MOLECULAR AND CELLULAR BIOLOGY, June 1991, p. 3326–3330
0270-7306/91/063326-05$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 11, No. 6

Telomeric Location of *Giardia* rDNA Genes

RODNEY D. ADAM,[1]* THEODORE E. NASH,[2] AND THOMAS E. WELLEMS[2]

Department of Internal Medicine, Section of Infectious Disease and Department of Microbiology/Immunology, University of Arizona, Tucson, Arizona 85724,[1] and Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892[2]

Received 31 December 1990/Accepted 17 March 1991

Giardia lamblia **telomeres have been isolated from a library enriched for repaired chromosome ends by (i) screening with a *Plasmodium falciparum* telomere and (ii) differential hybridization with *Bal* 31-digested and total *G. lamblia* DNA. Analysis of three clones isolated by this strategy has identified multiple tandem repeats of the 5-mer TAGGG. An oligonucleotide containing these repeats recognizes *Bal* 31-sensitive bands in Southern hybridizations and detects all *G. lamblia* chromosomes in pulsed-field gel electrophoresis separations. An abrupt transition from the *G. lamblia* rDNA sequence to telomeric repeats has been found in all three clones. In two of the clones the transition occurs at the same site, near the beginning of the large subunit rDNA sequence. In the third clone the transition occurs at a site in the intergenic spacer sequence between the rDNA genes. Hybridization of an rDNA probe to a pulsed-field separation of *G. lamblia* chromosomes indicates that rDNA genes are present on several chromosomes but vary in location from isolate to isolate. These results suggest that rRNA genes are clustered at telomeric locations in *G. lamblia* and that these clusters are mobile.**

Giardia lamblia is a flagellated protozoan parasite that commonly causes diarrheal disease throughout the world. Infection occurs following ingestion of cysts which excyst in the small intestine to form trophozoites that are responsible for the symptoms of the illness. Cysts contain four nuclei, while individual trophozoites contain two morphologically identical nuclei that replicate at approximately the same time (28). Both nuclei in the trophozoite appear to be transcriptionally active, and both contain copies of rDNA genes (11). In a previous study of the *G. lamblia* genome, we showed that trophozoites contain multiple copies of closely related chromosomes arranged in five sets (2). In certain *Giardia* clones and subclones, distinct minor bands were evident that stained less intensely with ethidium bromide than the five major bands but always demonstrated homology with a major band. Each minor band was found to represent a chromosomal DNA molecule at a stoichiometric level of one copy per nucleus, while each major band contained several chromosomal DNA molecules (2). Densitometric measurements were used to estimate that each trophozoite contained approximately 30 to 50 chromosomal DNA molecules and 8×10^7 to 12×10^7 bp of DNA. In comparison, C_0t analysis has yielded genomic complexity values of 3.0×10^7 to 3.2×10^7 bp (16) and 8×10^7 bp (5). More detailed physical mapping of the chromosomes would be useful in further addressing the genome size and complexity as well as the ploidy of the organism.

To further address the chromosome organization of *G. lamblia*, we have performed experiments to identify telomeric sequences of the parasite. Telomeres contain repetitive DNA sequences located at the chromosome ends of every eukaryote and allow chromosome replication without loss of essential coding material. The repetitive sequences are similar in a wide variety of organisms and consist of short repeats rich in guanosine. In some organisms there is variability in the exact sequence from repeat to repeat (3). Knowledge of the human telomeric sequence has been used to provide evidence that all or most vertebrate telomeres contain the same TTAGGG repeating sequence (15). In contrast, the diversity of telomeric sequences among the protozoans is much greater, perhaps reflecting the great genetic diversity of the protozoans. *G. lamblia* has been proposed as one of the most primitive of all eukaryotic organisms (25). Previous work indicated the lack of cross-hybridization between the trypanosome telomeric sequence (TTAGGG) and *G. lamblia* DNA (4). In this work, we have used the observation that the telomeric sequence from *Plasmodium falciparum* (TTT/CAGGG) cross-hybridizes to *G. lamblia* DNA at low stringency to clone telomeric DNA from *G. lamblia*. Analysis of three clones containing *G. lamblia* telomeric repeat sequences shows an abrupt transition from rDNA sequence to the telomeric repeat, TAGGG.

MATERIALS AND METHODS

Construction of telomere-enriched library. A library of *Giardia* DNA enriched for telomeric sequences was constructed (Fig. 1). DNA from the WB *Giardia* isolate was digested with *Pst*I and size selected for fragments greater than 1.3 kb. Vector DNA was prepared by digestion of pUC13 DNA with *Pst*I and *Sma*I. Ligation was performed with 0.1 U of T4 DNA ligase overnight at 4°C to optimize for overlapping-end (*Pst*I-*Pst*I) ligation. The unligated ends of *Giardia* DNA were then treated with Klenow fragment of DNA polymerase I with all four nucleotides. Ligation was performed under conditions designed to optimize blunt-end ligation of the *Giardia* DNA to the *Sma*I end of the pUC 13 vector (Fig. 1).

***Bal* 31 digestion.** *Giardia* DNA from the ISR isolate was digested with fast *Bal* 31 (International Biotechnologies, Inc., New Haven, Conn.). Aliquots removed at specified time intervals were treated by phenol-chloroform extraction and ethanol precipitation. Recovered DNA was digested with *Hin*dIII or *Bam*HI, separated by agarose electrophoresis, and blotted to Nytran. When used for secondary screening of the library, DNA was digested for 30 min with *Bal* 31, phenol-chloroform extracted, ethanol precipitated, and re-

* Corresponding author.

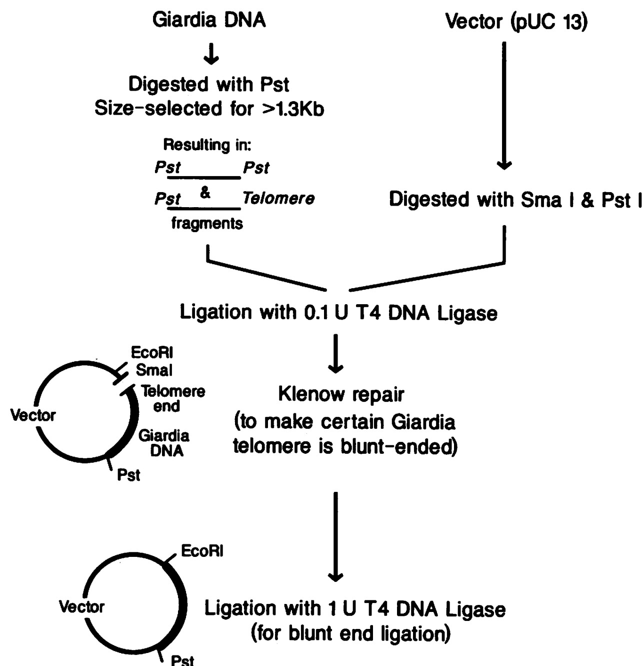

FIG. 1. Strategy of construction of a library enriched for telomeric sequences.

suspended in TE for radiolabeling by random hexamer priming.

Identification of recombinant clones containing telomeric sequences. Competent DH5α *Escherichia coli* cells (Life Technologies, Inc., Grand Island, N.Y.) were transformed by using the above-described ligation product. The library was screened with the purified insert from a *P. falciparum* telomeric clone, pC4.H11 (23). Hybridization was performed at 50°C with 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), and washing was performed at 45°C in 1× SSC. Replica blots of positive colonies were probed with (i) total *Giardia* DNA and (ii) total *Giardia* DNA which had been digested for 30 min with *Bal* 31. DNA was prepared from the colonies that hybridized with total *Giardia* DNA but not with *Bal* 31-digested DNA. Three clones contained inserts released by *Pst*I and *Eco*RI (on the opposite side of the *Sma*I site from the *Pst*I site) but not by *Pst*I or *Eco*RI alone.

Sequencing. DNA sequencing was performed by standard methods with modified T7 polymerase (Sequenase; United States Biochemical Corp., Cleveland, Ohio) or Klenow fragment of DNA polymerase I with plasmid preparations and M13 subclones. An acrylamide gel containing 50% formamide was used to resolve compression artifacts in GC-rich regions.

Restriction digests and DNA hybridization. Restriction digests were performed under conditions recommended by the suppliers. After agarose gel electrophoresis, DNA was transferred to Nytran (Schleicher & Schuell, Inc., Keene, N.H.) in 10× SSC or by alkaline transfer (0.4 M NaOH) (21). Oligonucleotides were synthesized by Operon Technologies, Inc. (San Pablo, Calif.) ($TAGGG_5$), or by the University of Arizona Biotechnology Resource Facility (GGTCGCTGGG ATTCGTCTTCTTCA). The oligonucleotides were end labeled with [γ-^{32}P]ATP and T4 kinase, and unincorporated nucleotides were separated by Sephadex G25 chromatography. Hybridization and final washing were performed at 50°C in 6× SSC. Double-stranded DNA was labeled with [^{32}P]dCTP by random priming (9), and unincorporated nucleotides were removed by Sephadex G50 chromatography. Hybridization was performed at 50°C in 4× SSC, and final washing was performed at 50°C in 0.2× SSC.

PFGE. Pulsed-field gel electrophoresis (PFGE) was performed as described (2) on an OFAGE unit (7). Conditions used for separation of chromosomes were 80 V for 3 days at a switching interval of 11 min, 65 V for 2 days at 25 min, and 50 V for 2 days at 60 min.

RESULTS

Identification of telomeric sequence. A *Bal* 31-*Hin*dIII digest of *Giardia* DNA was probed with the pC4.HII *P. falciparum* telomere clone and washed at low stringency. Autoradiographic signals indicated the presence of a *Bal* 31-sensitive signal (data not shown). The pC4.H11 probe was therefore used to screen the telomere-enriched *Giardia* DNA library. Twenty clones identified by low-stringency hybridization with pC4.H11 were then lifted on duplicate filters and probed with two radiolabeled samples of *Giardia* DNA: (i) total DNA and (ii) DNA that had been treated for 30 min with *Bal* 31. Nineteen clones showed intense hybridization to total *Giardia* DNA and little or no hybridization to *Bal* 31-digested *Giardia* DNA. Plasmid preparations of these clones were digested with *Eco*RI plus *Pst*I or with *Pst*I alone. Three clones (pG1T, pG2T, and pG3T) contained inserts that were released by *Pst*I plus *Eco*RI but not by *Pst*I alone, as would be expected for an insert with the blunt end cloned into the *Sma*I site and the other into the *Pst*I site. In addition to the *Eco*RI-*Pst*I insert, clone pG3T also contained two additional inserts that were excised by *Pst*I alone; these inserts were not analyzed further. The other 16 clones contained inserts excised by *Pst*I alone and most likely resulted from vector DNA that had been incompletely digested by *Sma*I.

Sequence analysis of the pG1T clone revealed a 5-bp repeat (CCCTA), extending approximately 0.6 kb from the blunt-end *Sma* I site. Since telomeric sequences are customarily represented in the opposite direction (i.e., sequencing toward the chromosome end), the sequence is referred to as TAGGG [in comparison, the *Plasmodium* telomere consensus repeat is $T_2(T/C)AG_3$] (20). Clones pG2T and pG3T had 40 and 34 copies, respectively, of the telomeric repeat. The telomeric repeats were uniform in each of the clones except that the 11th repeat from the blunt end of both pG1T and pG3T was TAAG rather than TAGGG.

An oligonucleotide consisting of five telomeric repeats was used to probe *Bal* 31-*Bam*HI digests of *G. lamblia* DNA (Fig. 2). The autoradiographic signal revealed hybridization to 9.4- and 6.4-kb fragments and a broad band extending from approximately 2 to 4 kb that almost totally disappeared within 5 minutes of *Bal* 31 digestion, confirming the *Bal* 31 sensitivity of the repeat sequence, TAGGG. The broad bands recognized by the telomeric oligonucleotide are characteristic of telomeric restriction fragments and stand in contrast to the much fainter bands at 6.8, 5.8, and 3.7 kb (Fig. 2B). The latter bands most likely represent a few copies of the TAGGG repeat in a *Bam*HI fragment internal to the end of the chromosome and confirm that the DNA was not internally degraded by *Bal* 31. In addition, a control probe hybridized to distinct bands of each lane on the same blot

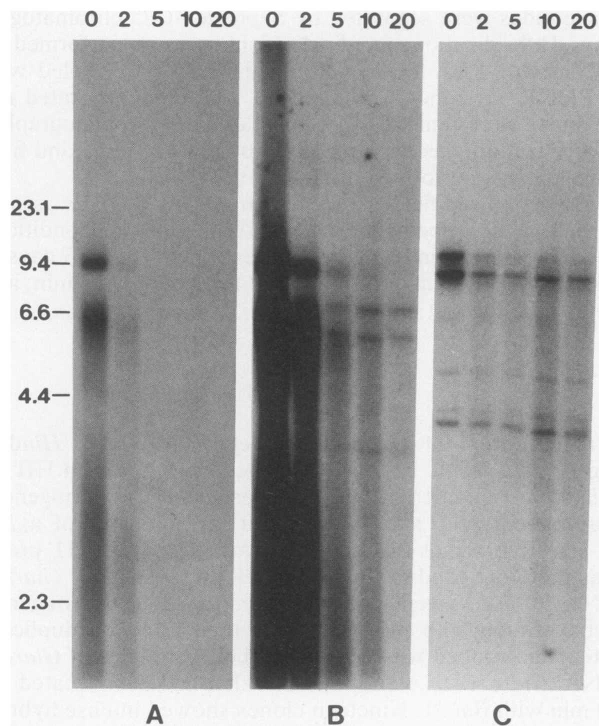

FIG. 2. A Southern blot of a *Bam*HI-Bal 31 digest of *G. lamblia* DNA was probed with the telomeric repeat oligonucleotide, and the film was exposed for 18 h (A) or 12 days (B). The blot was also hybridized with a control probe (an oligonucleotide based on the sequence of the 195-bp repeat of CRP170 [1]), confirming that the DNA was not nonspecifically degraded by *Bal* 31 (C). Sizes (in kilobases) are on the left. Durations of *Bal* 31 digestion (in minutes) are on the top.

(Fig. 2C), providing further evidence that the *Bal* 31 degraded only DNA at the ends of the chromosomes.

Telomeric location of rDNA genes. The sequence from the region of pG1T immediately adjacent to the telomeric repeats was compared with the GenBank and EMBL data banks by using the Pearson and Lipman FastA algorithm (18). Results of this computer search revealed a 126-bp region of 96% identity with the *G. lamblia* 5.8S rRNA sequence. The sequences of each of the telomere clones were aligned with the *Giardia* rDNA gene sequence published by Healey et al. (10), using the BestFit local homology algorithm of Smith and Waterman (24) (Fig. 3). Clones pG1T and pG3T contained inserts extending from the *Pst*I site in the region encoding the small subunit rRNA to base 1108, before an abrupt transition to the TAGGG repeat. The only difference between pG1T and pG3T was the number of telomeric repeats (120 versus 34). The presence of two additional inserts in the pG3T clone indicates that pG1T and pG3T were not derived from the same cloning event. However, the sequence similarity suggests the possibility that the pG1T and pG3T inserts could have derived from the same chromosome end, with stretches of telomeric repeats deleted during cloning in *E. coli*. In contrast, the pG2T clone extended from the *Pst*I site near the end of the coding region of the large subunit RNA to base 4102, where the telomeric repeat motif began. The sequences immediately prior to the transition, CCCCGGAGG for pG1T and pG3T and CCCC GGACGGC for pG2T, included a 7-base stretch of perfect identity (CCCCGGA) between the two.

Chromosome heterogeneity of rDNA gene location. A PFGE separation of the chromosomes of four *Giardia* isolates was probed with the telomeric oligonucleotide (Fig. 4A and B). Autoradiographic signals show that a telomeric repeat sequence is present on each of the chromosomes, including the more faintly staining minor bands. The intensity of hybridization is approximately proportional to the intensity of ethidium bromide staining. The blot was reprobed with the rDNA gene (kindly provided by T. Edlind) (8) (Fig. 4C). Results showed that rDNA genes occur on multiple *Giardia* chromosomes but that their location varies from isolate to isolate. For the E11 clone of the ISR isolate (Fig. 4C, lane 2), most of the hybridization was to minor band b, with less hybridization to other chromosomal bands. The hybridization patterns for each of the other isolates were

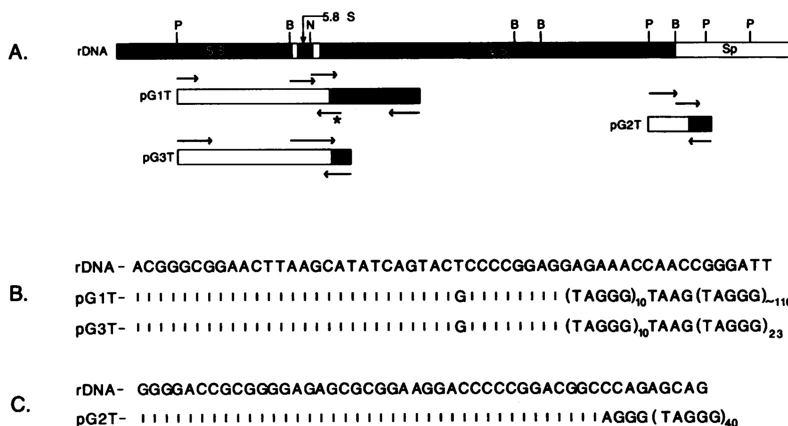

FIG. 3. (A) Restriction map of the *Giardia* rDNA 5,566-bp gene, determined from the sequence reported by Healey et al. (11) beginning with the coding region for the small subunit (SS) rRNA (base 4929), and demonstrating the 5.8S, large subunit (LS), and intergenic spacer (Sp) regions. The *Bam*HI (B) and *Pst*I (P) sites are shown, as well as the *Nar*I (N) site nearest the telomeric repeats of pG1T and pG3T. Telomeric repeats are crosshatched, and sequencing strategy is noted by arrows in the appropriate directions. *, M13 subclone which sustained a spontaneous deletion of most of the telomeric repeats. (B) Alignment of the breakpoint sequences from pG1T and pG3T with the rDNA gene. Note: other reported sequences (GenBank #M35013 and X05396/Y00331) contain a G rather than T where a T-to-G substitution is shown in the figure. (C) Alignment of the breakpoint sequence of pG2T with the rDNA gene.

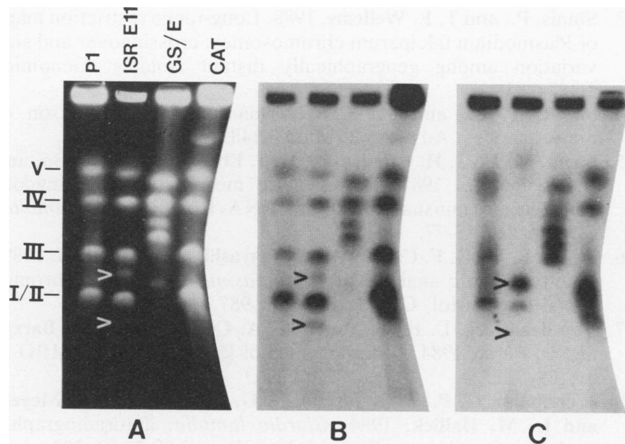

FIG. 4. Hybridization of *Giardia* telomeric sequence and rDNA to PFGE separations of *Giardia* DNA. (A) Ethidium bromide gel. (B) Autoradiograph of Southern blot of gel in panel A probed with the telomeric oligo repeat. (C) Autoradiograph of blot of panel A stripped and probed with the rDNA gene (provided by T. Edlind). Arrows identify minor band a (lower arrow) and minor band b (upper arrow).

unique, indicating variable numbers of copies of the rDNA gene on several chromosomes.

DISCUSSION

The *Giardia* telomeric repeat, TAGGG, preserves the guanosine-rich trait that is a characteristic of telomeric sequences (3). It is present on all *G. lamblia* chromosomal molecules separated by PFGE, consistent with our previous observations that these molecules are linear (2). Relative intensities of hybridization by a telomeric probe to each of the major and minor PFG bands are comparable to ethidium bromide staining intensities, consistent with the observation that these bands contain different numbers of chromosomal DNA molecules.

The location of copies of the rDNA gene on several chromosomes of the *Giardia* isolates (Fig. 4C) suggests that their telomeric position has given them increased mobility, consistent with a telomeric location where translocations and deletions may be common. In most organisms, the rDNA genes are present in multiple tandemly repeated copies (12). In mammals, the rDNA genes are usually present on clusters located on several chromosomes, while the rDNA genes of the lower eukaryotes are most commonly present on a single chromosome. For example, the rDNA genes of *Saccharomyces cerevisiae* (22), *Candida albicans* (13), *Neurospora crassa* (17), *Histoplasma capsulatum* (26), and the kinetoplastids (27) are most likely present on single chromosomes. The rDNA genes of *P. falciparum* are interspersed over several chromosomes (29). However, *G. lamblia* is the only organism so far for which a subtelomeric location for the rDNA genes has been found.

The subtelomeric regions of other organisms, including humans, *S. cerevisiae*, *Drosophila melanogaster*, and *P. falciparum*, consist of regions of repetitive DNA that can vary among different chromosomes (30). In humans, chromosomes can be classified into subsets by their subtelomeric regions (6). In *G. lamblia*, the abrupt transition from rDNA to the telomeric repeats suggests an analogy with the breaking and healing events that have been reported with gene deletions and chromosome size polymorphisms in *P. falciparum* (19). Analysis of the *P. falciparum* breakpoint sequences revealed the presence of the CA dinucleotide at each of three sequences examined (19). Breakpoint analysis of the sequences associated with nonhomologous recombination in mammals frequently reveals the presence of repeats of 7 or fewer bp (14). The CCCCGGA sequences we identified at the breakpoint of *G. lamblia* rDNA have not been found at breakpoints in other organisms. Further investigation will be required to determine the sequences associated with these and other breaking and healing events in *G. lamblia*.

ACKNOWLEDGMENTS

We thank Tom Edlind for providing a clone of the *Giardia* rDNA gene, Lynda Schurig for technical assistance, and Melanie Rucker for editorial assistance.

This work was supported in part by a Smith Kline & French Young Investigator Award awarded to Rodney D. Adam by the Infectious Diseases Society of America.

ADDENDUM IN PROOF

The location of *Giardia* rDNA genes on multiple chromosomes has also been reported by Jonckheere et al. (Mol. Biochem. Parasitol. **29**:23–28, 1990).

REFERENCES

1. **Adam, R. D., A. Aggarwal, A. A. Lal, V. F. De La Cruz, T. McCutchan, and T. E. Nash.** 1988. Antigenic variation of a cysteine-rich protein in *Giardia lamblia*. J. Exp. Med. **167**:109–118.
2. **Adam, R. D., T. E. Nash, and T. E. Wellems.** 1988. The *Giardia lamblia* trophozoite contains sets of closely related chromosomes. Nucleic Acids Res. **16**:4555–4567.
3. **Blackburn, E. H.** 1990. Telomeres: structure and synthesis. J. Biol. Chem. **265**:5919–5921.
4. **Blackburn, E. H., and P. B. Challoner.** 1984. Identification of a telomeric DNA sequence in *Trypanosoma brucei*. Cell **36**:447–457.
5. **Boothroyd, J. C., A. Wang, D. A. Campbell, and C. C. Wang.** 1987. An unusually compact ribosomal DNA repeat in the protozoan *Giardia lamblia*. Nucleic Acids Res. **15**:4065–4084.
6. **Brown, W. R. A., P. J. MacKinnon, A. Villasante, N. Spurr, V. J. Buckle, and M. J. Dobson.** 1990. Structure and polymorphism of human telomere-associated DNA. Cell **63**:119–132.
7. **Carle, G. F., and M. V. Olson.** 1984. Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. Nucleic Acids Res. **12**:5647–5664.
8. **Edlind, T. D., and P. R. Chakraborty.** 1987. Unusual ribosomal RNA of the intestinal parasite *Giardia lamblia*. Nucleic Acids Res. **15**:7889–7901.
9. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
10. **Healey, A., R. Mitchell, J. A. Upcroft, and P. F. Boreham.** 1990. Complete nucleotide sequence of the ribosomal RNA tandem repeat unit from Giardia intestinalis. Nucleic Acids Res. **18**: 4006.
11. **Kabnick, K. S., and D. A. Peattie.** 1990. In situ analyses reveal that the two nuclei of *Giardia lamblia* are equivalent. J. Cell Sci. **95**:353–360.
12. **Lewin, B.** 1980. Gene expression—eucaryotic genomes, p. 875–906. *In* B. Lewin (ed.), Genes III. John Wiley & Sons, Inc., New York.
13. **Magee, B. B., Y. Koltin, J. A. Gorman, and P. T. Magee.** 1988. Assignment of cloned genes to the seven electrophoretically separated *Candida albicans* chromosomes. Mol. Cell. Biol. **8**:4721–4726.
14. **Meuth, M.** 1989. Illegitimate recombination in mammalian cells,

p. 833–860. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.

15. **Meyne, J., R. L. Ratliff, and R. K. Moyzis.** 1988. Conservation of the human telomere sequence (TTAGGG)n among vertebrates. Proc. Natl. Acad. Sci. USA **86:**7049–7053.
16. **Nash, T. E., T. McCutchan, D. Keister, et al.** 1985. Restriction-endonuclease analysis of DNA from 15 Giardia isolates obtained from humans and animals. J. Infect. Dis. **152:**64–73.
17. **Orbach, M. J., D. Vollrath, R. W. Davis, and C. Yanofsky.** 1988. An electrophoretic karyotype of *Neurospora crassa*. Mol. Cell. Biol. **8:**1469–1473.
18. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence analysis. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
19. **Pologe, L. G., and J. V. Ravetch.** 1988. Large deletions result from breakage and healing of *P. falciparum* chromosomes. Cell **55:**869–874.
20. **Ponzi, M., T. Pace, E. Dore, et al.** 1985. Identification of a telomeric DNA sequence in *Plasmodium berghei*. EMBO J. **4:**2991–2995.
21. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
22. **Schwartz, D. C., and C. R. Cantor.** 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell **37:**67–75.
23. **Sinnis, P., and T. E. Wellems.** 1988. Long-range restriction maps of Plasmodium falciparum chromosomes; crossingover and size variation among geographically distant isolates. Genomics **3:**287–295.
24. **Smith, T. F., and M. S. Waterman.** 1981. Comparison of bio-sequences. Adv. Appl. Math. **2:**482–489.
25. **Sogin, M. L., J. H. Gunderson, H. J. Elwood, R. A. Alonso, and D. A. Peattie.** 1989. Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. Science **243:**75–77.
26. **Steele, P. E., G. F. Carle, G. S. Kobayashi, and G. Medoff.** 1989. Electrophoretic analysis of *Histoplasma capsulatum* chromosomal DNA. Mol. Cell. Biol. **9:**983–987.
27. **Van der Ploeg, L. H. T., A. W. C. A. Cornelissen, J. D. Barry, and P. Borst.** 1984. Chromosomes of kinetoplastida. EMBO J. **3:**3109–3115.
28. **Weisehahn, G. P., E. J. Jarroll, D. G. Lindmark, E. A. Meyer, and L. M. Hallick.** 1984. *Giardia lamblia*: autoradiographic analysis of nuclear replication. Exp. Parasitol. **58:**94–100.
29. **Wellems, T. E., D. Walliker, C. L. Smith, V. E. do Rosario, W. L. Maloy, R. J. Howard, R. Carter, and T. F. McCutchan.** 1987. A histidine-rich protein gene marks a linkage group favored strongly in a genetic cross of Plasmodium falciparum. Cell **49:**633–642.
30. **Zakian, V. A.** 1989. Structure and function of telomeres. Annu. Rev. Genet. **23:**579–604.