Development 101 Supplement, 101–106 (1987)
Printed in Great Britain © The Company of Biologists Limited 1987

Telomeres of the human X and Y chromosomes

HOWARD J. COOKE

MRC Clinical and Populations Cytogenetics Unit, Outstation at Kings Buildings, West Mains Road, Edinburgh EH9 3JT, UK

Summary

Analysis of the nature of the DNA sequences at the telomeres of the short arms of the human sex chromosomes suggests parallels with the structures found at telomeres of a number of lower eukaryotes. The exact nature of the end of the DNA has not yet been established but it is clear that there are multiple levels of variability both between and within individuals in this region of the genome.

Key words: telomeres, X chromosome, Y chromosome, human.

Introduction

The human sex chromosomes have been the subject of intensive genetic and molecular study. This has been driven largely by the ease of both genetic and molecular analysis afforded by their monozygosity in the male and the absence of a Y chromosome in the female. In the case of both the X and the Y chromosomes, a large set of anonymous DNA markers have been assigned to these chromosomes and localized or mapped cytogenetically or by recombination. In the case of the Y chromosome, the rationale for this work has been the search for the locus responsible for male sex determination. Several by-products have developed from this search, including the realization that the Y chromosome is a mosaic of sequences which are X related, autosome related and Y specific and the proof that a portion of the chromosome is genetically autosomal. The subject of this article is the nature of the telomeres of the X and Y chromosomes. The finding of telomeric sequences was again a by-product of a search for genes present on the Y chromosome. It is not surprising that, in the quest for understanding and control of gene expression, differentiation, carcinogenesis and inherited disease, genes are often regarded as the only functional entity that resides in the DNA. This is a failing of molecular biologists and especially of those of us who work on human systems. We have a good excuse; human genetics would not be regarded as 'real' genetics by the standards of bacterial, yeast, *Drosophila* or even mouse genetics. The problems of human molecular biology parallel those of human genetics; manipulations possible in yeast are not currently possible in mammalian systems. To find a wider view of possible functions of DNA other than genes one must borrow from other organisms and disciplines. Cytology provides the concepts of chromosomal elements such as centromeres and telomeres which, although not genes, have important functions. Organisms such as yeast and *Tetrahymena* have been utilized to work out the molecular structure of these noncoding but nevertheless functional regions of DNA. In this article, I will review our limited understanding of human telomeres borrowing heavily from yeast and other lower eukaryotes to construct a model.

Telomeres are the natural end of a linear DNA molecule. They have a number of functions to fulfil. All known DNA polymerases require a primer, usually a RNA. Excision of this primer would leave one of the daughter DNA strands shorter than its parent and over many rounds of replication this would lead to extensive deletions at the ends of the chromosomes. A number of replication schemes have been proposed which could avoid this problem (for a review see Blackburn & Szostak (1984)). Another function that the sequences at the telomeres must display is the protection against degradation or fusion of free ends of DNA in the cell. Fragments of DNA introduced into mammalian cells are normally integrated into a chromosome whereas the chromosomes themselves are stable. At the cellular level, telomeres are associated with the nuclear membrane and have a specific distribution within the nucleus (Agard & Sedat, 1983). What features of the structure of telomeres can be involved in these functions?

Table 1. *Terminal repeats*

Tetrahymena	$(CCCCAA)_n$
S. cerevisiae	$(C_{1-3}A)_n$
Trypanosoma	$(CCCTAA)_n$
Oxytricia	$(CCCCAAAA)_n$

Terminal repeats

In those organisms in which the chromosome termini have been studied at the molecular level, a variable number of short terminal repeats have been found. These sequences are presented in Table 1. Within a cloned population of cells, the number of repeats found at the end of a particular chromosome varies from cell to cell. One result of this is that restriction fragments that terminate at telomeres are heterogeneous in size. In yeast, the average number of repeats is under genetic control. CDC17, a temperature-sensitive lethal mutation which causes a block in the cell cycle, increases the length of terminal restriction fragments by increasing the number of repeats (Carson & Hartwell, 1985). Two other genes are known in *S. cerevisiae* which are also lethal and give rise to telomeres which have fewer terminal repeats than wild-type cells. These genes may be involved in turnover processes which subtract and add repeats from the telomeres (Lustig & Petes, 1986). Only one such process has been directly demonstrated. In *Tetrahymena* an activity has been isolated which can add the *Tetrahymena* macronuclear terminal repeats, in a terminal transferase-like manner, to either *Tetrahymena* or yeast terminal repeat sequences acting as primers (Greidler & Blackburn, 1985).

The first clues suggesting that we were observing human telomeres came from observations (Cooke *et al.* 1985) that a probe isolated from a Y-chromosome-derived cosmid detected heterogeneously sized restriction fragments in DNA samples from single individuals when digested with some restriction enzymes whilst detecting homogeneous fragments when the DNAs were digested with other enzymes. This was directly comparable with the observations described above on the telomeres of yeast and other organisms. Other criteria that any naturally occurring end of a DNA molecule must satisfy are that this end should be nuclease sensitive in intact DNA and should appear as a restriction site for all restriction enzymes when a restriction map is generated. One obvious additional criterion that a telomere should demonstrate is that sequences adjacent to it should map to the ends of a chromosome cytologically. The sequences detected by this probe fulfilled all these criteria.

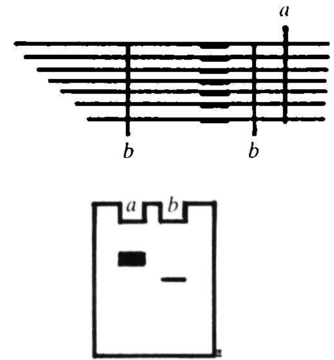
The observation of heterogeneity in DNA fragments that originate from a single individual is consistent with the idea that human telomeres are like the telomeres of other organisms which have been studied at the molecular level in that there is a variation from cell to cell in the number of short repeats found at the end of the chromosome. The repeat length is such that the quantal nature of this variation cannot be seen. In this case, our probe is more than 15 kb from the chromosome end and we would not expect that a repeat unit of less than 100 bp would be resolved. The repeat unit may, of course, be very much smaller than this and its existence must remain a hypothesis until it has been cloned and sequenced. It remains, however, the most likely explanation for the observation of heterogeneity.

The extent of the heterogeneity is the same within different tissues of a single individual with two exceptions. When sperm and blood DNAs from a single individual are compared the distance between the terminal *Bam*HI site and the telomere is greater by about 5 kb in the germ line DNA and the heterogeneity is either reduced or absent (Cooke & Smith, 1986). This situation is reminiscent of the phenomenon of germ-line diminution found in the nematode *Ascaris*. In this organism, up to 25 % of the germ-line DNA is present as heterochromatic blocks of tandemly repeated satellite DNA which is lost during the development of the somatic tissues (Blackburn & Szostak, 1984). A loosely comparable situation exists in the holotrichous ciliates such as *Tetrahymena* and *Oxytricha* in which the micronucleus constitutes the germ line. After conjugation and fusion of the micronuclei, the DNA is fragmented and in this process specific DNA sequences are lost. It appears that a related process is taking place on a smaller scale at human telomeres. I stress that this is a loose analogy, the bulk of the complex changes that occur in the development of the ciliate nuclei are not reflected in the human system. The second exception is that a lymphoblastoid line, although heterogeneous, shows a much reduced range of terminal fragment lengths in comparison to blood DNA from the same individual. The simplest explanation is that the number of terminal repeats is loosely fixed in cell lineages but varies considerably between lineages.

Changes in the number of terminal repeats have been observed in trypanosomes which suggest that one repeat is added per cell division in a process which can continue through many cell cycles and is terminated by a sudden loss of repeats (Bernards *et al.* 1983). Human chromosomes, at least in cultured cells, do not seem to be subject to such a process since a comparison of early and late passages of a lymphoblastoid line reveals no differences in terminal restriction fragment length. Perhaps, after the loss of a variable amount of DNA from the chromosome ends at some stage during development, a steady state is

reached in which processes of addition and deletion of repeats are balanced giving rise to the small degree of heterogeneity seen in transformed cell lines (Fig. 2).

In *S. cerevisiae* the only structures necessary for stabilization and replication of artificial chromosome ends are the terminal repeats. When terminal repeats from *Tetrahymena* rDNA are used as the ends of a linear plasmid, yeast terminal repeats are added to them (Shampay *et al.* 1984). The general structure of these repeats in single-cell eukaryotes is conserved. The probable reason for this is that these sequences are involved in interactions with specific proteins. In *Oxytricha*, two different proteins have been isolated tightly bound to the macronucleus telomeres. One of these proteins will bind *in vitro* to a synthetic telomere



Bam *Bgl* *Hin* *Eco*RI *Pst* *Sfi* *Sst* *Taq*

B

Fig. 1. (A) Restriction of heterogeneous ends. Digestion with a restriction enzyme that cuts at a site such as *a* gives a pattern of heterogeneous fragments when the probe (the thickened line in this drawing) is located between the restriction site and the end of the molecule. An enzyme that cuts at sites *b* gives rise to homogeneously sized fragments detected by the same probe. (B) In this blot *Bam*, *Eco*RI and *Sfi* give a type *a* pattern and the other enzymes type *b* patterns.

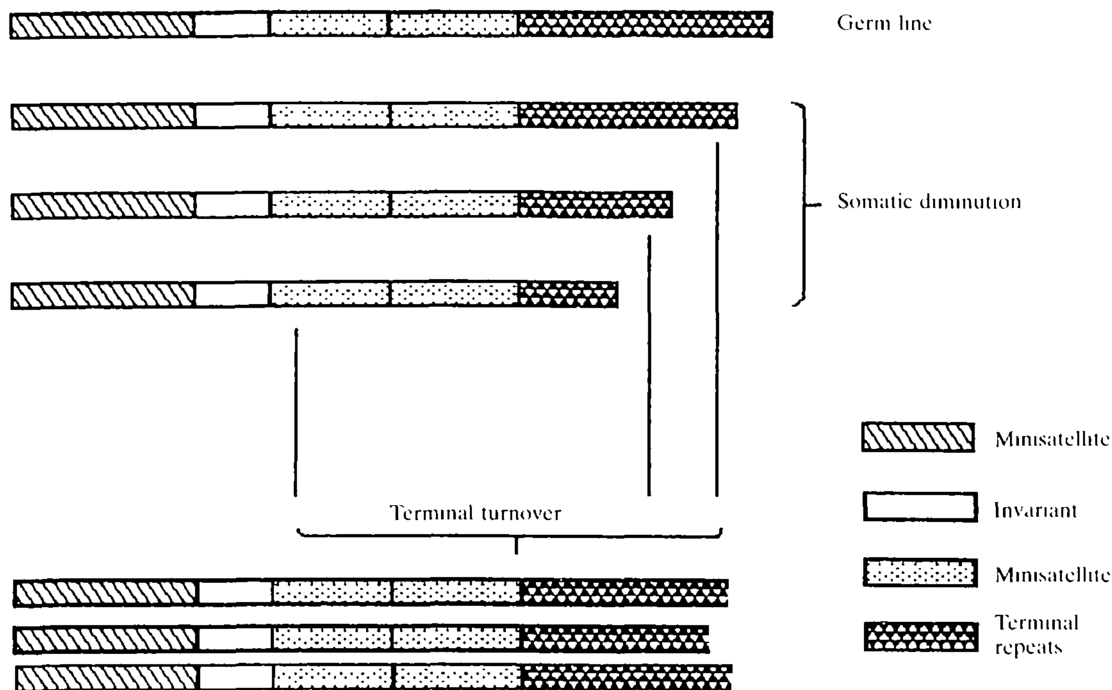


Fig. 2. Variations in the numbers of terminal repeats. Deletion of sequences from the ends of the chromosomes occurs during development. This process is arrested at different stages in different cell lineages. A turnover of repeat units may then occur giving rise to microheterogeneity seen in cell lines. Integral regions are unaffected by these processes.

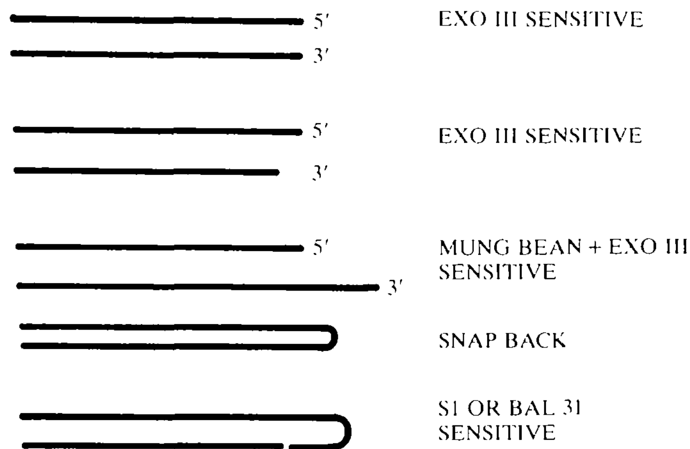


Fig. 3. Possible structures for the end of a chromosome.

in a manner that suggests that the presence of a single-strand tail and duplex region, which must have the natural DNA sequence, are both recognized by the protein (Gottschling & Zakian, 1986). In *S. cerevisiae* telomere binding activity has also been detected but here the sequence specificity is simpler as both internal and external copies of the yeast terminal repeat $C_{1-3}A$ are recognized (Berman *et al.* 1986). Do human telomeres fall into this general pattern? The indirect evidence strongly suggests the presence of terminal repeats but their structure and the presence of bound proteins awaits direct proof.

What is the end?

The nature of the DNA terminus has been investigated in a number of ciliated protozoa using *in vitro* labelling techniques. These methods are impractical for the study of human telomeres because there are only 92 ends per nucleus. *Oxytricha* for example has approx. 10^7 ends per macronucleus. Instead nucleases with different specificities can be used as probes for the nature of the terminus. Exonuclease III will degrade molecules with blunt ends or 5′ extensions but will not degrade molecules with 3′ extensions. Treatment of high molecular weight human DNA with exonuclease III leaves telomeric restriction fragments intact. If the terminus was a 3′ extension it should be possible to remove this with an enzyme such as mung bean nuclease to give molecules which are then sensitive to digestion with exonuclease III. This is not the case for the X/Y telomeres. The remaining possibility is that there is a loop structure present. If this was a covalent joining of the two strands of the DNA then these molecules would 'snap-back' into a duplex when renatured at low concentration. Again this does not happen leaving the most probable structure as that of a hairpin which is not covalently closed. Such a structure would be predicted to be *Bal*BI sensitive, as is observed to be the case. This is consistent with one model for telomere replication and with the results from labelling methods in a number of ciliated protozoa.

Subterminal variability

Within 30 kb of the X/Y short-arm telomeres, there is a great deal of variability in the population (Cooke & Smith, 1986). The basis of this variability is a number of blocks of minisatellite sequences. The probe that defines DXYS14 contains over 1 kb of a minisatellite composed of multiple imperfect copies of a 31-base pair repeat. This minisatellite lies within a block of sequences, which may be duplicated or not, on different chromosomes. The size of the minisatellite block is itself variable. In combination, these two levels of variability result in at least 150 different alleles at this locus. Separated from this region by about 5 kb of invariant DNA is a block of a different minisatellite, about 20 kb long, which shows internal restriction site polymorphisms. Neither of these minisatellites are present elsewhere in the human genome nor are they converved outside the higher primates. This distribution contrasts sharply with those minisatellites described by Jeffries *et al.* which are present at multiple locations (Jeffreys *et al.* 1985). Because of this distinction, these telomere-associated minisatellites cannot be involved in any function that involves all telomeres such as replication or chromosome 'capping'. Perhaps they are involved in positioning or identifying telomeres of chromosomes that pair. If so, one would predict that this class of sequence would exist at all telomeres but that the repeat on different chromosomes would have a different sequence. The only human telomere-associated sequences known are at the X/Y pairing region and so this hypothesis remains untested.

S. cerevisiae also has telomere-associated repeated sequences which are more complex in structure than the terminal repeats. All yeast chromosomes have a copy of a sequence designated as X with between 0 and 4 copies of a sequence Y'. This last sequence was originally isolated as containing a yeast origin of replication. Stretches of C_{1-3}rA repeats separate these elements and Y' elements when these are tandemly repeated. Like the human telomere-associated repeats these sequences are highly variable and have been shown to be involved in recombination (both meiotic and mitotic) (Horowitz *et al.* 1984) and to exist as free circular forms. Although the similarities in position of these yeast and human telomere-associated repeats invite the conclusion that they may have some similar role which is perhaps more subtle than either replication of the ends or 'capping' of the DNA termini, the distribution of these sequences is different.

Many questions remain to be answered about telomeres in general and human telomeres in particular. I have posed some here already. The major problem is that we have no sequence information about the end of the chromosome. The most centromere-distal sequences that we have so far been able to clone are a minimum of 2 kb from the end of the shortest X/Y chromosome in the cell populations that we use. The most reasonable model would suggest that terminal repeats on all chromosomes would be capable of being replicated by the same enzymatic machinery and that this constraint would result in related, if not identical, sequences at the chromosome ends. Cloning of such sequences from the X/Y chromosome short arms should then allow the analysis of the telomeric region of a number of different chromosomes. From this data, it will be possible to draw some conclusions about conserved features which may be significant for telomere function.

A knowledge of the terminal sequences will allow a search for proteins that recognize these sequences and either bind to them or are involved in their replication. It will be possible to look for the presence of cryptic telomeres in the human genome, i.e. sequences that are normally present internally in a chromosome but that are potentially capable of telomere function perhaps as a result of breakage. The most interesting prospect is that of the construction of synthetic chromosomes for mammalian cells. Functional telomeres are only one requirement for this. Centromeres will also be necessary and there may be other, as yet undefined, essential elements. We are close to at least one telomere. Intriguing data from injection of viral constructs into mouse eggs suggest that centromere function may be exercised by relatively small DNA elements (Rassoulzadegan *et al.* 1986). I do not think it is too far-fetched to suggest that constructs that mimic at least some aspects of mammalian chromosome behaviour will be available in the near future.

References

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

Berman, J., Tachibana, C. Y. & Tye, B-K. (1986). Identification of a telomere-binding activity from yeast. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3713–3717.

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

Blackburn, E. H. & Szostak, J. W. (1984). The molecular structure of centromeres and telomeres. *A. Rev. Biochem.* **53**, 163–194.

Carson, M. J. & Hartwell, L. (1985). CDC17: An essential gene that prevents telomere elongation in yeast. *Cell* **42**, 249–257.

Cooke, H. J., Brown, W. R. A. & Rappold, G. A. (1985). Hypervariable telomeric sequences from the

human sex chromosomes are pseudoautosomal. *Nature, Lond.* **317**, 687–692.

COOKE, H. J. & SMITH, B A (1986). Variability at the telomeres of the human X/Y pseudoautosomal region. *Cold Spring Harbor Symp. Quant. Biol.* **51**, 213–219.

GOTTSCHLING, D. E. & ZAKIAN, V. A. (1986). Telomere proteins: Specific recognition and protection of the natural termini of oxytricha macronuclear DNA. *Cell* **47**, 195–205.

GREIDER, C. W & BLACKBURN, E. H. (1985). Identification of a specific telomere terminal transferase activity in tetrahymena extracts. *Cell* **43**, 405–413.

HOROWITZ, H., THORNBURN, P. & HABER, J. E. (1984). Rearrangements of highly polymorphic regions near telomeres of *Saccharomyces cerevisiae. Molec. cell. biol.* **4**, 2509–2517.

JEFFREYS, A. J., WILSON, V. & THEIN, S. L. (1985). Hypervariable 'minisatellite' regions in human DNA. *Nature, Lond.* **314**, 67–73.

LUSTIG, A. J. & PETES, T. D. (1986). Identification of yeast mutants with altered telomere structure. *Proc. natn. Acad. Sci U.S.A.* **83**, 1398–1402.

RASSOULZADEGAN, P., LÉOPOLD, P., VAILLY, J. & CUZIN, F. (1986). Germ line transmission of autonomous genetic elements in transgenic mouse strains. *Cell* **46**, 513–319.

SHAMPAY, J., SZOSTAK, J. W. & BLACKBURN, E. H. (1984). DNA sequences of telomeres maintained in yeast. *Nature, Lond.* **310**, 154–157.