H.J. (1990) Hypervariable ultra-long telomeres in mice. *Nature* **347**, 400–402.
8. Starling, J.A., Maule, J., Hastie, N.D. and Allshire, R.C. (1990) Extensive telomere repeat arrays in mouse are hypervariable. *Nucleic Acids Res.* **18**, 6881–6888.
9. NIH/DOE Mouse Working Group (1993) A plan for the mouse genome project. *Mammal. Genome* **4**, 293–300.
10. Dietrich, W.F., Miller, J.C., Steen, R.G., Merchant, M., Damron, D., Nahf, R., Gross, A., Joyce, D.C., Wessel, M., Dredge, R.D., Marquis, A., Stein, L.D., Goodman, N., Page, D.C. and Lander, E.S. (1994) A genetic map of the mouse with 4006 simple sequence length polymorphisms. *Nature Genet.* **7**, 220–245.
11. Cross, S.H., Allshire, R.C., McKay, S.J., McGill, N.I. and Cooke, H.J. (1989) Cloning of human telomeres by complementation in yeast. *Nature* **338**, 771–774.
12. Kipling, D., Ackford, H.E., Taylor, B.A. and Cooke, H.J. (1991) Mouse minor satellite DNA genetically maps to the centromere and is physically linked to the proximal telomere. *Genomics* **11**, 235–241.
13. Neil, D.L., Villasante, A., Fisher, R.B., Vetrie, D., Cox, B. and Tyler-Smith, C. (1990) Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome vectors. *Nucleic Acids Res.* **18**, 1421–1428.
14. Patek, C.E., Olszewska, E., Kipling, D., Armstrong, J.F., Sealey, P.G., Lee, M., Speed, R.M., Hooper, M.L. and Jones, K.W. (1994) DYEpa1: a novel mouse Y chromosome-specific repetitive DNA element. *Mouse Genome* **92**, 696–698.
15. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment tool. *J. Mol. Biol.* **215**, 403–410.
16. Gish, W. and States, D.J. (1993) Identification of protein coding regions by database similarity search. *Nature Genet.* **3**, 266–272.
17. European Backcross Collaborative Group (1994) Towards high resolution maps of the mouse and human genomes—a facility for ordering markers to 0.1 cM resolution. *Hum. Mol. Genet.* **3**, 621–628.
18. Taylor, B.A., Frankel, W.N., Burmeister, M. and Bryda, E. (1993) Mouse chromosome 10. *Mammal. Genome* **4**, S154–163.
19. Bayne, R.A.L., Broccoli, D., Taggart, M.H., Thomson, E.J., Farr, C.J. and Cooke, H.J. (1994) Sandwiching of a gene within 12 kb of a functional telomere and alpha satellite does not result in silencing. *Hum. Mol. Genet.* **3**, 539–546.
20. Craig, J.M. and Bickmore, W.A. (1993) Chromosome bands—flavours to savour. *Bioessays* **15**, 349–354.
21. Kipling, D. (1995) Telomerase: immortality enzyme or oncogene? *Nature Genet.* **9**, 104–106.
22. Brown, W.R.A., MacKinnon, P.J., Villasanté, A., Spurr, N., Buckle, V.J. and Dobson, M.J. (1990) Structure and polymorphism of human telomere-associated DNA. *Cell* **63**, 119–132.
23. Broccoli, D., Miller, O.J. and Miller, D.A. (1992) Isolation and characterization of a mouse subtelomeric sequence. *Chromosoma* **101**, 442–447.
24. Hayashi, T., Ohtsuka, H., Kuwabara, K., Mafune, Y., Miyashita, N., Moriwaki, K., Takahashi, Y. and Kominami, R. (1993) A variant family of mouse minor satellite located on the centromeric region of chromosome 2. *Genomics* **17**, 490–492.
25. Kipling, D., Wilson, H.E., Mitchell, A.R., Taylor, B.A. and Cooke, H.J. (1994) Mouse centromere mapping using oligonucleotide probes that detect variants of the minor satellite. *Chromosoma* **103**, 46–55.
26. Elliott, R.W. and Yen, C-H. (1991) DNA variants with telomere probe enable genetic mapping of ends of mouse chromosomes. *Mammal. Genome* **1**, 118–122.
27. Eicher, E.M., Lee, B.K., Washburn, L.L., Hale, D.W. and King, T.R. (1992) Telomere-related markers for the pseudoautosomal region of the mouse genome. *Proc. Natl Acad. Sci. USA* **89**, 2160–2164.
28. Burke, D.T., Carle, G.F. and Olson, M.V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**, 806–812.
29. Burgers, P.M. and Percival, K.J. (1987) Transformation of yeast spheroplasts without cell fusion. *Anal. Biochem.* **163**, 391–397.
30. Bates, G.P., MacDonald, M.E., Baxendale, S., Sedlacek, Z., Youngman, S., Romano, D., Whaley, W.L., Allitto, B.A., Poustka, A., Gusella, J.F. and Lehrach, H. (1990) A yeast artificial chromosome telomere clone spanning a possible location of the Huntingdon Disease gene. *Am. J. Hum. Genet.* **46**, 762–775.
31. Pietras, D.F., Bennett, K.L., Siracusa, L.D., Woodworth-Gutai, M., Chapman, V.M., Gross, K.W., Kane-Haas, C. and Hastie, N.D. (1983) Construction of a small *Mus musculus* repetitive DNA library: identification of a new satellite sequence in *Mus musculus*. *Nucleic Acids Res.* **11**, 6965–6983.
32. Farr, C., Fantes, J., Goodfellow, P. and Cooke, H. (1991) Functional reintroduction of human telomeres into mammalian cells. *Proc. Natl Acad. Sci. USA* **88**, 7006–7010.