SEGMENTING TELOMERES AND CHROMOSOMES IN CELLS

*Steven S.S. Poon, Rabab W. Ward, *Peter M. Lansdorp*

Center for Integrated Computer Systems Research; Department of Electrical and Computer Engineering
2366 Main Mall, University of British Columbia, Vancouver, BC, Canada

*Terry Fox Laboratory for Hematology/Oncology, BC Cancer Research Centre;
Department of Medicine, University of British Columbia, Vancouver, BC, Canada

ABSTRACT

The very end of every chromosome is a region called the telomere. Telomeres are nucleo-protein complexes containing specific DNA repeat sequences whose lengths are strongly believed to give indications to aging and tumor progression. In order to study the role these repeat sequences play in the cell, we developed a fluorescence microscopy imaging system and associated image analysis methods to accurately measure these telomere lengths. To visualize the image of the tiny telomeres, we captured 2 spectrally different images of the same cell. One image contains only telomeres and the other contains only chromosomes. We next apply successful and novel methods to segment the telomere and chromosome images and then to link each chromosome with its telomeres. Our system is so far the only existing system available for this purpose and has already been in use in many research laboratories in Western Europe, North America, and Hong Kong.

1. INTRODUCTION

Telomeres play important roles in the function of the cell [3],[22]. They contain proteins and unique repetitive DNA sequences which protects the natural ends of chromosomes in eukaryotic (nucleated) cells from degradation and end-to-end fusions. Telomeres are also involved in gene regulation as the length of the telomere may determine if particular genes at the ends of chromosomes are expressed. In addition, telomeres play a crucial role in cell division. It has been recently shown that telomere lengths shorten with age until the telomeres reach a certain length which prevents the cell from further dividing [2],[8],[13],[20]. Cancer cells, however, are able to maintain the length of their telomeres after repetitive cell divisions [7],[9],[12].

The conventional technique for measuring the length of telomere repeat sequences is based on a gel electrophoresis method known as the Southern analysis method [1],[6]. There are many drawbacks to this technique which are resolved by our system. Instead of around 100,000 cells, less than 30 cells are required by our system to measure the telomere length distribution. This makes it possible to carry biological studies when only a limited number of cells are available for analysis. In addition, with our system, telomere length studies can now be carried out on each cell as well as individual chromosomes in every cell instead of a population of cells. Furthermore, the size of the telomeres obtained is not under or over-estimated and the obtained telomere signals are not biased in their lengths. Our new technique is based on the fluorescence in-situ hybridization technology which relies on probes (nucleic acid sequences) that can hybridize (bind) to specific sites in denatured chromosomes. By attaching a fluorescent marker onto the probe, the location of specific sites in the chromosome can be identified under a fluorescence microscope.

To visualize the tiny telomeres only, 2 sets of images of the cell at different spectral wavelengths are captured and analyzed by our system. The first one is an image of the telomeres only. The second is an image of the chromosomes but without the telomeres. The quantitative nature of the fluorescent marker allows us to estimate the telomere length based on the observed fluorescence intensity from the first image. The second image is then used to help determine which telomere belongs to which chromosome. In this paper, we will give an overview of our image acquisition system and describe the segmentation algorithms for the telomeres and chromosomes. We will then discuss the results obtained.

2. IMAGING SYSTEM

We built our imaging acquisition system by selecting, optimizing and integrating basic commercial components such that consistent results can be obtained over time. The major components of our acquisition system are the fluorescence microscope, the camera and the computer (Fig. 1). The fluorescence microscope (Zeiss Axioplan) with a 63x magnification objective lens (Plan Apochromat 63x/1.4, Zeiss) transforms and magnifies the telomeres and chromosomes for visualization. We equipped the microscope with a hybrid mercury/xenon lamp (200W, OptiQuip distributed by Zeiss) because this lamp fluctuates less than the traditional mercury lamp in time and is more intense than the xenon lamp at the wavelengths of interest (405nm and 546nm peak wavelengths used to excite the DAPI and CY3 probes, respectively). To reduce the image registration problem of matching images, we added a filter wheel (Pacific Scientific Inc.) which has openings for up to 8 different excitation filters. A single multi-spectrum dichroic mirror and emission filter assembly is then used to

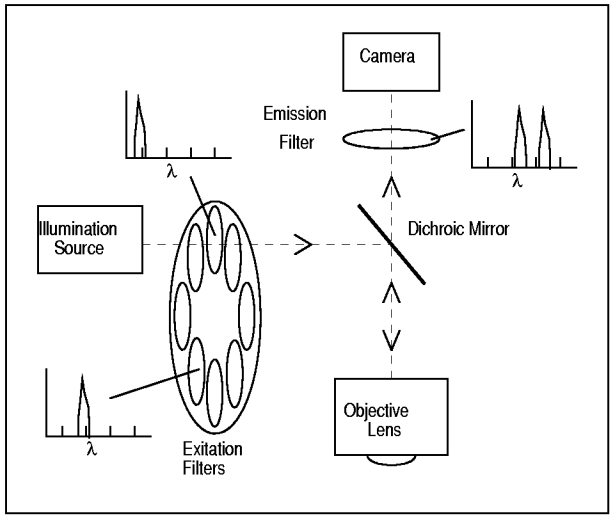

Figure 1. Block diagram of the excitation and emission filter system. An 8-position filter wheel is used to select the excitation wavelength in the illumination path. A double band pass dichroic and emission filter is used in the imaging path.

image all probes in the experiment. To capture the image, we chose the MicroImager MI1400-12 digital camera (Xillix Technologies Corp.) because it meets the following requirements: i) high spatial resolution and large field of view, ii) sufficient photometric resolution, high sensitivity, and large dynamic range, and iii) multi-spectral image acquisition capability and relatively fast readout rates [10],[17]. The captured images are then stored and analyzed by the computer.

3. TELOMERE SEGMENTATION AND LENGTH MEASUREMENTS

The length of the telomere is correlated to the integrated fluorescence intensity (IFI) value which is a measure of the total amount of fluorescence emitted from the object. The major problem in accurately quantifying the IFI of telomeres lies in defining the region in which to sum the fluorescence intensities. Thus, the correct segmentation, i.e. determining the exact boundaries of each telomere, is important. Most telomeres are relatively easy to detect since they appear as bright tiny spots. Approximate locations of these spots can be found by thresholding or edge detection methods. However, the inherent noise in the system (optics and illumination aberrations, camera noise, sample preparation noise, etc.) and interference of nearby telomeres make it difficult to define the "true" background/normalization level and the segmented regions for the IFI calculation (Fig. 2a).

Thus, we developed a segmentation algorithm called the Average Difference Filter [16]. This filter is similar in operation to the Laplacian filter [19]. This filter detects bright spots and deletes edge pixels. By removing the edge pixels, the filter separates touching telomeres. For each pixel, the average intensity value of its surrounding pixels is subtracted from its intensity, $I(x,y)$ to generate an edge image, $E(x,y)$, that is:

$$E(x,y) = I(x,y) - \frac{1}{9}\sum_{i=-1}^{1}\sum_{j=-1}^{1} I(x-i, y-j)$$

Using the appropriate threshold measured from the background noise level, our algorithm eliminates noise and detects intensity peaks. In addition, telomeres that are close to each other are separated using this technique. To recover the lost edge pixels, removed by this algorithm, we first label the segmented telomere and then dilate the region of each telomere to avoid accidental combination of multiple telomeres into one (Fig. 2b).

4. CHROMOSOME SEGMENTATION

To determine which chromosome each segmented telomere belongs to, we need to segment and identify each chromosome. The variability in the chromosome texture (intensity) within individual chromosomes and amongst different chromosomes makes it difficult to find the exact borders of each chromosome (Fig. 2c). In addition, the high noise levels associated with low light level fluorescence images pose another difficulty for segmentation. Another segmentation difficulty lies in defining the boundaries of touching and overlapping chromosomes. All available segmentation programs at present cannot correctly segment all chromosomes. For example, commercial chromosome analysis systems from Applied Imaging Inc., Biological Detection Incorporated, Vysis Inc., and Oncor Instrument Systems tend to first use a simple semi-automated segmentation algorithm to generate an initial estimate of the chromosome borders and then allow the user to interactively verify and correct the results. By using our method, a vast majority of chromosomes are correctly segmented. Hence, less user interaction is required in the manual verification process resulting in a less tedious and a more economical overall interactive analysis.

Our chromosome segmentation algorithm consists of a combination of different segmentation methods since no single technique produced good results. Each step in the sequence improves on the results obtained by the previous one. Thresholding is first used to define the rough regions occupied by chromosome region, $I(x,y)$ *(Fig. 2c).* Texture information from the thresholded region is then used to give a more defined region. In this step, we first detect the local high intensity pixels using the following Average Difference Filter (a 5x5 filter region to smooth out more noise and texture present in chromosome images):

$$J(x,y) = I(x,y) - \frac{1}{25}\sum_{i=-2}^{2}\sum_{j=-2}^{2} I(x-i, y-j)$$

We then impose a non-negative constraint on the difference image $J(x,y)$ to get image $K(x,y)$. As a result of this operation, the background regions, which are close to the chromosomes, and most of the chromosome edges are removed from the segmented region. Most of the points in-between touching chromosomes are also eliminated, since they have negative difference values. We then use our Rank Difference filter ($R_{7,1}[S(K(x,y))]$) as a

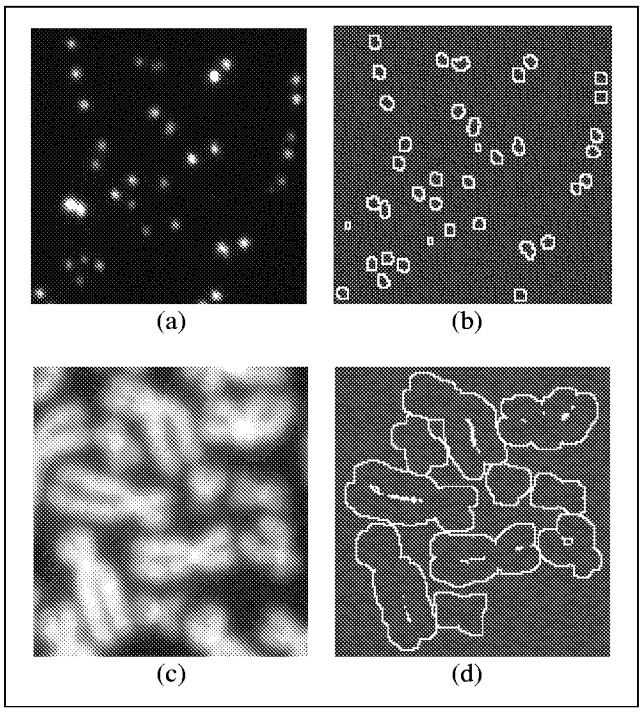

Figure 2. Results of segmentation. Images (a) and (c) are the captured fluorescence images of the cell showing only the telomeres and chromosomes, respectively. Images (b) and (d) are the corresponding segmentation results.

morphological operator to merge detected pixels into different chromosome regions and at the same time further separate touching chromosomes. The Rank Difference filter is a general form of the maximum local contrast filter [19]. Instead of limiting the difference to be between the maximum and minimum rank filters, the difference can be performed between any two ranks filters number (i.e. an upper rank $R_u[S(i(x,y))]$ and a lower one $R_l[S(i(x,y))]$) as follows:

$$R_{u,l}[S(i(x,y))] = R_u[S(i(x,y))] - R_l[S(i(x,y))]$$

where minimum rank $\leq l < u \leq$ maximum rank, defined over a region $S(i(x,y))$ of the image $i(x,y)$. The resulting Rank Difference image is then binarized by setting all negative values to 0 and all others to 255 to get image $L(x,y)$.

The next step in our segmentation process first finds the edges of the chromosome regions using our Rank Difference filter ($R_{9,1}[S(L(x,y))]$) again, but this time it functions as an edge detector to get image $M(x,y)$. We chose to use our Rank Difference filter instead of the difference of Gaussian [14] or Canny [5] filters because of the following reasons. First, the algorithm is already available within the program. Second, the algorithm uses only integer operations and is less complex and hence it is faster to compute. Third and most important, our filter gives thick edges at the appropriate locations such that some of the remaining touching chromosomes that are not segmented in previous approximations is now separated. The edges generated are then used to refine the borders for each detected chromosome region and separate the touching chromosomes. In this operation, a new chromosome region, $N(x,y)$ is generated based on the logical AND (•) of the previous chromosome segmented region, $M(x,y)$, with the logical NOT (—)of the newly calculated edge image, $O(x,y)$ as follows:

$$N(x, y) = M(x, y) \bullet \overline{O(x, y)}$$

The regions found so far are 2 pixels smaller than the actual regions of the chromosomes. However, they are mostly distinct and isolated from one another. Each object is next labeled such that each isolated region is given a distinct number. The size of the region of each labeled object is then increased such that it is representative of the size of the chromosomes. This increasing process is accomplished by dilating each labeled region twice using a 3x3 dilation filter. As different label numbers are used in the dilation process, regions that touch one another after the dilation are kept distinct with different label numbers.

Finally, segmented objects whose fluorescence intensities are too weak and whose sizes are too small to be chromosomes are rejected. To further refine both the telomere and chromosome segmented regions, both the telomere and chromosome images are used to determine if each chromosome region contains 4 corresponding telomere spots (Fig. 2d). Other chromosome segmentation algorithms do not have our added advantage of having corresponding telomere images which can help in defining the ends of the chromosomes.

5. DISCUSSION AND CONCLUSIONS

The accuracy of the telomere segmentation is important for determining the telomere length measurement which is the ultimate goal of our work. Thus, it is more important to validate the telomere length measurements, which will indirectly validate the accuracy of our segmentation results. As no direct method for validating the accuracy of our telomere length measurements is available, we resorted to indirect methods to validate our fluorescence measurements. For this purpose, we used test objects of known fluorescence intensities that resemble telomeres. These objects included i) simulated objects of different shapes and sizes, ii) fluorescence beads of known size and relative fluorescence intensities, and iii) plasmids with known telomere insert lengths (which are typically an order of magnitude less in length than the telomeres in the cells). Our algorithm estimated the integrated fluorescence intensity (IFI) of simulated objects of varying shapes and sizes to within 3%. The estimated mean IFI values correlated well (correlation coefficient of 0.99) with the size of the fluorescence beads and with the length of telomere insert in plasmids [15]. The standard deviation in the estimation ranged from 2% for the 1μm beads to 13% for the 0.2μm beads to 29% for the 0.1μm beads. The standard deviation was larger for the smaller beads because it is more difficult to fabricate them. The standard deviation for telomere inserts in plasmids was around 20% of the mean estimated IFI value. This variance is most likely due to the variable efficiency of the hybridization procedure (binding of the probe).

To conclude, our results show that using our segmentation method, the telomere length could be measured to within 3%

accuracy in a population of cells and around 20% for individual telomeres. Although the variation appears to be large for individual telomeres, we observed that by averaging the results of 10 or more cells, a good indication of the differences in the lengths between any 2 specific telomeres can be obtained (i.e. differences between groups with a significance level less than 0.05 using the Wilcoxon rank sum test).

In addition, we have developed algorithms to segment chromosomes including those that are just touching. The segmentation results and the calculated telomere IFI values are then presented to the user for verification and editing. The automation of the telomere and chromosome extraction and IFI calculation process has simplified the user verification and editing process. On average, over 90% of the chromosomes are segmented properly. The success rate in segmentation is dependent on the metaphase sample, which typically contains a few overlapping chromosomes.

Our telomere analysis system has already enabled numerous studies concerning the role of telomeres in aging or genetic disorders [4],[11],[12],[15],[21],[23]. A number of improvements could still be made to the system. A cooled integrating CCD camera would be useful to obtain better quality images. With such a camera, the accuracy of the IFI algorithm would be improved since it is no longer necessary to estimate the value of the faulty pixel by taking the average of its surroundings. Another improvement would be to implement a more objective method of focussing [18]. This would result in more consistency in image acquisition and provide a more accurate estimate of the telomere length.

6. REFERENCES

[1] Allshire RC, Gosden JR, Cross SH, Cranston G, Rout D, Sugawara N, Szostak JW, Fantes PA, and Hastie ND: Telomeric repeat from T. thermophilia cross-hybridizes with human telomeres. Nature 332:656-659, 1998.

[2] Allsopp R, Vaziri H, Patterson C, Goldstein S, Younglai E, Futcher B, Greider CW, and Harley CB: Telomere length predicts replicative capacity of human fibroblasts. Proc. Natl. Acad. Sci. USA 89:10114-10118, 1992.

[3] Benbow R.M. [1992] Chromosome structures. Sci. Progress Oxford. 76:425-450.

[4] Blasco MA, Lee HW, Hande MP, Samper E, Lansdorp PM, DePinho RA, Greider CW: Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. Cell 91:25-34, 1997.

[5] Canny JF: A computational approach to edge detection. IEEE Trans. of Pattern Analysis and Machine Intelligence 8(6):679-698, 1986.

[6] de Lange T, Shiue L, Myers RM, Cos DR Naylor SL Killery AM, and Varmus HE: Structure and variability of human chromosome ends. Mol. Cell. Biol. 10:518-527, 1990.

[7] de Lange T: Telomere Dynamics and Genome Instability in Human Cancer. in Telomeres. eds. Blackburn E. H., and Greider C. W. Cold Spring Harbour Laboratory Press, New York, 265-293, 1995.

[8] Harley CB, Futcher AB, and Greider CW: Telomeres shorten during aging of human fibroblasts. Nature 345:458-460, 1990.

[9] Hastie ND, Dempster M, Dunlop MG, Thompson AM, Green DK, and Allshire RC: Telomere reduction in human colorectal carcinoma and with ageing. Nature 346:866-868, 1990.

[10] Jaggi B, Pontifex B, Swanson J, Poon S.S.S: Performance Evaluation of a 12-Bit, 8Mpel/s Digital Camera. SPIE Proceedings, Cameras Scanners, and Image Acquisition Systems, 1901:99-108, 1993

[11] Lansdorp PM, Poon S, Chavez E, Dragowska V, Zijlmans M, Bryan T, Reddel R, Egholm M, Bacchetti S, and Martens U: Telomeres in the hematopoietic system. CIBA Foundation Symposium No:211. Telomeres and Telomerase. John-Wiley & Sons Ltd. Chichester U.K., 209-219, 1997.

[12] Lansdorp PM, Verwoerd NP, van de Rijke FM, Dragowska V, Little MT, Dirks RW, Raap AK, and Tanke HJ: Heterogeneity in telomere length of human chromosomes. Human Molecular Genetics 5(5):685-691, 1996.

[13] Levy MZ, Allsopp RC, Futcher AB, Greider CW, and Harle CB: Telomere End-replication Problem and Cell Aging. 225:954-960, 1992.

[14] Marr D: Vision, WH Freeman and Company, New York, 1982.

[15] Martens UM, Zijlmans JMJM, Poon SSS, Dragowska V, Yui J, Chavez EA, Ward RK, and Lansdorp PM: Short telomeres on the p-arm of human chromosome 17. Nature Genetics 18:76-80, 1998.

[16] Poon SSS: Telomere Length Measurements Using Fluorescence Microscopy. Thesis disertation, Univ. of British Columbia, 1998.

[17] Poon SSS, and Hunter DB: Electronic cameras to meet the needs of microscopy specialists. Advance Imaging. 9(7):64-67, 84, 1994.

[18] Poon SSS, Ward RK, and Palcic B: Feature extraction from three-dimensional images in quantitative microscopy. Microns and Microscopica Acta, 23(4):481-489, 1992.

[19] Russ JC: Computer Assisted Microscopy: The Measurement and Analysis of Images. Plenum Press, New York, 1990.

[20] Vaziri H, Dragowska W, Allsopp RC, Thomas TE, Harley CB, and Lansdorp PM: Evidence for a mitotic clock in human hematopoietic stem cells: loss of telomeric DNA with age. Proc. Natl. Acad. Sci. USA, Cell Biology, 91:9857-9860, 1994.

[21] Wan TSK, Martens UM, Poon SSS, Tsao SW, Chan LC, Lansdorp PM: Absence or low number of telomere repeats at junctions of dicentric chromosomes. Genes Chromosomes Cancer. 1998 (in press).

[22] Wilson E, Larosche T, and Gasser M: Telomeres and the functional architecture of the nucleus. Trends Cell Biol. 3:128-134, 1993.

[23] Zijlmans JMJM, Martens UM, Poon S.S.S, Raap AK, Tanke HJ, Ward RK, and Lansdorp PM: Telomeres in the mouse have large inter-chromosomal variations in the number of T_2AG_3 repeats. Proc. of National Academy of Science USA, Genetics 94:7423-7428, 1997.