Telometric: A Tool Providing Simplified, Reproducible Measurements of Telomeric DNA from Constant Field Agarose Gels

BioTechniques 31:1314-1318 (December 2001)

J.D. Grant, D. Broccoli, M. Muquit, F.J. Manion, J. Tisdall, and M.F. Ochs
Fox Chase Cancer Center, Philadelphia, PA, USA

Cellular proliferation must be closely regulated by an organism to maintain health because cancers arise primarily through the uncontrolled growth of specific cells. Proliferation is controlled at least in part by telomere integrity. Telomeres, structural elements that confer stability at the end of chromosomes, contain 7–10 kb tandem repeats of 5′-TTAGGG-3′ in humans (6) and associated sequence-specific DNA binding proteins. DNA replication contributes to loss of telomeric DNA because typical DNA polymerases are unidirectional and require primers (8,10). Variation in telomeric length exists between individuals, between chromosomes, and between cells within an individual. This heterogeneity leads to difficulties in making a reproducible measurement of size distribution that is representative of the population. Such measurements help to assess the effect of tumorigenesis on telomeres, allowing the determination of the effect of drugs on tumors and the effect of potential chemopreventive agents.

To simplify estimates of the size of telomeric repeats, we needed to improve the speed and accuracy of measurements of the distributions of DNA molecular weights from agarose gels. Although other methods have been used to measure telomere length in acrylamide and pulsed-field gels (1,3) and directly in cells using fluorescent tags (9), we required an easy-to-use system for performing accurate, reproducible measurements on inexpensive constant field agarose gels. Such a system is not available, which has led many researchers to make "eyeball" measurements of the location of peak intensity. Accurate measurements are problematic because of the nonlinear migration of DNA on gels and because telomeric repeat regions hybridize multiple times to the probe (in this case, an end-labeled 24-mer), so that intensity distributions do not reflect copy number distributions. However, the correct determination of the size and distribution of the telomeric repeats requires measures on the distribution of copies, not intensities. We have developed easy-to-use, freely available software to determine the distributions by copy number and calculate statistics on them. This software, *Telometric*, is useful for studies of telomeric regions and for measuring DNA distributions in other contexts, such as cDNA library construction. It can be freely downloaded from the *BioTechniques* Software Library (www.BioTechniques.com).

The software is an NIH Image macro that analyzes an autoradiograph of an agarose gel hybridized with telomeric probe. The chromosomal DNA is prepared in standard fashion. Briefly, restriction digestion of phenol/chloroform-extracted genomic DNA is carried out with *Hin*fI and *Rsa*I. Following digestion, DNA concentrations are determined using the fluorescent DNA quantitation kit and VersaFluor™ fluorometer (Bio-Rad Laboratories, Hercules, CA, USA) following the manufacturer's recommendations, and equivalent amounts of DNA are loaded per lane, ranging from a minimum of 1.0 µg/lane to a maximum of 5.0 µg/lane, on 0.7% agarose gels. Following electrophoresis, the DNA is transferred to Hybond® N membranes (5). The membrane is hybridized with end-labeled oligonucleotides complementary to the telomeric repeats, 5′-TTAGGG-3′ and 5′-CCCTAA-3′, as described previously (2). The autoradiograph is scanned and imported into NIH Image (Figure 1). It is assumed that the grayscale intensity measured at each point in an image is directly proportional to the concentration of labeled DNA, essentially ignoring non-telomeric DNA attached to telomeres, so that such regions need to be either relatively small compared to the telomeric repeat region or relatively uncommon. Obviously for very small telomeric regions (<1 kb), this assumption may break down; however, the assumption seems justified above that size as shown by the distribution curves in Figure 2 that show no tail at short telomere lengths, which non-telomeric DNA would cause. Second, it is assumed that the image is unsaturated, which is easily checked by verifying that no intensity

Table 1. Useful Statistical Measurements Produced by *Telometric* for the Data in Figure 1

Lane	Median (kb)	Mode (kb)	SIR (kb)
1	11.8	7.7	5.1
2	8.2	5.5	2.8
3	14.2	10.0	4.8
4	12.8	9.7	3.8

For this gel with its large asymmetric distribution, the significant values for determining the average telomeric length and heterogeneity are the median and the semi-interquartile range (SIR). The SIR gives the spread in kilobases (divided by 2) spanning the 25th to 75th percentiles (e.g., for lane 1, 10.2 kb covers the middle 50% of the measurements). The low SIR for lane 2 shows how the eye can be fooled by the nonlinearity of an agarose gel. Although lane 2 in Figure 3 seems to have a broader distribution of DNA size, it is a result of the fact that smaller DNA fragments run faster and are spread out significantly more than larger size fragments. Figure 2 corrects for this and therefore shows the tighter distribution for lane 2.

curves show flat peaks (the program displays these curves).

The macro generates a profile of the grayscale intensity along a lane using standard NIH Image routines. The lanes are defined by the user and outlined by the software as shown in Figure 1. This profile is converted to a profile linear in molecular weight by a log linear interpolation of molecular weight to distance traveled (4), which assumes that the DNA is in the Ogston-sieving domain (7). This is adequate for most measurements of telomeric length, as the gels are prepared to keep most DNA in this domain. The background is removed by allowing the user to specify a background region from outside the lanes containing DNA (Figure 1), and the average pixel intensity in that region is subtracted from all pixels in the image. Next, the intensity (y-axis) is converted to relative copy number by adjusting for the number of probes bound to the DNA. For each lane of the gel, the relative copy number (C) is determined by

$$C_i = \frac{I_i - B_i}{L_i} \qquad \text{[Eq. 1]}$$

where C_i is the relative copy number associated with line i across the gel lane, I_i is the integrated grayscale intensity within line i, B_i is the background measurement for line i, and L_i is the DNA molecular weight at line i. Equation 1 provides a correction for the background intensity (B) and for the binding of multiple copies of the telomeric probe, which scales with the size of the DNA (L).

A plot of relative telomere copy number versus molecular weight is created, which gives the user a realistic picture of the actual distribution of telomeric lengths. In addition, the distribution permits statistics to be generated, including the mean, median, and

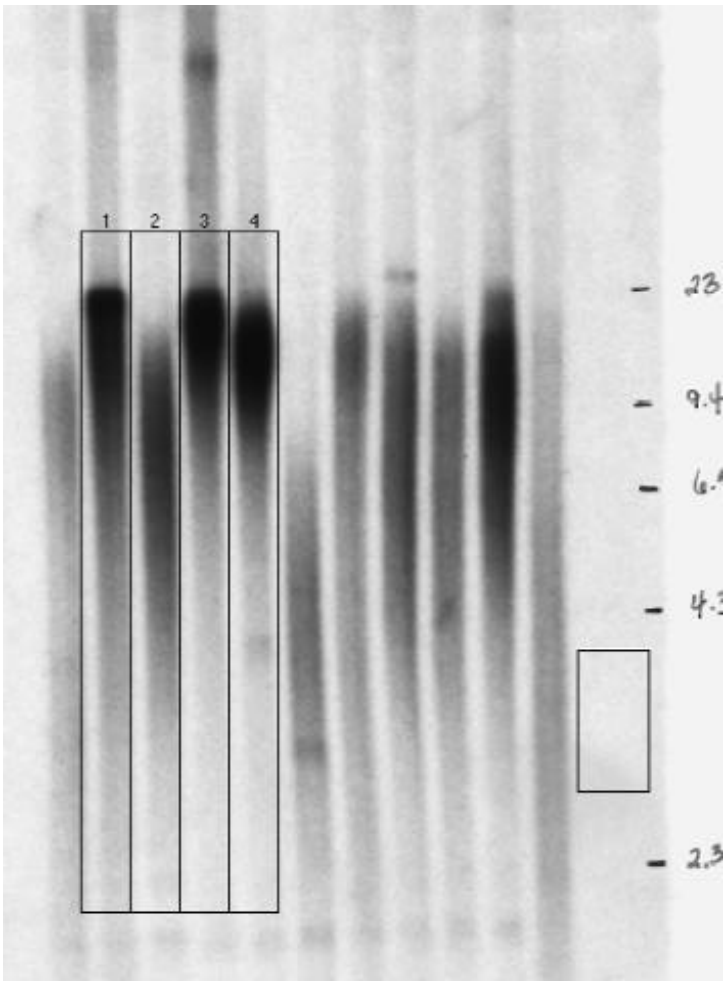

Figure 1. Image of a typical gel used to measure telomere length. This gel has been scanned and imported as a TIFF file into NIH Image. The lines on the right show the positions of λ*Hin*dIII DNA size markers that are used to calibrate the molecular weights in kilobases of DNA. The labeled boxes were added by the authors using the telometric software, and those four lanes were analyzed for this paper. The box on the right is the location of an area suitable for background estimation,as it is outside the lanes containing DNA. Note that, for the constant field agarose gels analyzed, we have not seen a "smile" effect and so typically run only a single marker lane.

mode of the molecular weight as well as the variance and semi-interquartile range (a useful measure of heterogeneity for asymmetric distributions). Users may use these statistics in various ways depending on the goal of their experiment. For bell-shaped distributions, the mean and variance provide excellent measures of typical length and heterogeneity of telomere length. For skewed distributions, the median and semi-interquartile range provide these measures, and we have found these values to be the most reliable in our own work. Table 1 contains the *Telometric* results for Figure 1.

We have validated the software in a number of ways. First, we ran λ*Hin*dIII markers side by side with a 1-kb ladder on 0.7% agarose gels and verified that we recovered the correct sizes for the ladder (R^2 = 0.99 for 14 measurements). Second, we tested the conversion of the intensity-distance distribution to the copy number-molecular weight distribution by measuring the relative copy number of the λ*Hin*dIII markers stained with ethidium bromide. Since ethidium bromide intercalates into the DNA, the intensity scales with the DNA length, just as with telomeric probes. We recovered the correct uniform distribution in relative copy number (standard deviation between markers of 3.5% in relative copy number). However, incorrect specification of background or poor gel quality can degrade this, as we have seen standard deviations as high as 13% when background correction is misapplied. This measurement also demonstrates the usefulness of the technique for analyzing DNA preparations for cDNA library construction. Finally, we checked gel-to-gel reproducibility by analyzing the size distributions for four identical telomere preparations run on five separate gels (one shown in Figure 1) on different days, giving three measurements each. The median values were 8.20 ± 0.17 kb, 9.77 ± 0.49 kb, 9.23 ± 0.49 kb, and 7.77 ± 0.32 kb, while the semi-interquartile ranges were 3.47 ± 0.12 kb, 4.37 ± 0.21 kb, 3.73 ± 0.38 kb, and 3.17 ± 0.76 kb.

Telometric is a new algorithm for measuring telomeric length from ubiquitously digested chromosomal DNA electrophoresed on constant field agarose gels. The method improves the knowledge of the distribution of the telomeric repeats by providing useful, consistent statistical measures. The software should also be able to make the same measurements on DNA prepared for cDNA library creation and stained with ethidium bromide. We soon hope to release a Java2 version of the software that will allow the code to run on all common computer platforms

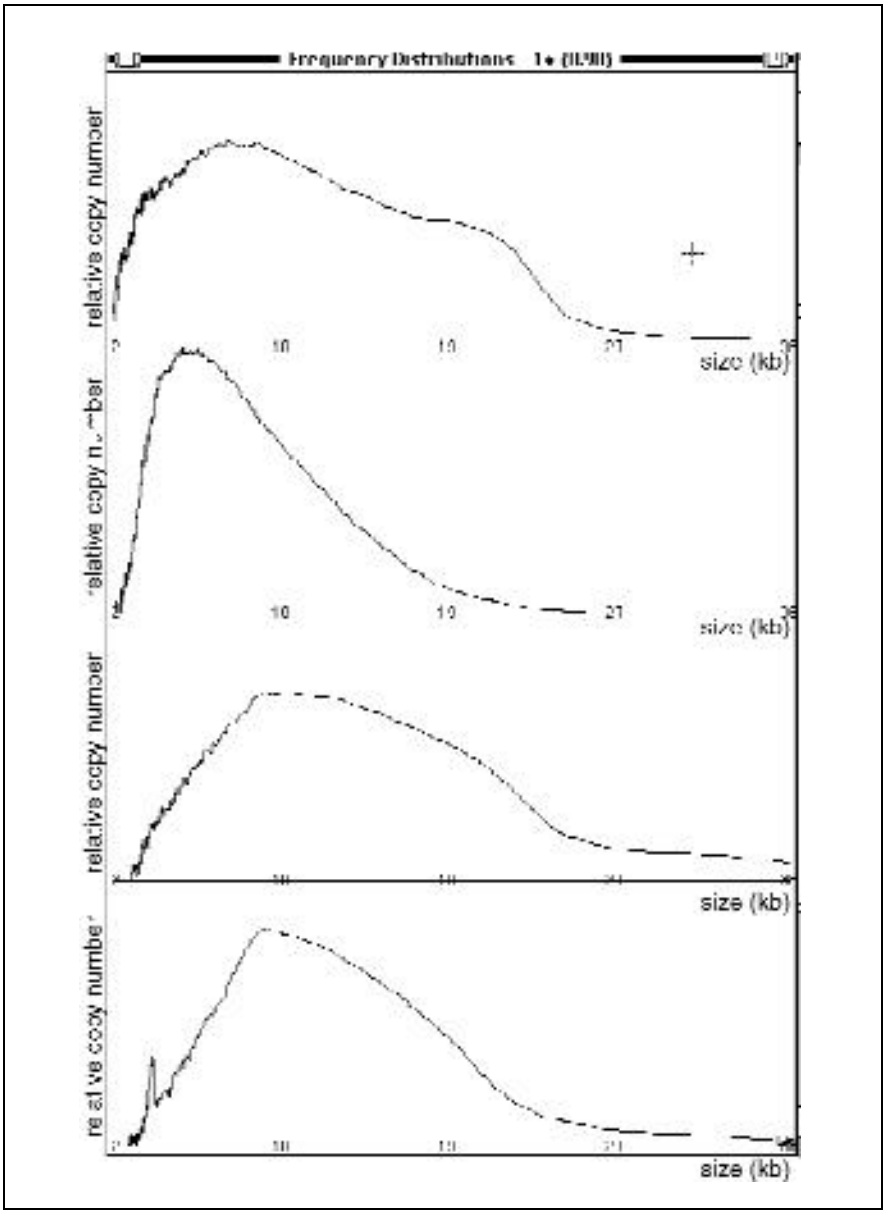

Figure 2. Plots of the relative copy number versus telomere length for the four lanes outlined in Figure 1 (lane 1 at the top to lane 4 at the bottom). This was produced by the application of Equation 1 to the raw data gathered from the TIFF image of Figure 1. Since the distributions are linear in telomere length and relative copy number, the statistics generated are useful for determining telomeric length distribution.

(Java version now available). All software and source code is freely available (http://bioinformatics.fccc.edu/software) and may be modified by users under the GNU public license (http://www.gnu.org/copyleft/gpl.html)

ACKNOWLEDGMENTS

We thank the National Institutes of Health, National Cancer Institute (grant no. CA06927 to R. Young) for support. We would also like to thank our reviewers for many helpful comments that improved this manuscript.

REFERENCES

1. **Allsopp, R.C., H. Vaziri, C. Patterson, S. Goldstein, E.V. Younglai, A.B. Futcher, C.W. Greider, and C.B. Harley.** 1992. Telomere length predicts replicative capacity of human fibroblasts. Proc. Natl. Acad. Sci. USA *89*:10114-10118.
2. **Broccoli, D., L.A. Godley, L.A. Donehower, H.E. Varmus, and T. de Lange.** 1996. Telomerase activation in mouse mammary tumors: lack of detectable telomere shortening and evidence for regulation of telomerase RNA with cell proliferation. Mol. Cell. Biol. *16*:3765-3772.
3. **Feng, Y.R., D. Norwood, R. Shibata, D. Gee, X. Xiao, M. Martin, S.L. Zeichner, and D.S. Dimitrov.** 1998. Telomere dynamics in HIV-1 infected and uninfected chimpanzees measured by an improved method based on high-resolution two-dimensional calibration of DNA sizes. J. Med. Primatol. *27*:258-265.
4. **Freifelder, D.M**. 1976. Physical Biochemistry: Applications to Biochemistry and Molecular Biology. W.H. Freeman, San Francisco.
5. **Luderus, M.E., B. van Steensel, L. Chong, O.C. Sibon, F.F. Cremers, and T. de Lange.** 1996. Structure, subnuclear distribution, and nuclear matrix association of the mammalian telomeric complex. J. Cell. Biol. *135*:867-881.
6. **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.
7. **Noolandi, J.** 1992. Theory of DNA gel electrophoresis, p. 2-57. *In* A. Chrambach, M.J. Dunn, and B.J. Radola (Eds.), Advances in Electrophoresis. VCH Verlagsgesellschaft mbH, Weinheim, Germany.
8. **Olovnikov, A.M.** 1973. A theory of marginotomy. J. Theor. Biol. *41*:181-190.
9. **Poon, S.S., U.M. Martens, R.K. Ward, and P.M. Lansdorp.** 1999. Telomere length measurements using digital fluorescence microscopy. Cytometry *36*:267-278.
10. **Watson, J.D.** 1972. Origin of concatemeric T7 DNA. Nature *239*:197-201.

Received 26 January 2001; accepted 14 June 2001.

Address correspondence to:

Dr. Michael Ochs
Manager, Bioinformatics
Fox Chase Cancer Center
7701 Burholme Avenue
Philadelphia, PA 19111, USA
e-mail: m_ochs@fccc.edu