Mol Gen Genet (1995) 246:29–36

ORIGINAL PAPER

Takanobu Higashiyama · Sinya Maki
Takashi Yamada

Molecular organization of *Chlorella vulgaris* chromosome I: presence of telomeric repeats that are conserved in higher plants

Received: 2 May 1994 / Accepted: 17 June 1994

Abstract The unicellular green alga *Chlorella vulgaris* (strain C-169) has a small genome (38.8 Mb) consisting of 16 chromosomes, which can be easily separated by CHEF gel electrophoresis. We have isolated and characterized the smallest chromosome (chromosome I, 980 kb) to elucidate the fundamental molecular organization of a plant-type chromosome. Restriction mapping and sequence analyses revealed that the telomeres of this chromosome consist of 5′-TTTAGGG repeats running from the centromere towards the termini; this sequence is identical to those reported for several higher plants. This sequence is reiterated approximately 70 times at both termini, although individual clones exhibited microheterogeneity in both sequence and copy number of the repeats. Subtelomeric sequences proximal to the termini were totally different from each other: on the left arm, unique sequence elements (14–20 bp) which were specific to chromosome I, form a repeat array of 1.7 kb, whereas a 1.0 kb sequence on the right arm contained a poly(A)-associated element immediately next to the telomeric repeats. This element is repeated several times on chromosome I and many times on all the other chromosomes of this organism.

Key words *Chlorella vulgaris*
Small-sized chromosomes
Pulsed field gel electrophoresis · Physical mapping
Telomeric repeats

Communicated by M. Sekiguchi

Takanobu Higashiyama · Sinya Maki · Takashi Yamada (✉)
Department of Fermentation Technology,
Faculty of Engineering, Hiroshima University,
Kagamiyama 1-4-1, Higashi-Hiroshima 724, Japan
EMBL/GenBank/DDBJ accession nos. D26372, D26373, D26374, and D26375

Introduction

Each chromosome of a eukaryotic organism contains the cis-acting DNA sequence elements required for chromosome maintenance, including functional centromeric DNA, chromosomal origins of DNA replication, and telomeric DNA. Most information about the molecular structure of these elements comes from the yeast *Saccharomyces cerevisiae* (Chan and Tye 1980; Clarke and Carbon 1980; Szostak and Blackburn 1982, Murray and Szostak 1983; Newlon 1988; Sambrook et al. 1989). Recent work, however, shows that the telomeric repeat 5′-TTTAGGG (as read toward the chromosome terminus) is common to several higher plants including *Arabidopsis thaliana*, maize, and tomato (Richards and Ausubel 1988; Ganal et al. 1991; Broun et al. 1992). A long tandem repeat array of a 180-bp sequence element is found at all five *A. thaliana* centromeres (Maluszynska and Heslop-Harrison 1991) and the presence of telomere-like sequences around centromeres of chromosome I of *A. thaliana* (Richards et al. 1991) has also been reported. The telomeric sequence of the unicellular green alga *Chlamydomonas reinhardtii* , 5′-TTTTAGGG, differs somewhat from the higher plant repeats (Petracek et al. 1990). In spite of these advances, the molecular organization of plant chromosomes is largely unknown because of their large size and structural complexity.

The unicellular green algal genus *Chlorella* includes a variety of species (Fott and Nováková 1969), some of which have served as model organisms in plant physiological and biochemical studies for several decades (Govindjee and Braun 1974). Recently, it was found that the genome of *Chlorella* strains consists of a set of relatively small chromosomes, which can be easily separated by pulsed-field gel electrophoresis (Higashiyama and Yamada 1991; Yamada 1993):for example, the karyotype established for *C. vulgaris* C-169 contains 16 chromosomes ranging from 980 kb to 4.0 Mb in size. To our knowledge, the 980-kb chromosome (chromosome I) of this organism is the smallest plant-type chromo-

some so far studied. Structural analysis of this chromosome should provide an excellent opportunity to elucidate the fundamental and detailed molecular arrangement of plant chromosomes. In this report, we have constructed a restriction map, made a genomic library, and cloned and sequenced both terminal regions, including telomeres and subtelomeric sequences of *C. vulgaris* chromosome I.

Materials and methods

Strains and growth conditions

Chlorella vulgaris (C-135, C-150, and C-169) and *C. saccharophila* (C-211) were obtained from the culture collection of the Institute of Molecular and Cellular Biosciences, University of Tokyo. *C. vulgaris* (211–11b) and *Chlorella.* sp. NC64A were from the Sammlung von Algenkulturen, Pflanzenphysiologisches Institut der Universität Göttingen (SAG), Germany and from J. L. Van Etten, University of Nebraska at Lincoln, USA, respectively. Cells were cultured photosynthetically in modified Bristol medium (MBM) as described previously (Higashiyama and Yamada 1991).

DNA preparation and techniques

Algal DNAs (nuclear, chloroplast, and mitochondrial) were isolated and fractionated by ultracentrifugation in a CsCl-Hoechst 33258 gradient as described (Yamada 1982). DNA was digested with Bal31 and restriction enzymes and analyzed by Southern blot hybridization under standard conditions (Sambrook et al. 1989). To detect telomeric repeat sequences by Southern blot hybridization, a synthetic 24 mer deoxyoligonucleotide ([5′-(TTTTAGGG)$_3$]) that corresponds to the telomere sequence of *Chlamydomonas reinhardtii* (Petracek et al. 1990) was used as a probe. The probe was labeled with [γ-^{32}P]ATP (Amersham) and polynucleotide kinase (Takara Shuzo). For hybridizations using chromosome I DNA and its cloned fragments as probes, nonradioactive digoxygenin-dUTP labeling of the probe was carried out with a Boehringer kit according to the manufacturer's manual. Hybridization was performed in a mixture containing 50% formamide, 5 × SSC (1 × SSC is 0.15 M NaCl plus 15 mM sodium citrate), 5% blocking reagents (Boehringer Mannheim), 0.1% Sarkosyl, and 0.02% sodium dodecyl sulfate (SDS) for 20 h either at 42°C (high stringency) or at 30°C (low stringency).

Construction of genomic libraries

A genomic library of *C.vulgaris* C-169 nuclear DNA (nDNA) was constructed as follows: chromosomal DNA, separated by ultracentrifugation in a CsCl-Hoechst 33258 gradient, was partially digested with *Mbo*I (fragments 9–20 kb in size) and ligated to *Bam*HI-digested λDASHII arms (Stratagene). After packaging with GigapackII Gold packaging extract (Stratagene), the phages were grown in *Escherichia coli* SRB(P2). The preparation yielded approximately 2×10^7 pfu/μg DNA (Yamada et al. 1993). A library enriched for telomeric sequences was prepared by treating chromosomal DNA successively with Bal31 and *Pst*I, ligating it to *Sma*I + *Pst*I-digested pUC19, and transforming it into *E. coli* SURE (Toyobo) cells. Bal31 digestion was as follows: *Chlorella* chromosomal DNA (10 μg) in 500 μl of 1 × Bal31 buffer (20 mM TRIS-HCl pH 7.2, 0.6 M NaCl, 12.5 mM $MgCl_2$, 12.5 mM $CaCl_2$, 1 mM EDTA) was treated with 5 U of Bal31 (Takara) at 28°C for 1 min. Under these conditions, approximately 50–100 bp were degraded from each DNA fragment. The reaction was terminated by adding EGTA to a final concentration of 50 mM. DNA samples were phenol-extracted and precipitated with ethanol before digestion with restriction enzyme. The *Hin*dIII fragments derived from the chromosome I termini were cloned as follws: after treatment with Bal31 (1.2 U per 0.1 μg DNA) at 28°C for 45 sec, chromosome I DNA was digested with *Hin*dIII and ligated to pUC19 digested with *Hin*dIII and *Sma*I. The resulting plasmids were introduced into, and amplified in *E. coli* SURE cells. Positive clones were screened and analyzed by colony hybridization with labeled pCHt-2 as a probe.

Pulsed-field gel electrophoresis

Chromosomal DNA molecules of *Chlorella* cells were separated by pulsed-field gel electrophoresis (CHEF) as described previously (Higashiyama and Yamada 1991). CHEF (contour-clamped homogeneous electric field) gel electrophoresis was carried out in a 1% agarose (GTG, Takara) gel in 0.5 × TBE (45 mM TRIS, 45 mM boric acid, 1 mM EDTA at pH 8.3) with a 12 min switching interval at 2 V/cm for 4 days using a CHEF-DRII system (Bio-Rad) to separate all 16 chromosomes. To isolate chromosome I, CHEF gel electrophoresis was performed in a 1% low-melting-point agarose (InCert agarose, Takara) gel in the same buffer with a 3 min swiching interval at 5V/cm for 24 h.

For digestion with restriction enzymes, gel blocks containing chromosome I were cut from the CHEF gel, washed successively in $T_{10}E_{0.1}$ (10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA) containing 0.2% SDS, and 1 mg/ml proteinase K at 42°C for 1.5 h and in $T_{10}E_{0.1}$ containing 1 mM phenylmethanesulfonyl fluoride (PMSF, Sigma) at 30°C for 2 h. The blocks were transferred into restriction enzyme buffers (300 μl) and incubated at room temperature for 1 h before addition of restriction enzyme. For complete digestion, the gel block (0.5 cm × 1.0 cm) in 300 μl buffer was incubated with 100 U of restriction enzyme for 6 h followed by addition of another 100 U of enzyme and further incubation for 6 h. For partial digestion, the gel block was treated in the same way, except that the digestion was for 6 h with 50 U of restriction enzyme. The digestion was stopped by adding 3 μl of 0.5 M EDTA. DNA fragments were separated by CHEF gel electrophoresis.

DNA sequencing and analysis

Restriction fragments of both terminal regions of chromosome I were cloned into M13 mp18 and mp19, respectively. Single-stranded DNA was sequenced by the chain termination procedure with a kit (Auto Read sequencing kit, Pharmacia) using an Automated Laser Fluorescence (ALF) DNA sequencer (Pharmacia). Some fragments were sequenced with a Taq DNA sequencing kit (Takara) and a terminally ^{32}P-labeled primer to obtain direct ladder patterns.

Results

Isolation and mapping of chromosome I of *C. vulgaris*

Chromosome I of *C. vulgaris* C-169 was separated by CHEF gel electrophoresis as described in Materials and methods. Gel blocks containing chromosome I were cut out of the gel, washed , and treated with several rare-cutter restriction enzymes, including *Not*I, *Sfi*I, *Sse*8387I, *Srf*I, and *Swa*I. After CHEF gel electrophoresis, only the NotI-digested sample gave a well resolved fragmentation pattern (Fig. 1A). Six fragments were resolved here, with sizes of 440 kb, 220 kb (doublet), 70 kb, 30 kb, and 9 kb. The sum of these fragments coincided well with the total size of this chromosome (980 kb). To establish linkage among these fragments, *Not*I partial digests of chromosome I were probed with a few DNA

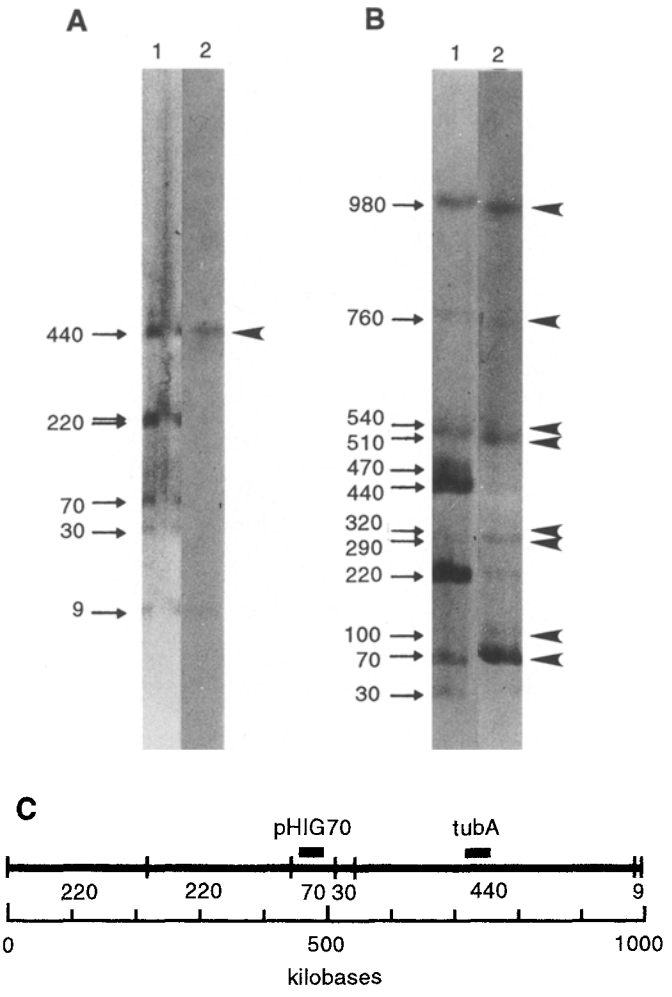

Fig. 1A–C Physical mapping of *C. vulgaris* chromosome I. **A** *C. vulgaris* chromosome I DNA isolated by CHEF gel electrophoresis was completely digested with *Not*I, size-fractionated by CHEF gel electrophoresis, and hybridized with digoxiygenin-labeled chromosome I DNA (lane 1) and the *tubA* gene clone (Higashiyama and Yamada 1991) (lane 2) as a probe. The arrowhead indicates the hybridizing band. **B** *C. vulgaris* chromosome I DNA was partially digested with *Not*I, size-fractionated by CHEF gel electrophoresis and, hybridized with digoxygenin-labeled chromosome I DNA (lane 1) and pHIG70 (Higashiyama and Yamada 1991) (lane 2) as a probe. *Arrowheads* indicate the bands used to construct a linkage map. **C** *Not*I restriction map of *C. vulgaris* chromosome I. Putative positions for *tubA* and pHIG70 are indicated by boxes. Sizes are shown in kilo bases

clones that had previously been shown to be encoded on this chromosome (Higashiyama and Yamada 1991). As shown in Fig. 1B, eight additional bands were discernible in *Not*I partial digests; these were 980 kb, 760 kb, 540 kb, 510 kb, 470 kb, 320 kb, 290 kb, and 100 kb. Telomere clones (described below) hybridized to the 220 kb and 9 kb fragments, which indicated that these fragments derived from each terminus.

When clone pHIG70, which specifically hybridizes to the 70 kb *Not*I fragment, was hybridized to the *Not*I partial digests, 100 kb and 290 kb bands were seen (Fig. 1B). Based on these results, a physical map with five *Not*I sites could be constructed as shown in Fig. 1C.

pCHt-1:

AAGCTTGTGG ATGTGACCAC ATGTGGTGAA AGCCCATTCT AGCTCTTGAG GAATTATGTT 60

AGGGTTTAGG GTTTACATTT TAGGGTGTAG GGGTTAGGGT TTAGGGTTTA GGGTTTAGGG 120

TTTAGGGTTC AGGGTTTACG GTTTACGGTT TAGGGTTTAG GGTTTAGGGT TTAGGGTTTA 180

GGGTTTAGGG TTTAGGGTTT AGGGTTTAGG GTTTAGGGTT TAGGGGTTAG GGTTTAGGGT 240

TTAGGGTTTA GGGTTTAGGG TTTAGGGTTT AGGGTTTAGG GTTTAGGGTT TAGGGTTTAG 300

GGTTTAGGGT TTAGGGTTTA GGGTTTAGGG TTTAGGGTTT AGGGTTTAGG GTTTAGGGTT 360

TAGGGTTTAG GGTTTAGGGT

pCHt-2:

GATCGCGCCA TCGCTGGCTC GAAATGACTA TATTTTAGGG TTTAGGGTGT AGGGTGTAGG 60

GTTTAGGGTT TAGGGTTCAG GGTTCAGGGT TCAGGGTTCA GGGTTCAGGG TTTAGGGTTT 120

AGGGTTCAGG GTTCAGGGTT CCGGGTTCAG GGTTTAGGGT TTAGGGGTTA GGTTTAGGGT 180

TTAGGGTTAG GGTTAGGGTT TAGGGTTTAG GGTTTTGGTT AGGGTTTAGG GTTTAGGGTT 240

TAGGGTTTGG GTTTAGGGTT TAGGGTTTAG GGTTTTAGGG TTTAGGGTTT AGGGTTTAGG 300

GTTTTAGGGT TTTAGGGTTT AGGGTTTAGG GTTTAGGGTT TAGGGTTTAG GGTTTAGGGT 360

TTAGGGTTTA GGGTTT

Fig. 2 Nucleotide sequences of pCHt-1, and pCHt-2. The 5'-TTTAGGG motif characteristic of higher plant telomeric repeats is *underlined*

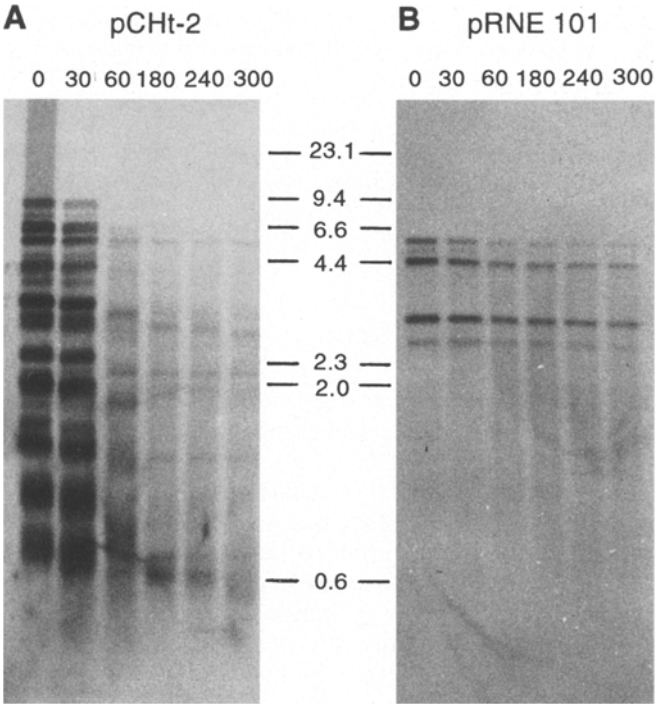

Fig. 3A,B The pCHt-2 DNA hybridizes to chromosomal termini of *C.vulgaris*. Chromosomal DNA of *C. vulgaris* was pretreated with nuclease *Bal*31 for 0, 30, 60, 180, 240, and 300 sec, as indicated at the top of the gel. The DNAs were then digested with *Pst*I, size-fractionated on an agarose gel and transferred to a nylon membrane. The filter was hybridized with digoxygenin-labeled pCHt-2 DNA **A** or pRNE 101 (Aimi et al. 1994); which contains an array of *Chlorella* rDNAs **B**. Sizes are shown in kilobases

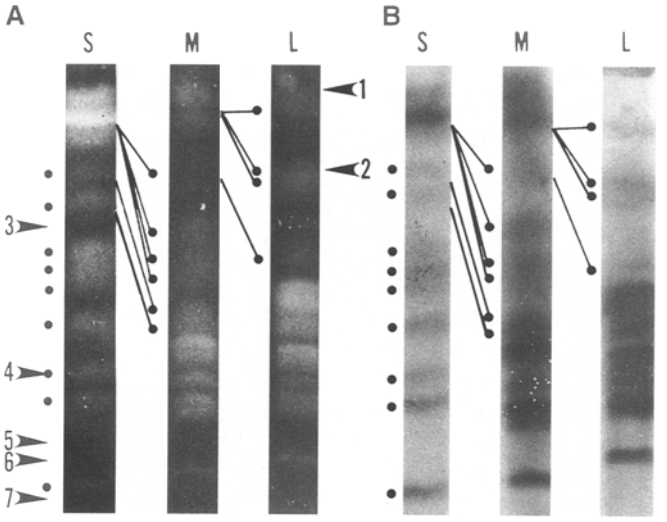

Fig. 4A,B The telomeric clone pCHt-2 hybridizes to all 16 chromosomes of *C. vulgaris*. **A** *C. vulgaris* chromosomal DNAs were separated by CHEF gel electrophoresis under three different conditions (Higashiyama and Yamada 1991): a combination of a 5-min switching interval at 3.3 V/cm for 24 h and a 8-min interval at 2.6 V/cm for 24 h to separate smaller chromosomes (S); a 12-min switching interval at 2 V/cm for 4 days to separate chromosome molecules with intermediate sizes (M) and a switching interval of 25 min at 1.6 V/cm for 3 days and with a 20-min interval at 2 V/cm for 2 days to separate larger molecules (L). The corresponding chromosomal DNA bands in different lanes are connected to each other by lines. *Arrowheads* indicate the positions of size markers of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* chromosomal DNAs: 1–7, 4600, 3500, 2200, 1600, 1125, 1020, and 945 kb, respectively. **B** Southern blot hybridization of the chromosomal DNAs in **A** with digoxygenin-labeled pCHt-2 DNA as a probe

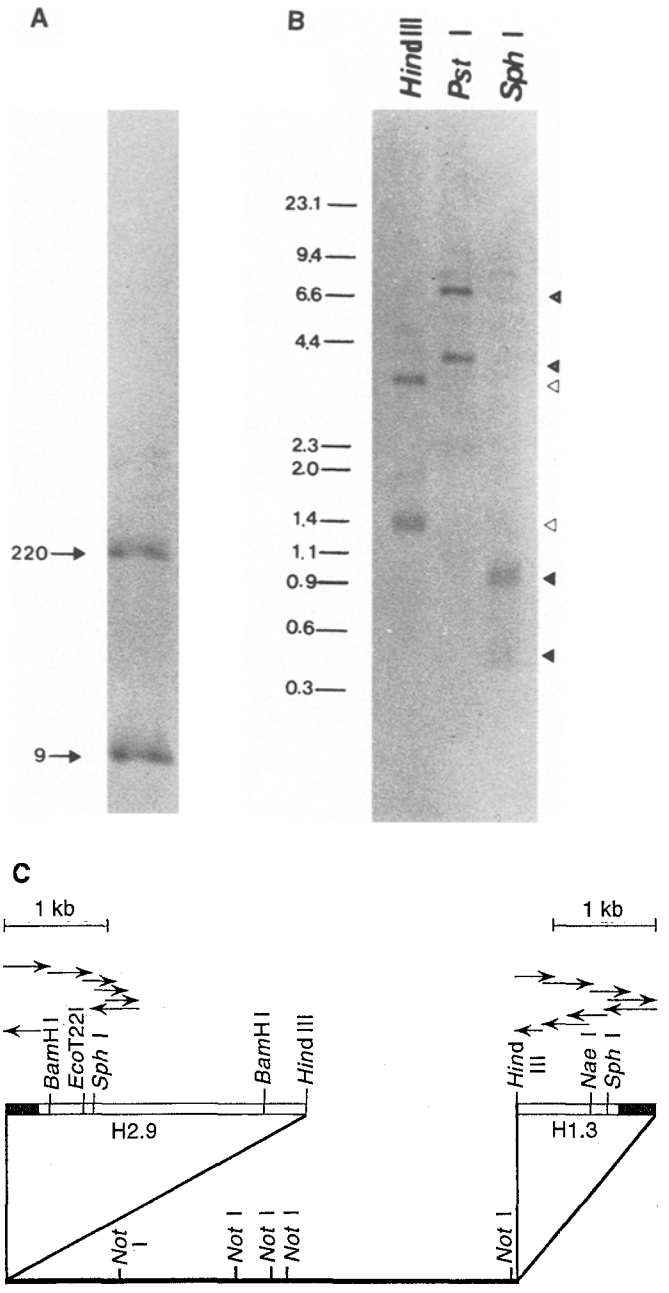

Fig. 5A-C Telomeric repeats of *C. vulgaris* chromosome I. **A** *Not*I fragments of chromosome I DNA were separated by CHEF gel electrophoresis as shown in Fig. 1 and probed with labeled pCHt-2 by Southern blot hybridization. Two strong hybridizing bands (9 kb and 220 kb) were discernible (*arrows*). Sizes are shown in. **B** Southern blot analysis of chromosome I DNA digested with *Hin*dIII, *Pst*I, or *Sph*I and probed with labeled pCHt-2. Hybridizing bands are indicated by *triangles*. Sizes are shown in kb. **C** Restriction maps of the 1.3 kb (H1.3) and 2.9 kb (H2.9) *Hin*dIII fragments derived from the *C. vulgaris* chromosome I termini. *Shaded* parts of H1.3 and H2.9 are the telomeric repeats. The sequencing strategy is shown above the map. Sizes are shown in kb

The *tubA* gene for α-tubulin (Yamada et al. 1993) was mapped to the 440 kb *Not*I fragment by Southern blot analysis (Fig. 1A).

Detection and cloning of telomeric fragments

The next step in the characterization of *Chlorella* chromosome I was to detect and clone its telomeric repeats. When *Pst*I digests of *Chlorella* nuclear DNA that had been pretreated with Bal31 nuclease for various periods were probed with a synthetic oligonucleotide ([5′-(TTTAGGG)$_3$]), corresponding to the telomeric repeat sequence of *Chlamydomonas reinhardtii* (Petracek et al. 1990), many hybridizing bands appeared (data not shown). They gradually disappeared with extended Bal31 digestion. However, the hybridizing bands were unexpectedly sharp and showed relatively weak hybridization. Some clones of these hybridizing fragments were obtained by colony hybridization under low stringency with the synthetic oligonucleotide probe from a terminal fragment library (see Materials and methods). Nucleotide sequences of two of these clones (pCHt-1 and pCHt-2) were determined and are shown in Fig. 2(B and C). Both clones contained, on the blunt-end side, a regular tandem array of the repeated sequence 5′-TTTAGGG instead of 5′-TTTTAGGG, which is exactly the same as the telomeric repeats reported for higher plants such as *Arabidopsis thaliana* and maize (Richards and Ausubel 1988). To confirm that these sequences rep-

resent genuine telomeric repeats of *C. vulgaris* chromosomes, Bal31 + *Pst*I-treated nDNA samples were hybridized with digoxygenin-labeled pCHt-2 as a probe: almost all hybridizing bands, which were usually broad but gave strong signals, gradually decreased in size with Bal31 digestion, and after 300 sec of Bal31 treatment no bands could be detected (Fig. 3). In contrast with this, the same nDNA probed with an *rrn* clone, pRNE101 (Aimi et al. 1993), gave four hybridizing bands whose sizes remained the same even after prolonged treatment with Bal31 (Fig. 3). The pCHt-2 probe hybridized to all 16 chromosomal DNA bands of *C. vulgaris* separated by CHEF gel electrophoresis (Fig. 4). These results indicated that the repeat sequences in pCHt-2 indeed represent *Chlorella* telomeric repeats.

Telomeric repeats of chromosome I

When *Not*I fragments of *C. vulgaris* chromosome I, separated by CHEF gel electrophoresis, were probed with pCHt-2 by Southern blot analysis, two bands of 9 kb and 220 kb gave strong hybridization signals (Fig. 5A). The result indicates that these bands derive from each terminus of chromosome I, as shown in Fig. 1. Further digestion of chromosome I DNA with *Pst*I, *Hin*dIII, and *Sph*I yielded pairs of hybridizing bands of 3.5 kb and 7.0 kb (*Pst*I), 1.3 kb and 2.9 kb (*Hin*dIII), and ≈0.5 kb and ≈0.9 kb (*Sph*I)(Fig. 5 B). There were no other discernible hybridizing bands under high stringency conditions, therefore these bands represent telomeric fragments from chromosome I. To determine the detailed terminal structure of chromosome I, we obtained 16 clones for the 1.3 kb *Hin*dIII fragment (H1.3) and 3 clones for the 2.9 kb *Hin*dIII fragment (H2.9). A unique portion of H1.3 hybridized with the 9.0 kb *Not*I fragment at the right terminus of chromosome I, as described below. A centromere-proximal part of H2.9 hybridized with the 220 kb *Not*I fragment from the left terminus (data not shown). Restriction maps and nucleotide sequencing strategies for these clones are shown in Fig. 5C.

Nucleotide sequences determined for the centromere-distal region of these clones (1284 bp for H1.3 and 1169 bp for H2.9) shown in Fig. 6, revealed, as expected, an extended repeat array of the telomeric sequences for both clones. Although the reiterated sequences were well conserved within and between the

Fig. 6A,B Nucleotide sequences of the telomeric repeats and flanking regions in both left and right arms of *C. vulgaris* chromosome I. **A** A left arm sequence (1169 bp) containing 32 copies of the telomeric sequence (*underlined*), a spacer sequence, and 16 copies of the subtelomeric repeat sequence (*dotted line*). **B** A right arm sequence (1284 bp) containing 52 copies of the telomeric sequence (*underlined*) and an adjacent poly(A)-type retroposon-like sequence. The poly(A) tail is *underlined* with a *wavy line*

A

GCCGTTCGAC AATCAAAGCC GCCGACAAAG AAAACCCTTC GACAATCAAG GCCGTCGACA ATGAAAGGCC GCCGCGACAA TTAAATGCCG TTGCGACAAT CAAAGTTGTG TCGACAATGA 120

AAGTTGTGTT GCGACAATCA AAGGCGCGCC GCGACATTCA AAGGCCCCGC GACAATGAAA TCCGTCGCGA CAATGAAATC CGTTGCGACA ATCAAAGGCG CGCGTGACAA TGAAATCCGT 240

CGCGGCAATG AAAGGGGGGC CGCGACAATC AAAGGCGCGC CGCGACAATA AAAGGCCGTT GCGACAATCA AAATGTTTAG GGAGGTTGGG GTTCAGTATC TAGCTTTGCC TCCCATGTGA 360

AGCCTTGCAT GCTGTAGGGT GGCATACCCC CTTGGCTAAG TTTCATATAT GACCCTTTTC AAGGGGTGAA TGCATGGCTT CAAGGGGATT CCATCTAATG GCACTTATGC CCAAGACCTA 480

CTTTCCCATG TTGGCATGTA ACTTGGCAAC ACTTGGCAAC ACTTGGCAAG ACTTGGCCTT GGGCTGCAGA CAACCACTGC CTTGGCTCAT ACTTGTTGTT TGGGATGTAT GAGACCCCGT 600

TGTGAGATTT GAGGGCCCTT GGACCCACTG CACATGCACA GCACATGCTC CAAGTGGGCA CAGGATGCAG CCATGCACCT GGCATATGTA TACACCATGC TGCACACCTG CACAGGGTGC 720

CGCTTCACTT GGCACCCATG CAGGATCATG TATGGGCCCC ATATGAGTCC TCCCAGCCAT GCCAAGGGTC TTGGATCCCC TGCACACCTC CAAGGCAACC CCATGTGCCT TGCAAGCCAC 840

TGCAAACCTT ATTTAAATTC ACACCCACCT CCATAAATTG TGAGCCACCT CTGGCAACCC ATAACATGGC CTCCTGCATG GGACTGTGAA TGTTTAGGGT TTAGGGTTTA GGGTTTAGGG 960

TTTAGGGTTT AGGGTTTAGG GTTTAGGGTT TAGGGTTAGG GTTAGGGTTT AGGGTTTAGG GTTTAGGGTT TAGGGTTTAG GGTTTAGGGT TTAGGGTTTA GGGTTTAGGG TTTAGGGTTT 1080

AGGGTTTAGG GTTTAGGGTT TAGGGTTTAG GGTTTAGGGT TTAGGGTTTA GGGTTTAGGG TTTAGGGTTT AGGGTTTAGG GT

B

AAGCTTTAGG CTATCAATAG TTTTCCAGTG ATCCTCCGAT GTTTCTGTCT CACCGGCAGC CCCAACAATC CCCTTGGCCT GTCATCCGTG CCTTCTGCCT TAGTGTCTTC CTTCACCCAA 120

TTCACAGTAG TTCATTGAAG GCTATCTGAA GGGTTTAGGT TTCCAACCCC AGTATCCACC CTTGAACCTG AGCGGGTGTC TAGGACCTTG GACATGATGC CTCCTGCCTA TCCCCGCCTC 240

CAATGTGGTG GGTGTGTGTA TCTGGTCGAG TTCTCGCCAG CCTACCTGTT TTTCGCCTCA TTTGCGGCTT CATATCTCCC ATATGCCACC GTTTGCTTGT GCCCCACCTT GCTGGCAGAT 360

TGAGGCATAG TTTAGGGTCA CATTGCCAGC CCCGAGCGGC ACCTAGTGGT AATGGGTCCG AACGCGATCC TGTTCCATCC CTGTCGGCCA ACCGTATAAA CGTGCCTTTC ATGTCTGTTC 480

CACCACTGGT TGCCCGCACA TCCATAATTG CCTGCCAGTG ATTGATGATT GCCTGCTTTG CTGTGACCCC TGCCTCCAGC TGCAAGCCCT AGGCTTGCCC CGTGGGTCCT CCAACGTCAC 600

CTCAGCGGTG TTGCGTTGCC GGCGGGCTGG AGTGCGCCTC TGGGGATGAG GTAACCCTCT GGTGCTTGGG ACATGTCCAT CGCTAGGAGC CTTTCTGTTG CCTTTCTGTG GTATGCCTTC 720

ATGCTTGCGC TTCTCTTAAG TATCGCATTA TCCTCTGCGG GAGTGCTCTG CGTGCCCCAA CTCTGACAGG TTTCCCCAAA TATCCTATTG CATGCGTGCT GGGAAGCGTG CGACTGGCAC 840

CTGTGTAATA GCGGGGGCTT CCCCTTTTAA AAAAAAAAAA AAAGGGTTTA GGGTTTAGGG TTTAGGGTTT AGGGTTTAGG GTTTAGGGTT TAGGGTTTAG GGTTTAGGGT TTAGGGTTAC 960

CTCATCCCCT ATATTTAGGG TTTTAGGGTT TTAGGGTTTT AGGGTTTAGG GTTTAGGGTT TAGGGTTTTT AGGGTTTAGG GTTTAGGGTT TAGGGTTTAT ATTCAGGGTT TAGGGTTTAG 1080

GGTTTAGGGT TTAGGGTTTA GGGTTTAGGG TTTAGGGTTT AGGGTTTAGG GTTTAGGGTT TAGGGTTTAG GGTTTAGGGT TTAGGGTTTA GGGTTTAGGG TTTAGGGTTT AGGGTTTAGG 1200

GTTTAGGGTT TAGGGTTTAG GGTTTAGGGT TTAGGGTTTA GGGTTTAGGG TTTAGGGTTT AGGGTTTAGG GTTTAGGGTT TAGG

two terminal fragments (H1.3 and H2.9), microheterogeneity was present in some positions, as indicated. This kind of microheterogeneity was even observed among different clones of the same fragment: the sequences of three different clones of H1.3 are compared in Fig. 7, in which occasional differences in the sequence and reiteration frequency of the telomeric repeats were seen. Since the clones of H1.3 and H2.9 were obtained after treatment with Bal31, they were missing some portions of the telomeric repeats. To determine the full size of the chromosome I telomeres, the digestion pattern of chromosome I DNA with *Sph*I was useful because a *Sph*I site is located very close to the telomere on both sides (Fig. 5C). The results shown in Fig. 5B indicated that both chromosome I telomeres have almost the same size; this ranged from 350–500 bp due to the size microheterogeneity described above, and thus corresponds to 50–70 copies of the telomeric sequence.

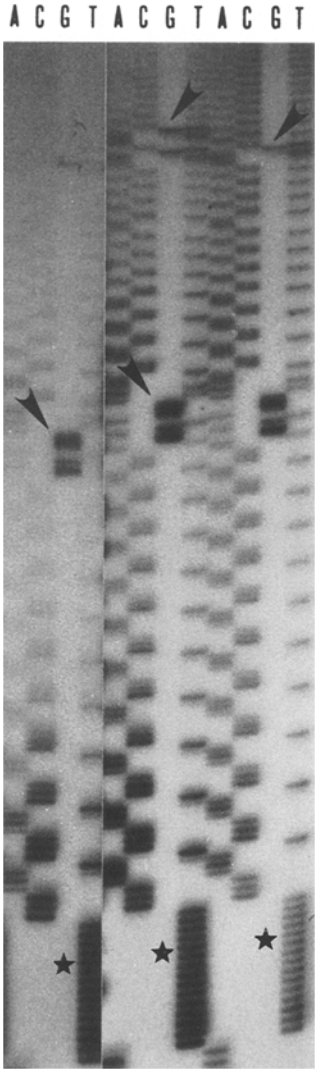

Fig. 7 Comparison of Nucleotide sequences of the *C. vulgaris* telomeric repeats in three different subclones of the 1.3 kb *Hin*dIII fragment (H1.3) from the right terminus. The subclones were obtained by digesting different H1.3 clones with *Sph*I and *Eco*RI and religating the fragment to the *Sph*I-*Eco*RI sites of M13mp18. The AC-rich strand was sequenced by the chain termination method with an end-labeled universal primer. The poly(A) [poly(T)] end of the retroposon is indicated by an *asterisk*. *Arrowheads* indicate regions of microheterogeneity among these clones

Subtelomeric or telomere-flanking sequences of chromosome I

As shown in Fig. 6, nucleotide sequences of the regions flanking the telomeric repeats on both sides of chromosome I differed totally from each other. An approximately 1.7 kb region adjacent to the left arm sequence shown in Fig. 6A consisted of an extended array of repeated sequence elements denoted as 5′-N_{3-9} CGACAATC/GAAA. Sixteen copies of the element are present in the 5′ part of the sequence in Fig. 6A. These sequence elements seemed to be specific for the left arm of chromosome I because Southern hybridization experiments with this sequence as a probe did not show any positive signals for other parts of chromosome I or for the other fifteen chromosomes (data not shown). On the other hand, the sequence flanking the right arm telomere possessed an unexpected feature: an approximately 1.0 kb sequence immediately next to the telomeric repeats has a unique structure including a 3′ polyA stretch (Fig. 6B). Similar elements are also detectable in other regions of chromosome I; Southern blot analysis of *Hin*dIII fragments of chromosome I DNA probed with the 800 bp *Sph*I-*Hin*dIII fragment of H1.3 revealed six positive bands ranging from 900 bp to 10.0 kb in size; a 1.3 kb band corresponded to the right arm element (Fig. 8A). All of these elements except for the most

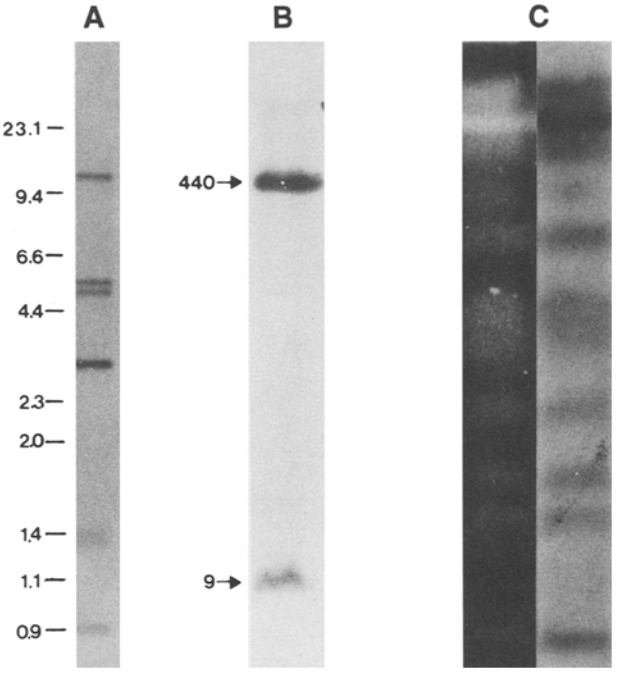

Fig. 8A-C Distribution of the poly(A) type retroposon on the *C.vulgaris* chromosomal DNAs. Chromosome I DNA was digested with *Hin*dIII **A** or *Not*I **B**, separated by gel electrophoresis, and probed with digoxygenin-labeled retroposon DNA (a 800 bp *Sph*I-*Hin*dIII fragment of H1.3). **C** Chromosomal DNAs of *C. vulgaris* were separated by CHEF gel electrophoresis and analyzed by Southern blot hybridization with the same probe as used in **A** and **B**. The CHEF gel electrophoresis pattern was the same as lane S in Fig. 4. Sizes are shown in kb

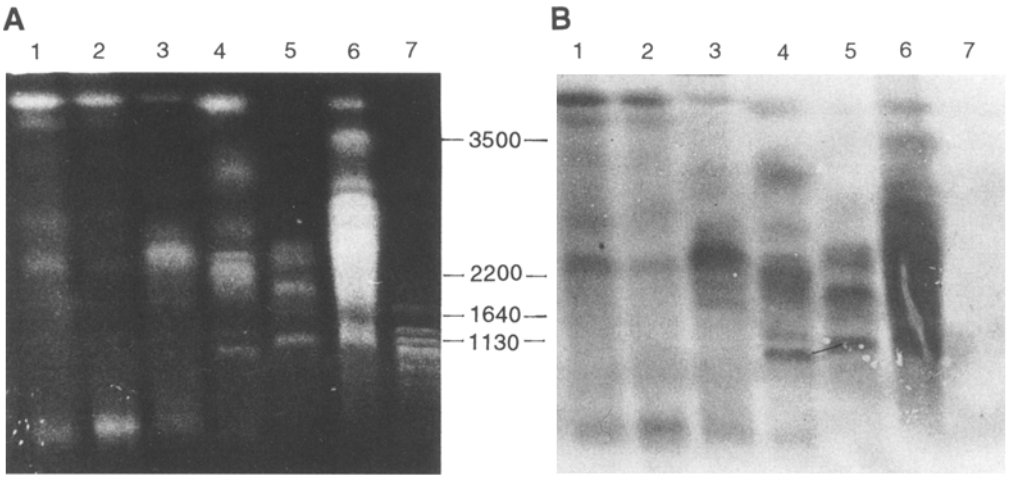

Fig. 9A,B Conservation of the telomeric sequence in *Chlorella* strains. Chromosomal DNAs of 6 strains of *Chlorella* species were separated by CHEF gel electrophoresis and analyzed by Southern blot hybridization with labeled pCHt-2 as a probe. Lanes 1 to 7 contain chromosomal DNAs of *C. vulgaris* C-135, C-150, and 211–11b, *Chlorella* sp. NC64A, *C. saccharophila* C-211, *C. vulgaris* C-169, and *Saccharomyces cerevisiae* (size markers), respectively. Sizes are indicated in kb

terminal one on the 9 kb *Not*I fragment were clustered within the 440 kb *Not*I fragment (Fig. 8B). Furthermore, all fifteen other chromosomes of *C. vulgaris* separated by CHEF gel electrophoresis exhibited strong hybridization with the probe for this element (Fig. 8C), indicating that homologous sequences with some degree of heterogeneity are dispersed in the *C. vulgaris* genome.

High conservation of the telomeric sequence in *Chlorella* strains

Previously we demonstrated that karyotypes of *Chlorella* strains differ considerably from strain to strain, even if they belong to the same species, and that there are two major chromosomal length polymorphism (CLP) groups (Higashiyama and Yamada 1991). In Fig. 9, chromosomal DNAs separated by CHEF gel electrophoresis of six *Chlorella* strains were hybridized with a probe of the telomeric repeat. All chromosomal DNA species of each strain exhibited strong hybridization signals under high stringency. This result indicated that the telomeric repeat sequence of *C. vulgaris* C-169, 5′-TTTAGGG, is highly conserved among *Chlorella* strains. Under the same hybridization conditions, the *Chlamydomonas reinhardtii* telomeric sequence 5′-TTTTAGGG showed different and much weaker hybridization patterns (data not shown).

Discussion

In this report, we characterized the terminal regions of chromosome I of the unicellular green alga *Chlorella vulgaris* C-169. We first expected that the *Chlorella* telomeric sequence would be identical or similar to that of another green alga, *Chlamydomonas reinhardtii* (Petracek et al. 1990) and used a synthetic 24mer of the *Chlamydomonas* sequence as a hybridization probe. The *Chlorella* sequence identified was, however, 5′-TTTAGGG running 5′ to 3′ towards the chromosomal terminus, which is exactly the same as the telomeric repeat of higher plants (Richards and Ausubel 1988; Ganal et al. 1991; Broun et al. 1992). The telomeric sequence is repeated regularly about 70 times (comprising approximately 500 bp) in tandem array at both termini of chromosome I (980 kb) of *C. vulgaris* C-169. This high degree of conservation of the telomeric sequence implies that the features of the telomerase which synthesizes telomeric DNA (Zakian 1989; Blackburn 1991, 1992) and of the telomere-associated proteins (Zakian 1989; Lustig et al. 1990; Gilson et al. 1993) might also be conserved between higher plants and *Chlorella*. Therefore, chromosome I of *C.vulgaris*, which is small enough to handle with ease (Higashiyama and Yamada 1991), serves as an excellent experimental subject for molecular analyses.

The microheterogeneity seen in the telomeric repeats among individual clones of the same terminal fragments of *C. vulgaris* chromosome I could be due to erorrs made during replication by telomerase. The error patterns show a bias, changing the sequence 5′-TTTAGGG to predominantly 5′-TTTTTAGGG, 5′-TTTTAGGG, or 5′-TTAGGG (Figs. 3 and 5C). By analogy to the *Tetrahymena* enzyme (Blackburn 1991, 1992), if the primer sequence of *Chlorella* telomerase is 5′-AAACCCUAAA, then slippage might occur in pairing the 3′ end of DNA (-TTT-3′) and the 3′ end of telomere RNA (-AAA-3′). Sometimes short streches of other sequences were observed within the telomeric repeats, for example 5′-ATATTC or 5′-TTACCTCATCCCCTATA in a specific clone (Fig. 6). The origin of these sequences is totally unknown. Since the lengths of the telomeric repeats on both arms of chromosome I were almost the same, some mechanisms should operate to control the telomerase reaction (Lustig et al. 1990). It is not known whether the length of telomeric repeats changes with changing physiological conditions, such as those that affect growth rate (McEachern and Hicks 1993). No difference was detected between cells grown photosynthetically in the light and cultures grown heterotrophically in the dark (unpublished data).

The telomeric repeats on the right arm are flanked by a sequence element with a poly A 3′ stretch (15 As) (Fig. 6B). An approximately 1.0 kb sequence, upstream from the poly A tail, is repeated several times on chromosome I; although their exact positions were not determined, these repeats are clustered on the 440 kb *Not*I fragment (Fig. 8B). Southern blot analysis showed that this kind of element is dispersed throughout all the other fifteen chromosomes of *C. vulgaris* C-169. From the characteristic structure with a long stretch of As, we consider these elements to be retroposons, like LINEs or SINEs (Weiner et al. 1986; Berg and Howe 1989; Boeke and Corces 1989). Homology searches, however, could not find within their nucleotide sequences any obvious motifs, of tRNAs characteristic or 7S RNA which are transcribed by RNA polymerase III (Geiduschek 1988) and are characteristic of SINEs (Deininger 1989). Detailed molecular characterization of these elements and determination of the possible contribution of these elements to dynamic rearrangements of *C. vulgaris* chromosomes will be interesting and, in addition to the further molecular characterization of chromosome I, is one of our future research goals.

Acknowledgments We thank Tokichi Miyakawa, Eiko Tsuchiya, and Dai Hirata for helpful discussions. This work was supported in part by a grant from Technical Research Institute, Chugoku Electric Power Co., Inc.

References

Aimi T, Yamada T, Murooka Y (1993) A group-I self-splicing intron in the nuclear small subunit rRNA-encoding gene of the green alga, *Chlorella ellipsoidea* C-87. Gene 134:65–71

Berg DE, Howe MW (1989) Mobile DNA, American Society for Microbiology, Washington DC

Blackburn EH (1991) Structure and function of telomeres. Nature 350:569–573

Blackburn EH (1992) Telomerases. Annu Rev Biochem 61:13–129

Boeke JD, Corces VG (1989) Transcription and reverse transcription of retrotransposons. Annu Rev Microbiol 43:403–434

Broun P, Ganal MW, Tanksley SD (1992) Telomeric arrays display high levels of heritable polymorphism among closely related plant varieties. Proc Natl Acad Sci USA 89:1354–1357

Chan CSM, Tye B-K (1980) Autonomously replicating sequences in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 77:6329–6333

Clarke L, Carbon J (1980) Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature 305:504–509

Deininger PL (1989) SINEs: short interspersed DNA elements in higher eucaryotes. In: Berg DE, Howe MM (ed) Mobile DNA. American Society for Microbiology, Washington DC, pp 619–636

Fott B, Nováková M (1969) A monograph of the genus *Chlorella*. The freshwater species. In: Fott B (ed) Studies in phycology. Verlag Akademische Wissenschaft, Prague, pp 10–74

Ganal MW, Lapitan NLV, Tanksley SD (1991) Macrostructure of tomato telomeres. Plant Cell 3:87–94

Geiduschek EP (1988) Transcription by RNA polymerase III. Annu Rev Biochem 57:873–914

Gilson E, Laroche T, Gasser SM (1993) Telomers and the functional architecture of the nucleus. Trends Cell Biol 3:128–134

Govindjee, Braun BZ (1974) Light absorption, emission and photosynthesis. In: Stewart WDP (ed) Algal physiology and biochemistry. University of California Press, Berkeley – Los Angeles, pp 346–390

Higashiyama T, Yamada T (1991) Electrophoretic karyotyping and chromosomal gene mapping of *Chlorella*. Nucleic Acids Res 19:6191–6195

Lustig AJ, Kurtz S, Shore D (1990) Involvement of the silencer and UAS binding protein RAP1 in regulation of tlomere length. Science 250:549–553

Maluszynska J, Heslop-Harrison JS (1991) Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. Plant J 1:159–166

McEachern MJ, Hicks JB (1993) Unusually large telomeric repeats in the yeast *Candida albicans*. Mol Cell Biol 13:551–560

Murray AW, Szostak JW (1983) Construction of artificial chromosomes in yeast. Nature 305:189–193

Newlon C (1988) Yeast chromosome replication and segregation. Microbiol Rev 52:568–601

Petracek ME, Lefebvre PA, Silflow CD, Berman J (1990) *Chlamydomonas* telomere sequences are A+T-rich but contain three consecutive G:C base pairs. Proc Natl Acad Sci USA 87:8222–8226

Richards EJ, Ausubel FM (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. Cell 53:127–136

Richards EJ, Goodman HM, Ausubel FM (1991) The centromere region of *Arabidopsis thaliana* chromosome I contains telomere-similar sequences. Nucleic Acids Res 19:3351–3357

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual (2nd edn). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Szostak JW, Blackburn EH (1982) Cloning yeast telomeres on linear plasmid vectors. Cell 29:245–255

Weiner AM, Dinninger PL, Efstratiadis A (1986) Nonviral retro transposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu Rev Biochem 55:631–661

Yamada T (1982) Isolation and characterization of chloroplast DNA from *Chlorella ellipsoidea*. Plant Physiol 70:92–96

Yamada T (1993) Cyanobacteria and algae. In: Murooka Y, Imanaka T(ed) Recombinant microbes for industrial and agricultural applications. Marcel Dekker, New York pp 701–712

Yamada T, Maki S, Higashiyama T (1993) Nucleotide sequence of a *Chlorella vulgaris* α-tubulin gene. Plant Physiol 103:1467–1467

Zakian VA (1989) Structure and function of telomeres. Annu Rev Genet 23:579–604