MOLECULAR AND CELLULAR BIOLOGY, Dec. 1994, p. 8028–8036
0270-7306/94/$04.00+0

Vol. 14, No. 12

A Family of Complex Tandem DNA Repeats in the Telomeres of *Chironomus pallidivittatus*

YU-JING ZHANG,† IRÉNE KAMNERT, CASIMIRO C. LÓPEZ, MARITA COHN,‡ AND JAN-ERIK EDSTRÖM*

Department of Molecular Genetics, University of Lund, S-223 62 Lund, Sweden

Received 27 December 1993/Returned for modification 11 August 1994/Accepted 24 August 1994

A family of 340-bp tandem telomere-associated DNA repeats is present in 50- to 200-kb blocks in seven of the eight paired chromosome ends in *Chironomus pallidivittatus*. It consists of four main subfamilies, differing from each other by small clusters of mutations. This differentiation may reflect different functional roles for the repeats. Here we find that one subfamily, D3, is consistently localized most peripherally and extends close to the ends of the chromosomes, as shown by its sensitivity to the exonuclease *Bal* 31. The amounts of D3 are highly variable between individuals. The repeat characteristic for D3 forms a segment with pronounced dyad symmetry, which in single-strand form would give rise to a hairpin. Evidence from an interspecies comparison suggests that a similar structure is the result of selective forces. Another subfamily, M1, is present more proximally in a subgroup of telomeres characterized by a special kind of repeat variability. Thus, a complex block with three kinds of subfamilies may occupy different M1 telomeres depending on the stock of animals. We conclude that subfamilies are differentially distributed between and within telomeres and are likely to serve different functions.

Most investigated eukaryotic chromosomes end with short repeats having an excess of guanines in the strand with 3′ termination. Such DNA is termed telomeric DNA, whereas more complex DNA at the chromosome ends is referred to as telomere-associated DNA (9) and is subtelomeric in species where telomeric DNA is present. Telomere-associated DNA may, however, be terminal, as in *Tetrahymena thermophila* linear mitochondrial DNA (33) or *Drosophila melanogaster* nuclear telomeres (26). Furthermore, what is normally subtelomeric DNA can in certain yeast mutants become terminal and permit survival and growth (31). Telomeric DNA, on the other hand, may occupy subterminal positions (18, 30, 48). Thus, there is no absolute topological distinction between the two types of DNA. One telomeric function, for which telomerase is responsible, is to restore DNA that is lost after replications (22, 23). Another is the maintenance of chromosomal integrity (see reference 2 and references therein) believed to be related to the propensity of telomeric DNA to form secondary structures (see reference 6 and references therein). Obviously, telomere-associated DNA, when in a terminal position, also has to fulfil these functions. Processes like unequal crossing over (34), gene conversion (39), and transposition (1, 3, 4, 16, 26, 45, 46) can be invoked to explain DNA additions. Little is known, however, about how telomere-associated DNA could maintain chromosomal integrity and about its subtelomeric roles.

Chromosome ends in the dipteran insect *Chironomus pallidivittatus* have a potential value in the study of telomere-associated DNA, because of the schematic structure of their long tandem arrays of complex 340-bp repeat units (TA repeats) (43). They occur in distinct subfamilies, three of which, D1, D2, and D3, are derived from a fourth one, M1, each by a cluster of mutations (12, 13). They are present in seven chromosome ends; the eighth one, at the telocentric end of chromosome 4, contains a different repeat (42). Here we ask whether this differentiation is correlated to properties that could explain the functions of telomere-associated DNA. We show that D3 is the most distal subfamily in the blocks of TA repeats and may extend to, or very close to, the termini. We also find that the region specific for D3 contains a region with pronounced dyad symmetry, i.e., the potential ability to form a hairpin in the single-stranded form. Interspecies comparison shows that this property is a result of selective forces. D3 is also highly variable in amount, even within stocks. Characteristic of M1, on the other hand, is its presence in a subgroup of telomeres, characterized by a specific type of variability involving a complex of three subfamilies. D1 and D2, finally, are exclusively associated with this complex and are always found together.

MATERIALS AND METHODS

Animals. Fourth-instar larvae of *C. pallidivittatus* (14) were used both for in situ hybridizations and for DNA preparations. Two different stocks, derived from the same mass culture in 1987, were used. One stock has been maintained in Lund, Sweden, and the other has been maintained in Göttingen, Germany.

In situ hybridization with one probe. Chromosome squashes were hybridized to digoxigenin-labeled ds1, ds2, and ds3 oligonucleotide as described previously (14) for subfamily D1, D2 and D3 localization, respectively. For subfamily M1 localization an M1 oligonucleotide, 5′-ATATATCGATGGATGC GTCTCATTAGTATCCGGAATTCACCCATACAAACC-3′, was labeled by Klenow extension of a 5′-CACAGGTT TG-3′ primer with a 6-base overlap with the M1 oligonucleotide. For digoxigenin labeling, 0.5 μg of a complementary oligonucleotide and 0.15 μg of primer were heated for 3 min in 15 μl of water at 70°C and cooled, after which 2 μl of 10×

* Corresponding author. Mailing address: Department of Molecular Genetics, University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden. Phone: 46-46-107253. Fax: 46-46-104023.

† Permanent address: Changchun University of Agriculture and Animal Sciences, Changchun 130062, Jilin, People's Republic of China.

‡ Present address: Department of Microbiology and Immunology, University of California, San Francisco, CA 94143-0414.

Klenow buffer was added; then 2 μl of deoxynucleoside triphosphate labeling mixture (Boehringer) and 1 μl of Klenow labeling-grade fragment (2 U/μl) were added. Incubation was done at 37°C overnight. Hybridization and washing (twice for 15 min) at 60°C were performed as previously described (like digoxigenin detection and fluorescence microscopy) (14).

In situ hybridization with two probes. Oligonucleotides were labeled with digoxigenin or biotin either by Klenow extension as described above or by tailing in terminal deoxynucleotidyl-transferase-catalyzed reactions and with digoxigenin-dUTP or biotin-16-dUTP as precursors as specified by the supplier (Boehringer). Digoxigenin-labeled probes were assayed with rhodamine-conjugated Fab antibody fragments (Boehringer) at a dilution of 1:250. Biotin-labeled hybrids were detected with fluorescein in an indirect method (44) involving avidin-fluorescein isothiocyanate conjugate (Sigma) followed by biotinylated anti-avidin D (Vector Laboratories).

Preparation of high-molecular-weight DNA. High-molecular-weight DNA was prepared as described previously (5) with minor modifications. The size of the DNA, determined by pulsed-field gel electrophoresis, was about 300 kb.

Pulsed-field gel electrophoresis. High-molecular-weight DNA was pipetted directly into the wells of 1.2% agarose gels. Separation was carried out with a Gene Navigator unit (Pharmacia-LKB) at 12°C and 200 V for 21 h in 0.5× TBE buffer (1× TBE is 0.045 M Tris-borate buffer [pH 8.0] and 0.001 M EDTA). The pulse times were 0.5 s for 30 min, 8 s for 30 min, 1 s for 3 h, 2 s for 3 h, 4 s for 6 h, and 8 s for 8 h.

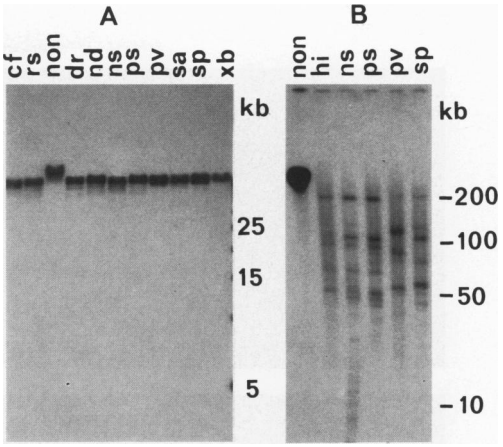

FIG. 1. (A) Electrophoretic separations of genomic DNA from the Göttingen stock cleaved with *Cfo*I (cf), *Rsa*I (rs), nondigested (non), *Dra*I (dr), *Nde*I (nd), *Nsi*I (ns), *Pst*I (ps), *Pvu*II (pv), *Sac*I (sa), *Spe*I (sp), and *Xba*I (xb). Electrophoresis was run in 0.3% agarose (SeaKem Gold; FMC Bioproducts) in 0.04 M Tris-acetate buffer (pH 8.0)–0.001 M EDTA with size markers (λ*Hin*dIII digest and φX 5 kb ladder) in an 18-cm-long gel at 1 V/cm for 18 h. Hybridization was done with pCp 306 insert, ^{32}P labeled by random priming. This is a cloned TA repeat of the M1 type (43). (B) Pulsed-field gel electrophoresis separations of restricted genomic DNA from the Göttingen stock: nondigested (non), *Hin*dIII (hi), *Nsi*I (ns), *Pst*I (ps), *Pvu*II (pv), and *Spe*I (sp).

RESULTS

Cytology. Each of the four chromosomes is a pair of somatically united homologs, three large pairs (chromosomes 1 to 3) and the small one (chromosome 4). The large chromosomes are metacentric, and chromosome 4 is telocentric, the left end, 4L, containing the centromere. Since a dipteran "chromosome" is a bivalent, ends from individual chromosomes can be seen separately only exceptionally. A telomere with a simple set of repeat types in one chromosome and a more complex mixture in the homolog would therefore be scored as complex.

Blocks of TA repeats. Genomic DNA was restricted, with enzymes having 4- or 6-base recognition sites, not represented in any known TA repeat sequence. The restricted DNA was electrophoresed in 0.3% agarose and hybridized to a ^{32}P-labeled insert of the pCp306 clone (43) containing a 340-bp TA repeat unit. Figure 1A shows that essentially all DNA is cleaved into fragments 50 kb or larger. For some digests the size distribution of these fragments was investigated by pulsed-field gel electrophoresis. Figure 1B shows that DNA is cut into 6 to 10 bands, 50 to 200 kb in size (one enzyme, *Nsi*I, also produces some weak, smaller fragments). The bands probably represent blocks of TA repeats from different telomeres with some flanking DNA extending to the restriction site in question. Weak bands contributing to numbers in excess of seven may be due to polymorphisms.

TA repeat subfamilies. The TA repeat unit is 339 or 340 bp long, with two pairs of alternating subrepeats, separated by four linker regions which are nonrepetitive within the unit (Fig. 2). There is one master unit, M1, from which three other units, D1, D2, and D3, are derived, in each case by one cluster of mutations in the parental M1 segment. The mutated segments are ds1, ds2, and ds3, respectively. The D3 unit, like D2 1 bp shorter than the others, is somewhat more complex, since linker region L4 is converted to L2. The ds3 is localized in this converted linker region. Most derived units contain one ds region each, but there are also a few complex units with two or three regions (including an unusual ds4) (13). DNA that hybridized with both ds1 and ds2 oligonucleotides was found in 3L and 4R, whereas sequences that hybridized with ds3 oligonucleotide were found consistently in 1R, 3R, 3L, and 4R and irregularly in 1L, 2L, and 2R.

Intertelomeric mapping. Three subfamilies, D1, D2, and D3, have previously been localized to individual chromosome ends in our local stock (Lund) of *C. pallidivittatus* after in situ oligonucleotide hybridizations carried out in 1992 (14). We have now found that it is possible to apply such a procedure to localization of M1. In this case the oligonucleotide, a 51-mer, covers a segment containing all degenerate segments, i.e., ds1, ds2, and ds3. Figure 3B shows that the M1 oligonucleotide hybridizes five chromosome ends, 1L, 2L, 2R, 3L, and 4R, of the seven containing TA repeats (Fig. 3A). M1 units have been

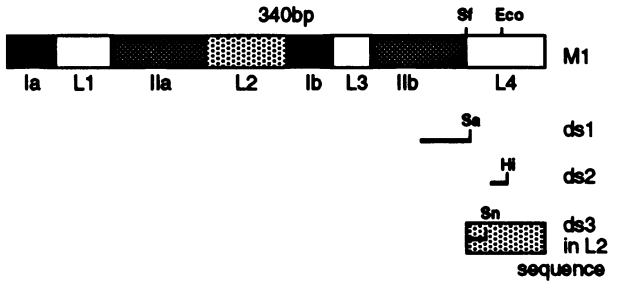

FIG. 2. Schematic representation of different TA repeats with relevant restriction sites. The 340-bp M1 unit is shown with two pairs of subrepeats, Ia and Ib and IIa and IIb, and linker regions, L1 to L4, nonrepetitive within the unit. The positions of ds1, ds2, and ds3 in relation to M1 are indicated. It is also shown that ds3 is usually present in an L2 sequence in the position of L4. Unique restriction sites in ds regions or corresponding parts of M1 are shown: Sf, *Sfa*NI; *Eco*, *Eco*RI; Sa, *Sau*3A; Hi, *Hin*fI; Sn, *Sna*BI.

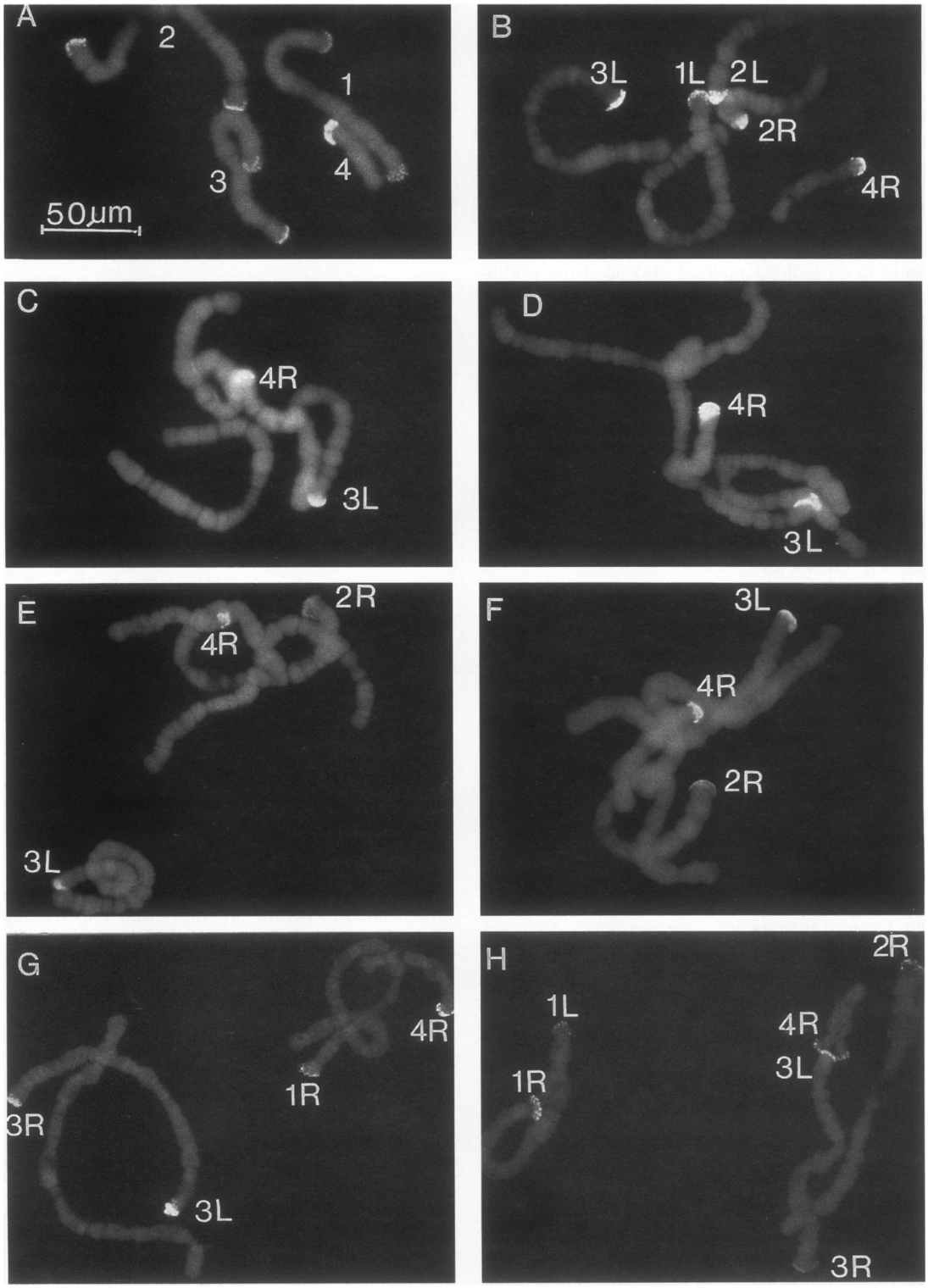

FIG. 3. In situ hybridization of digoxigenin-labeled probes to salivary gland chromosomes from *C. pallidivittatus*. (A) The probe is the whole TA repeat; (B) the probe is M1 oligonucleotide; (C and D) the probe is ds1 and ds2 oligonucleotide, respectively, applied to animals of the year 1992 type; (E and F) the probe is the ds1 and ds2 oligonucleotide, respectively, applied to animals of the year 1993 type; (G and H) the ds3 oligonucleotide is used for different animals of the year 1992 type, showing hybridization in four and six telomeres, respectively.

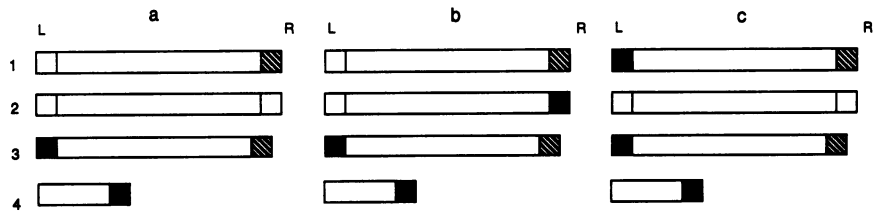

FIG. 4. Arrangement of different, constantly present subpopulations of TA repeats in the telomeres of *C. pallidivittatus*. (a) Diagram of animals of the year 1992 type (Lund stock); (b) the predominant type of the year 1993 animals (Lund stock); (c) pattern for the Göttingen stock. Telomeres indicated as specific for M1 (□) may or may not contain D3 as well in in situ hybridizations. Other symbols: ▧, D3; ■, M1, D1, D2, and D3.

recovered in these five ends by microcloning (12, 13). In the 1992 investigation of the Lund stock (14), D1 and D2 were found only in 3L and 4R. In the present study, 1 year later (1993), the majority of larvae from the same perpetuated culture had D1 and D2 not only in 3L and 4R but also in 2R (Fig. 3E and F), with a minority showing the year 1992 pattern (Fig. 3C and D). Furthermore, another stock (Göttingen) had D1 and D2 in a third pattern, namely, in 3L, 4R, and 1L (Fig. 4). In summary, therefore, three different patterns were observed for D1 plus D2, two subfamilies which always appear together. In all these cases M1 goes to the five telomeres 1L, 2L, 2R, 3L, and 4R. Without exception, D3 is the only subfamily present in 1R and 3R (Fig. 4).

D3 is consistently present, not only in 1R and 3R but also in the two or three chromosome ends that contain D1 plus D2. In the remaining two or three M1-containing telomeres, the presence of D3 is irregular (Table 1; Fig. 3G and H).

Isolation of subfamily blocks. To relate different subfamilies to one another within telomeres, we first had to learn if they are arranged in blocks. This can be done because of a high intrasubfamily sequence homogeneity (12, 13) and suitable restriction sites within ds regions as well as corresponding M1 segments. Useful sites are (Fig. 2) *Sfa*NI in M1 and D2, *Eco*RI in M1 and D1, *Sau*3A in D1, *Hin*fI in D2, and *Sna*BI in D3. Thus, *Sau*3A digests arrays of D1 to monomers. *Hin*fI cuts D2, and *Sna*BI cuts D3. Obviously, M1 units cannot be cleaved specifically. *Sfa*NI restricts both M1 and D2, whereas *Eco*RI isolates D1 in addition to M1. In practice this is a minor disadvantage because D1 is hardly present in blocks and *Eco*RI, therefore, is specific enough for M1. Figure 5, lanes a to d, shows how this application gives strong monomer bands for M1, D2, and D3, suggesting that these subfamilies, but not D1, are at least partially in tandem arrays.

Of greater value, however, is isolation of blocks of different subfamilies by digesting all except one repeat type. We can isolate arrays of M1 with *Sau*3A, *Hin*fI, plus *Sna*BI; D1 with *Sfa*NI, *Hin*fI, plus *Sna*BI; D2 with *Eco*RI, *Sna*BI, plus *Sau*3A; and D3 with *Eco*RI plus *Hin*fI plus Sau3A. Figure 5, lanes e to h, shows that this application supplements single-enzyme digestions. M1, D2, and D3 appear with strong bands at 9 kb (the resolving limit of the gel), in contrast to D1, with only a few weak bands larger than the monomer. Again, this suggests that M1, D2, and D3, but not D1, are in arrays, at least partially, as predicted from single-enzyme digestions.

Intratelomeric mapping. To localize subfamily repeats in relation to each other, we labeled one probe with biotin and another with digoxigenin. Avidin-fluorescein (FITC) complex and antibody conjugated with rhodamine were used for visualization. The fluorescein signal was reinforced with biotinylated anti-avidin. A double-pass fluorescence filter permitted simultaneous visualization of different subfamilies. The double-probe studies were done on year 1993 animals of the Lund stock.

With rhodamine to detect D3 labeled with digoxigenin and fluorescein for M1 marked with biotin, there is a distinct zone of rhodamine-generated red fluorescence distal to the green-yellow emission due to fluorescein (Fig. 6A and B). D3 had a similar relation to D1 (Fig. 6E) and D2 (Fig. 6D). With inverse use of fluorophors, an inverse color pattern was obtained. The zones are, however, less distinct, and higher magnification is necessary for clear results (Fig. 6C). D1, D2, and M1 were not spatially resolved from one another in any combination. In

TABLE 1. Hybridization to a ds3 oligonucleotide in individual salivary gland telomeres from nine animals of the Lund 1992 stock

Animal no.	Presence of D3 in:						
	1R	1L	2R	2L	3R	3L	4R
1	+	−	+	−	+	+	+
2	+	−	+	−	+	+	+
3	+	−	+	−	+	+	+
4	+	−	+	−	+	+	+
5	+	−	+	+	+	+	+
6	+	+	+	−	+	+	+
7	+	−	−	−	+	+	+
8	+	−	−	−	+	+	+
9	+	+	+	−	+	+	+

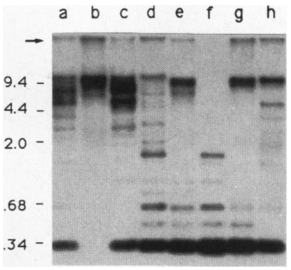

FIG. 5. Southern blot of digestions of genomic DNA from *C. pallidivittatus*, separated in a 1% agarose gel, hybridized with ^{32}P-labeled M1 repeat unit. Lanes a to d show single-enzyme digestions in which units in tandem arrays are isolated as 340-bp monomers. M1 (+D1) is in the monomer band in lane a after digestion with *Eco*RI; D1 in lane b has been isolated with *Sau*3A; D2 in lane c has been released with *Hin*fI; and D3 in lane d has been released with *Sna*BI. Lanes e to h show multiple digestions that release all repeats to monomers, except those belonging to one specific subfamily, which are retained in tandem arrays. In lane e, M1 blocks are isolated after digestion with *Sau*3A, *Hin*fI, and *Sna*BI; in lane f, D1 is isolated after *Sfa*NI, *Hin*fI, and *Sna*BI digestion (little D1 is present in tandem form); in lane g, D2 is isolated after *Eco*RI, *Sna*BI, and *Sau*3A digestion; and in panel h, D3 is isolated after *Eco*RI, *Hin*fI, and *Sau*3A digestion. The arrow shows the application wells.

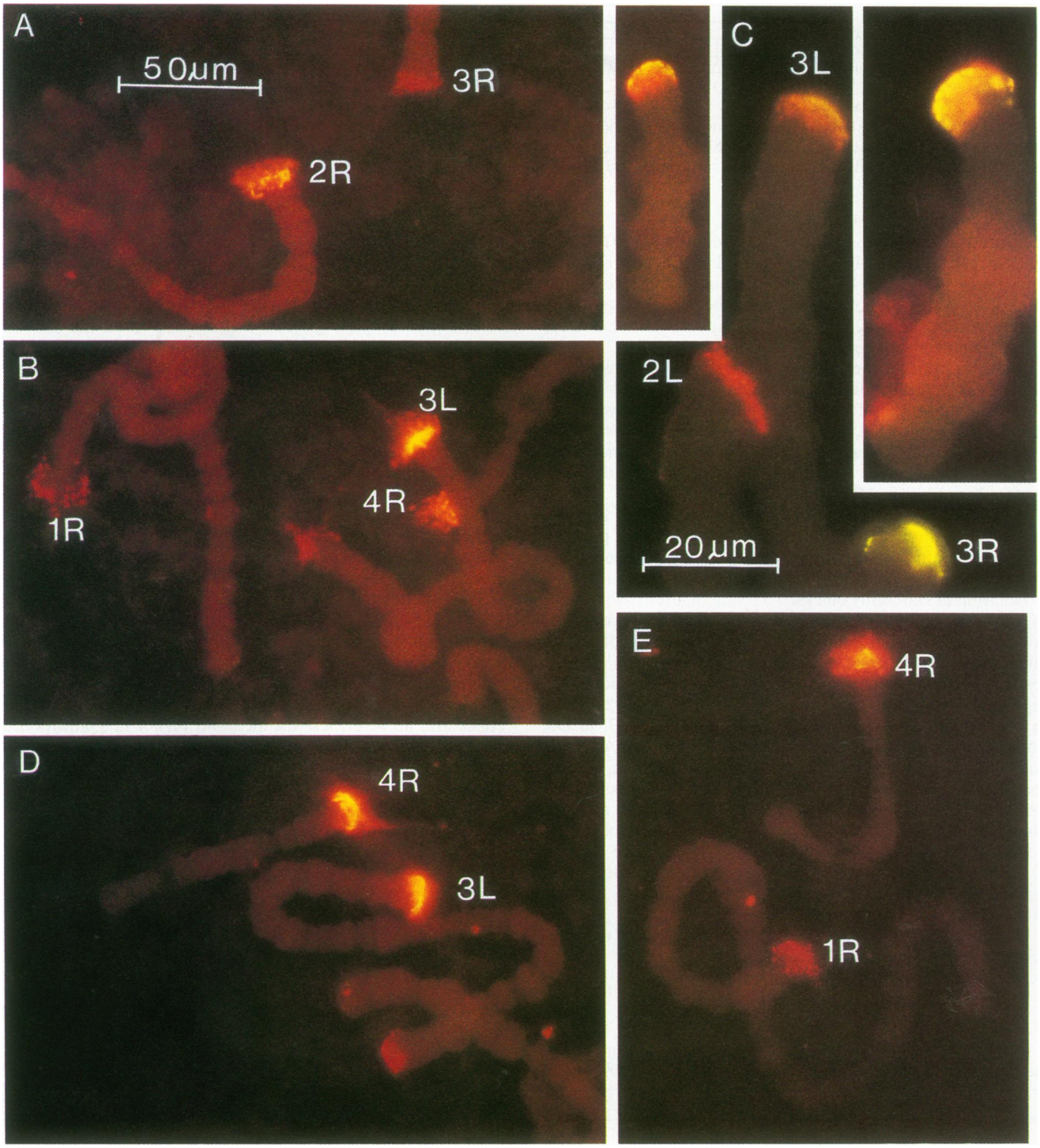

FIG. 6. Double-probe in situ hybridizations of TA repeat subfamilies. (A, B, D, and E) The ds3 oligonucleotide is labeled with digoxigenin and assayed with rhodamine-conjugated anti-digoxigenin antibody. For panels A and B simultaneous hybridization with biotin-labeled M1 oligonucleotide is performed and assayed with fluorescein in an indirect procedure. (C) Labeling procedures of the M1 and D3 oligonucleotides are reversed but other conditions are the same. The inserts in panel C show chromosome 4, also after reverse labeling at the same magnification as for the main illustration (right insert) and at a magnification intermediate between those of panels A and C (left insert). The ds2 oligonucleotide labeled with biotin is used as the second hybridization probe in panel D, and the ds1 oligonucleotide labeled with biotin is used in panel E. The magnification for panel A also applies to panels B, D, and E.

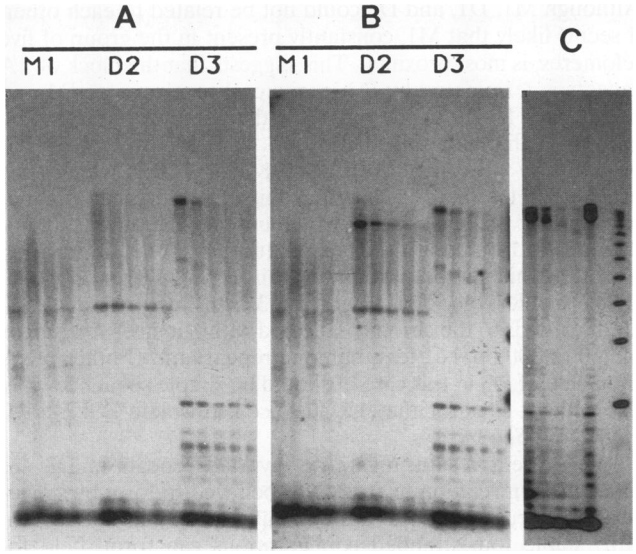

FIG. 7. (A) Exonuclease digestions of TA repeat blocks from genomic DNA of *C. pallidivittatus* (Lund stock). Three different reactions were run, M1, D2, and D3. In each reaction 5 μg of high-molecular-weight DNA was digested in 150 μl of buffer (600 mM NaCl, 12 mM $CaCl_2$, 12 mM $MgCl_2$, 20 mM Tris-Cl buffer [pH 8.0], 0.2 mM EDTA) (at 30°C with 2.1 U of *Bal* 31 per 100 μl). After 0, 30, 60, 90, and 120 min, 30-μl aliquots were withdrawn and added to 30 μl of 20 mM ethylene glycol-bis(β-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid (EGTA). Each of the 15 samples was then diluted with 120 μl of 50 mM Tris-Cl buffer (pH 7.4)–15 mM $MgCl_2$–1.5 mM dithiothreitol, after which the M1 samples were digested with *Hin*fI, the D2 samples were digested with *Eco*RI, and the D3 samples were digested with *Eco*RI-*Hin*fI. All samples were extracted with phenol-chloroform, precipitated, dried, and redissolved in 50 μl for digestion with *Sau*3A. The M1 and D2 samples were finally digested with *Sna*BI. After digestion, all samples were treated with phenol-chloroform, precipitated, dried, and redissolved for electrophoresis in 0.3% agarose, as described in the legend to Fig. 1; this was followed by hybridization to ^{32}P-labeled M1 repeat unit. (B) The filter has been rehybridized with labeled Balbiani ring 2-1 core repeat unit (21). The X-ray film from panel A is superimposed on the film from this hybridization. (C) D3 obtained from a second batch of animals from the same stock and digested with *Bal* 31 at a similar concentration as in panel A but with DNA samples withdrawn after 0, 60, 120, and 180 min. The rightmost genomic DNA separation is a control without *Bal* 31. A 5-kb ladder is shown to the right.

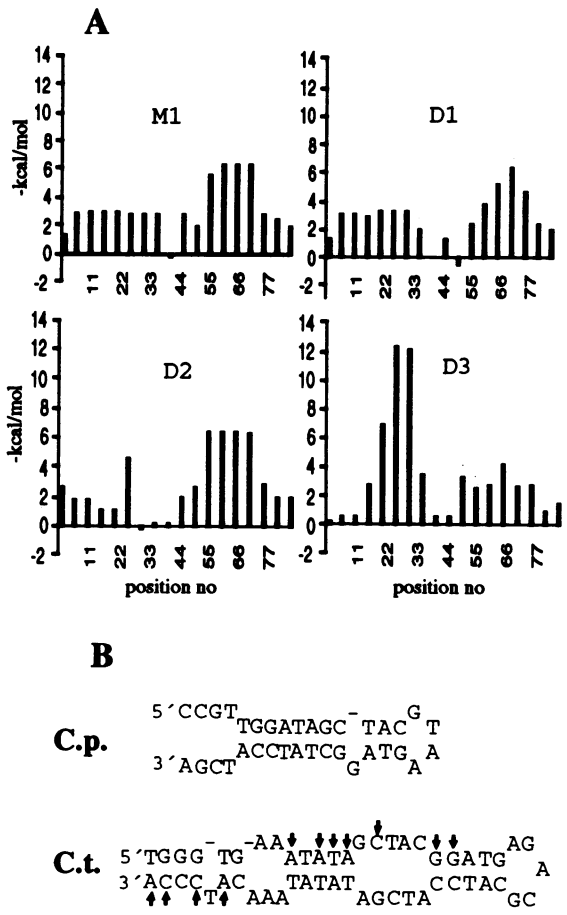

FIG. 8. (A) Scan along the G-rich strand (as defined in reference 36) of M1, D1, D2, and D3, from the end of linker region L4 in the reverse direction. The free energy of hypothetical hairpins is given within a window of 35 nt. Position 1 is the one in L4 that lies closest to subrepeat IA and that would have numbers 340 or 339 in earlier representations (12, 13, 43), in which C strands are given. (B) Secondary structures of the G strand in D3 (C.p.) at the peak of the scan and a 53-nt segment of the *C. tentans* repeat unit (C.t.) at the corresponding position. Differences in *C. tentans* from the M1 region (positions 265 to 318, in inverse, from pCp 306 [43]) in *C. pallidivittatus* are indicated by arrows.

summary, therefore, D3 appears more peripheral than all the other subfamilies.

***Bal* 31 digestions of subfamily blocks.** *Bal* 31 digestions were done with DNA from the Lund stock. Blocks of M1, D2, and D3 were produced after treatment of genomic DNA with *Bal* 31. The isolated arrays were separated in a 0.3% agarose gel and hybridized with labeled repeat unit. The results in Fig. 7A show a number of discrete bands in all three subfamilies. In D3 there are bands at 20 and 55 kb, digested by *Bal* 31. There are at least two fractions in the 55-kb band as shown by other separations. Plots of the size reduction of the 20-kb band as a function of digestion time show that it is linear from zero time within a margin of 200 to 300 bp. Occasionally, weak exonuclease-sensitive bands were also seen in M1 and D2 (results not shown).

To check that the size reduction is no artifact, the filter was rehybridized with core repeats of the Balbiani ring 2-1 gene (21). Figure 7B shows a ≈50-kb exonuclease-resistant band in the D2 separation.

Figure 7C shows DNA with a D3 separation from DNA of another batch of animals, qualitatively somewhat different from the one in Fig. 7A, with longer *Bal* 31 treatment to show exonuclease sensitivity of the 50- to 60-kb region.

Structure of the D3 unit. Evidence that D3 is usually most peripheral among the TA repeats or the exclusive component of such blocks prompts the question of the features that are specific for D3. Computer analysis of the region unique for D3 was performed with the MFold program of Zuker (53) (modified to work with the Genetics Computer Group package) to predict optimal and suboptimal secondary structures. The results were expressed as free energy for a hypothetical single strand. A window with an arbitrary size of 35 nucleotides (nt) was used to scan the whole D3 unit. The most stable hairpin (ΔG = −12.3 kcal/mol [−51.5 kJ/mol] at 37°C) encompassed a 27-nt region including ds3. A larger window will not increase the value. In the corresponding position, M1, like D1 and D2, would form considerably less-stable hairpins (−3.2 kcal/mol [−13.4 kJ/mol] for M1) (Fig. 8A). The scan showed a second

region at half the repeat distance, giving hypothetical hairpins with a somewhat lower negative peak value for free energy over the corresponding stretch in IIA and L2, which is common for all repeats (−9.7 kcal/mol [−40.6 kJ/mol]). Other regions providing less-stable hypothetical hairpins were also seen.

TA repeats in *C. tentans* have no known subfamilies (37). Here, however, the corresponding region in IIB-L4 also has dyad symmetry. In hypothetical single-strand form a hairpin would form over a 53-nt segment (−9.2 kcal/mol [−38.5 kJ/mol]) (Fig. 8B). The *C. tentans* region does not have the L2-for-L4 substitution or the ds3 stretch, but it has a cluster of mutations in IIB (37). The M1, like the corresponding *C. tentans* IIB-L4 region, has evolved from a common ancestor; during evolution, a total of 11 base differences have appeared in this stretch of 53 nt. Ten of these differences give a hydrogen-bonded base pair in *C. tentans* (Fig. 8B), and the remaining one does not lie in a base-paired region. These results suggest first that there is a selective pressure for a palindromic sequence in this region of the TA repeats and second that there is no selection for such a region when repeats have only a more interior localization.

DISCUSSION

Like in *D. melanogaster*, telomeric DNA has not been found in *Chironomus* species. We have looked for $TTAGGG_n$ (7, 35), $TTAGG_n$ (38), $TTGGGG_n$ (8), and the *Schizosaccharomyces pombe* $T_{1-2}ACA_{0-1}C_{0-1}G_{1-6,n}$ (see reference 51 and references therein) repeats and other variants (36), with negative results. Telomere-associated DNA may be derived from short repeats similar to telomeric repeats (36), and it remains to be decided if it also fulfils functions of such DNA. Two questions are of interest in this context, i.e., whether such repeats reach the chromosome termini and whether they help to maintain chromosome integrity.

We find telomere-associated DNA in 50- to 200-kb blocks, probably not interrupted by other DNA, since high-molecular-weight blocks can be isolated with several restriction enzymes with 6- or 4-base recognition sequences. The numbers of TA repeats in a telomere are highly variable (14), a property shared with similar DNA in other organisms; in several cases this is attributed to recombination between chromosome ends (15, 18, 20, 28–30).

It has been established for a number of cases, including human chromosomes (10, 17), that there is an intertelomeric differentiation in distribution of telomere-associated DNA, like for Y′ repeats in *Saccharomyces cerevisiae*, which can vary in numbers between chromosome ends in a strain- and species-dependent pattern (24, 25, 27, 52). Also in *D. melanogaster* there are examples of differential distribution of subfamilies of telomere-associated DNA (16, 26, 41, 46, 49). In *C. pallidivittatus* this differentiation is distinct. Two of the seven nontelocentric chromosome ends contain only D3, whereas M1 is restricted to the remaining five ends. Furthermore, the presence of M1 is correlated with an ability to carry a block with D1, D2, and D3 repeats, with a variable distribution between M1-containing telomeres. The genetic background for this variation is not yet known. The M1-containing ends lacking the D1 to D3 block vary in their content of D3, which here is present in levels down to nondetectable by in situ hybridization. The amounts of this repeat are highly variable in 1R and 3R as well (14) and probably in all seven ends. This variability is of a different kind from the previous variability, since it causes pronounced heterogeneity within stocks.

There is also an intratelomeric differentiation in *C. pallidivittatus*. D3 repeats are the most peripheral among the repeats. Although M1, D1, and D2 could not be related to each other, it seems likely that M1, constantly present in the group of five telomeres, is most proximal. This suggests that the block of TA repeats is terminated by D3, peripheral to a zone of D1 and D2, in turn lying outside of M1 in the complex chromosome ends. We find not only D3 blocks sensitive to *Bal* 31 but occasionally also weak bands of M1 and D2, which may originate in telomeres lacking D3 units. D1 is present at about one-third of the level (in molar amounts) of the other subfamily repeats (14) and is interspersed. It is always in combination with D2, which could partially be mixed with D1 units.

The smallest block of D3 with *Bal* 31 sensitivity was about 20 kb. It extends to the chromosome end within a measuring error of 200 to 300 bp, i.e., less than one repeat unit. If other DNA is outside of D3 in this case, it should be simple sequence type, i.e., without sites for the enzymes used to isolate D3: *Eco*RI, *Hin*fI, and *Sau*3A.

The present results regarding inverted repeats in D3 are interesting in relation to structures believed to form by single-strand protrusions of the G-rich strand in telomeric DNA. In vitro stable four-stranded configurations can form, held together by unconventional hydrogen bonds. It is not known whether these quadruplexes (G quartets) exist in vivo since they do not bind telomere-derived protein in vitro (40) or serve as substrate for telomerase (50). It has been shown, on the other hand, that a subunit of the *Oxytricha* telomere-binding protein catalyzes the formation of such structures, and it was proposed that this protein might also help to unfold four-stranded DNA, making it available for telomerase (19). One alternative to G quartets is a three-stranded structure (47), which can also explain some properties specific for chromosome ends; another alternative is the hairpin, which may provide more suitable stability and less resistance to telomerase action (11).

It is interesting that D3, close to the termini, has the best possibilities of the different repeats to form stable hairpins, in at least two regions per unit compared with once for the others. It is true that only part of the repeat unit would provide hypothetically favorable structures, in contrast to the short G-rich repeats. This could, however, have an interesting parallel in repeat units from a number of budding yeasts. McEachern and Blackburn (32) have shown that these species have telomeric repeats of variable lengths (8 to 25 bp), in which only a segment of 6 bp has a typical G+T-rich telomeric repeat sequence.

In the present case the postulated structure will provide potential approaches to obtain positive evidence for a specific terminal configuration. This may turn out to be particularly important for *C. pallidivittatus*, for which it may remain difficult to exclude entirely the presence of telomeric DNA, including its absence from germ cells.

ACKNOWLEDGMENTS

The present investigation was supported by the Swedish Cancer Society and the Philip-Sörensen Foundation.

We are grateful to Neal Sugawara, Waltham, Mass., for providing us with pSPT16 containing telomeric sequences and to Ulrich Grossbach, Göttingen, Germany, for sending us his stock of *C. pallidivittatus*.

REFERENCES

1. **Biessmann, H., L. E. Champion, M. O'Hair, K. Ikenaga, B. Kasravi, and J. M. Mason.** 1992. Frequent transpositions of *Drosophila melanogaster* HeT-A transposable elements to receding chromosome ends. EMBO J. **11**:4459–4469.
2. **Biessmann, H., and J. M. Mason.** 1992. Genetics and molecular biology of telomeres. Adv. Genet. **30**:185–249.

3. **Biessmann, H., J. M. Mason, K. Ferry, M. d'Hulst, K. Valgeirsdottir, K. L. Traverse, and M.-L. Pardue.** 1990. Addition of telomere-associated HeT DNA sequences "heals" broken chromosome ends in *Drosophila*. Cell **61**:663–673.
4. **Biessmann, H., K. Valgeirsdottir, A. Lofsky, C. Chin, B. Ginther, R. W. Levis, and M.-L. Pardue.** 1992. HeT-A, a transposable element specifically involved in "healing" broken chromosome ends in *Drosophila melanogaster*. Mol. Cell. Biol. **12**:3910–3918.
5. **Bingham, P. M., R. Levis, and G. M. Rubin.** 1981. Cloning of DNA sequences from the white locus of *D. melanogaster* by a novel and general method. Cell **25**:693–704.
6. **Blackburn, E. H.** 1991. Structure and function of telomeres. Nature (London) **350**:569–573.
7. **Blackburn, E. H., and P. B. Challoner.** 1984. Identification of a telomeric DNA sequence in *Trypanosoma brucei*. Cell **36**:447–457.
8. **Blackburn, E. H., and J. G. Gall.** 1978. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. J. Mol. Biol. **120**:33–53.
9. **Blackburn, E. H., and J. W. Szostak.** 1984. The molecular structure of centromeres and telomeres. Annu. Rev. Biochem. **53**:163–194.
10. **Brown, W. R. A., P. J. MacKinnon, A. Villasanté, N. Spurr, V. J. Buckle, and M. J. Dobson.** 1990. Structure and polymorphism of human telomere-associated DNA. Cell **63**:119–132.
11. **Choi, K.-H., and B.-S. Choi.** 1994. Formation of a hairpin structure by telomere 3′ overhang. Biochim. Biophys. Acta **1217**:341–344.
12. **Cohn, M., and J.-E. Edström.** 1991. Evolutionary relations between subtypes of telomere-associated repeats in *Chironomus*. J. Mol. Evol. **32**:463–468.
13. **Cohn, M., and J.-E. Edström.** 1992. Telomere-associated repeats in *Chironomus* form discrete subfamilies generated by gene conversion. J. Mol. Evol. **35**:114–122.
14. **Cohn, M., and J.-E. Edström.** 1992. Chromosome ends in *Chironomus pallidivittatus* contain different subfamilies of telomere-associated repeats. Chromosoma **101**:634–640.
15. **Corcoran, L. M., J. K. Thompson, D. Walliker, and D. J. Kemp.** 1988. Homologous recombination within subtelomeric repeat sequences generates chromosome size polymorphisms in *P. falciparum*. Cell **53**:807–813.
16. **Danilevskaja, O. N., D. A. Petrov, M. N. Pavlova, A. Koga, E. V. Kurenova, and D. L. Hartl.** 1992. A repetitive DNA element associated with telomeric sequences in *Drosophila melanogaster* contains open reading frames. Chromosoma **102**:32–40.
17. **de Lange, T., L. Shiue, R. M. Myers, D. R. Cox, S. L. Naylor, A. M. Killery, and H. E. Varmus.** 1990. Structure and variability of human chromosome ends. Mol. Cell. Biol. **10**:518–527.
18. **Dore, E., T. Pace, M. Ponzi, L. Picci, and C. Frontali.** 1990. Organization of subtelomeric repeats in *Plasmodium berghei*. Mol. Cell. Biol. **10**:2423–2427.
19. **Fang, G., and T. R. Cech.** 1993. The β subunit of *Oxytricha* telomere-binding protein promotes G-quartet formation by telomeric DNA. Cell **74**:875–885.
20. **Foote, S. J., and D. J. Kemp.** 1989. Chromosomes of malaria parasites. Trends Genet. **5**:337–342.
21. **Galler, R., H. Saiga, R. M. Widmer, M. Lezzi, and J.-E. Edström.** 1985. Two genes in Balbiani ring 2 with metabolically different 75S transcripts. EMBO J. **4**:2977–2982.
22. **Greider, C. W., and E. H. Blackburn.** 1985. Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. Cell **43**:405–413.
23. **Greider, C. W., and E. H. Blackburn.** 1989. A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. Nature (London) **337**:331–337.
24. **Horowitz, H., P. Thorburn, and J. E. Haber.** 1984. Rearrangements of highly polymorphic regions near telomeres of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4**:2509–2517.
25. **Jäger, D., and P. Philippsen.** 1989. Many yeast chromosomes lack the telomere-specific Y′ sequence. Mol. Cell. Biol. **9**:5754–5757.
26. **Levis, R. V., R. Ganesan, K. Houtchens, L. A. Tolar, and F.-M. Sheen.** 1993. Transposons in place of telomeric repeats at a *Drosophila* telomere. Cell **75**:1083–1093.
27. **Louis, E. J., and J. E. Haber.** 1990. The subtelomeric Y′ repeat family in *Saccharomyces cerevisiae*: an experimental system for repeated sequence evolution. Genetics **124**:533–545.
28. **Louis, E. J., and J. E. Haber.** 1990. Mitotic recombination among subtelomeric Y′ repeats in *Saccharomyces cerevisiae*. Genetics **124**: 547–559.
29. **Louis, E. J., and J. E. Haber.** 1992. The structure and evolution of subtelomeric Y′ repeats in *Saccharomyces cerevisiae*. Genetics **133**: 559–574.
30. **Louis, E. J., E. S. Naumova, A. Lee, G. Naumov, and J. E. Haber.** 1994. The chromosome end in yeast: its mosaic nature and influence on recombinational dynamics. Genetics **136**:789–802.
31. **Lundblad, V., and E. H. Blackburn.** 1993. An alternative pathway for yeast telomere maintenance rescues *Est-1* senescence. Cell **73**:347–360.
32. **McEachern, M. J., and E. H. Blackburn.** 1994. A conserved sequence motif within the exceptionally diverse telomeric sequences of budding yeasts. Proc. Natl. Acad. Sci. USA **91**:3453–3457.
33. **Morin, G. B., and T. R. Cech.** 1986. The telomeres of the linear mitochondrial DNA of *Tetrahymena thermophila* consist of 53 bp tandem repeats. Cell **46**:873–883.
34. **Morin, G. B., and T. R. Cech.** 1988. Mitochondrial telomeres: surprising diversity of repeated telomeric DNA sequences among six species of *Tetrahymena*. Cell **52**:367–374.
35. **Moyzis, R. K., J. M. Buckingham, L. S. Cram, M. Dani, L. L. Deaven, M. D. Jones, J. Meyne, R. L. Ratliff, and J.-R. Wu.** 1988. A highly conserved, repetitive DNA sequence (TTAGGG)n, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. USA **85**:6622–6626.
36. **Nielsen, L., and J.-E. Edström.** 1993. Complex telomere-associated repeat units in members of the genus *Chironomus* evolve from sequences similar to simple telomeric repeats. Mol. Cell. Biol. **13**:1583–1589.
37. **Nielsen, L., E. R. Schmidt, and J.-E. Edström.** 1990. Subrepeats result from regional DNA conservation in tandem repeats in *Chironomus* telomeres. J. Mol. Biol. **216**:577–584.
38. **Okazaki, S., K. Tsuchida, H. Maekawa, H. Ishikawa, and H. Fujiwara.** 1993. Identification of a pentanucleotide telomeric sequence, $(TTAGG)_n$, in the silkworm *Bombyx mori* and in other insects. Mol. Cell. Biol. **13**:1424–1432.
39. **Pluta, A. F., and V. A. Zakian.** 1989. Recombination occurs during telomere formation in yeast. Nature (London) **337**:429–433.
40. **Raghuraman, M. K., and T. R. Cech.** 1990. Effect of monovalent cation-induced telomeric DNA structure on the binding of *Oxytricha* telomeric protein. Nucleic Acids Res. **18**:4543–4552.
41. **Renkawitz-Pohl, R., and S. Bialojan.** 1984. A DNA sequence of *Drosophila melanogaster* with a differential telomeric distribution. Chromosoma **89**:206–211.
42. **Rovira, C., W. Beermann, and J.-E. Edström.** 1993. A repetitive DNA sequence associated with the centromeres of *Chironomus pallidivittatus*. Nucleic Acids Res. **21**:1775–1781.
43. **Saiga, H., and J.-E. Edström.** 1985. Long tandem arrays of complex repeat units in *Chironomus* telomeres. EMBO J. **4**:799–804.
44. **Scherthan, H., M. Köhler, P. Vogt, K. von Malsch, and D. Schweizer.** 1992. Chromosomal *in situ* hybridization with double-labeled DNA: signal amplification at the probe level. Cytogenet. Cell Genet. **60**:4–7.
45. **Traverse, K. L., and M. L. Pardue.** 1988. A spontaneously opened ring chromosome of *Drosophila melanogaster* has acquired He-T DNA sequences at both new telomeres. Proc. Natl. Acad. Sci. USA **85**:8116–8120.
46. **Valgeirsdottir, K., K. L. Traverse, and M. L. Pardue.** 1990. HeT DNA: a family of repeated sequences specific for heterochromatin in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **87**:7998–7982.
47. **Veselkov, A. G., V. A. Malkov, M. D. Frank-Kamenitskii, and V. N. Dobrynin.** 1993. Triplex model of chromosome ends. Nature (London) **364**:496.
48. **Walmsley, R. M., C. S. M. Chan, B.-K. Tye, and T. D. Petes.** 1984. Unusual DNA sequences associated with the ends of yeast chromosomes. Nature (London) **310**:157–160.
49. **Young, B. S., A. Pession, K. L. Traverse, C. French, and M. L.**

Pardue. 1983. Telomere regions in *Drosophila* share complex DNA sequences with pericentric heterochromatin. Cell **34:**85–94.
50. **Zahler, A. M., J. R. Williamson, T. R. Cech, and D. M. Prescott.** 1991. Inhibition of telomerase by G-quartet DNA structures. Nature (London) **350:**718–720.
51. **Zakian, V. A.** 1989. Structure and function of telomeres. Annu. Rev. Genet. **23:**579–604.
52. **Zakian, V. A., and H. M. Blanton.** 1988. Distribution of telomere-associated sequences on natural chromosomes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8:**2257–2260.
53. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244:**48–52.