Three-dimensional analysis tool for segmenting and measuring the structure of telomeres in mammalian nuclei

Bart J. Vermolen[a], Ian T. Young[a], Alice Chuang[b], Landon Wark[b], Tony Chuang[b], Sabine Mai[b] and Yuval Garini[a]

[a]Quantitative Imaging Group, Department of Imaging Science and Technology, Faculty of Applied Sciences, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands;
[b]Manitoba Institute of Cell Biology, CancerCare Manitoba, University of Manitoba, 675 McDermot Avenue, Winnipeg, MB, R3E OV9, Canada

ABSTRACT

Quantitative analysis in combination with fluorescence microscopy calls for innovative digital image measurement tools. We have developed a three-dimensional tool for segmenting and analyzing FISH stained telomeres in interphase nuclei. After deconvolution of the images, we segment the individual telomeres and measure a distribution parameter we call ρ_T. This parameter describes if the telomeres are distributed in a sphere-like volume ($\rho_T \approx 1$) or in a disk-like volume ($\rho_T \gg 1$). Because of the statistical nature of this parameter, we have to correct for the fact that we do not have an infinite number of telomeres to calculate this parameter. In this study we show a way to do this correction. After sorting mouse lymphocytes and calculating ρ_T and using the correction introduced in this paper we show a significant difference between nuclei in G2 and nuclei in either G0/G1 or S phase. The mean values of ρ_T for G0/G1, S and G2 are 1.03, 1.02 and 13 respectively.

Keywords: Telomeres, 3D imaging, image processing, fluorescence microscopy, FISH

1. INTRODUCTION

With the advent of sequence-specific DNA probes, the use of fluorescence microscopy in cancer and genetics research has steadily grown. Continuous improvements in fluorescence microscopy methods (hardware and software), specific labeling methods (wetware), and better understanding of the genome functions and structure now enable us to detect almost any DNA sequence, gene or chromosome region with high sensitivity and to address the central question: 'How do all these different aspects contribute to our understanding of biological processes?'

As data acquisition (the recording of images) in fluorescence methods is usually based on digital imaging, quantitative analysis can be used and has, in fact, become a crucial part of the methodologies. These methods, therefore, require suitable quantitative image analysis procedures and algorithms. As one of the last links in the chain, the algorithms being used must take into account the whole procedure that is being used including the optical properties of the microscope and system, nature of the probes and instrument parameters (e.g. numerical aperture of the objective, wavelength of light) for the acquisition.

The organization of the interphase nucleus has been studied since the late 19th century.[1] It is now well accepted that the position of chromosomes in the nucleus plays an important role in gene regulation.[2] Recently, interest has also focused on telomeres whose importance to genomic stability was recognized as early as in the 1930s.[3] We have developed a method of studying the spatial organization of the genome in the three-dimensional (3D) interphase nucleus. We analyze digital images of the 3D organization of the telomeres and how their positions change during the cell-cycle. This method enables us to determine that telomere organization is cell-cycle dependent with assembly of telomeres into a telomeric disk (TD) in G2 phase. Further, this disk formation is disrupted in tumor cells.[4]

Further author information: (Send correspondence to B.J. Vermolen)
B.J. Vermolen: E-mail: b.j.vermolen@tnw.tudelft.nl, Telephone: +31 (0)15 27 83221

Imaging, Manipulation, and Analysis of Biomolecules and Cells: Fundamentals and Applications III, D. V. Nicolau, J. Enderlein, R. C. Leif, D. L. Farkas, R. Raghavachari, eds., Proc. of SPIE Vol. 5699 (SPIE, Bellingham, WA, 2005) · 1605-7422/05/$15 · doi: 10.1117/12.590232

Table 1. Characteristics of the microscope system

$FWHM_{lateral}$	200 nm
$FWHM_{axial}$	400 nm
Δx	106 nm
Δy	106 nm
Δz	200 nm
M	63x
NA	1.4
Filters	DAPI, Cy3
Typical image size	200 x 200 x 100

We characterize the TD with a parameter called, ρ_T.[5] Calculation of this parameter is problematic with a finite number of measurement points. This is the case with telomeres; we only have a finite number. In this study we concentrate on a correction for this parameter using numerical analysis.

2. MATERIALS AND METHODS

2.1. Cell Preparation

We have studied a mouse B lymphocyte cell nuclei population. The immortalized cells were sorted according to their DNA content for the determination of G0/G1, S or G2 phase. Cell cycle fractions were quantified through fluorescent-activated cell sorting analysis.[4] Flow analysis were performed on a EPICS Altra^{TM}cytometer operating under Multicycle^{TM}software (Beckman-Coulter, France). Approximately 10-15 nuclei from each phase were analyzed for this study, representing a total of 35 cell nuclei.

For measurements of the telomeric disk the cells were first fixed and then telomere fluorescence *in situ* hybridization (FISH) was performed as described in Figueroa, et al.[6] using a Cy3-labeled PNA probe (DAKO, Glostrup, Denmark). DAPI was used as a DNA-specific counterstain. Telomere hybridizations were specific and we verified the correct number of telomeric signals observed at the ends of chromosomes prepared from primary cells using 2D FISH metaphase spreads. The lymphocytes were fixed in such a way that the 3D structure of the nuclei was conserved.[4]

2.2. 3D Image Acquisition

For analysis of the telomere distribution, images were acquired with a Zeiss Axioplan 2 with a cooled AxioCam HR CCD in combination with a PlanApo 63x/1.4 oil immersion objective (Zeiss). This gives a pixel (sampling) distance in the lateral plane of $\Delta x = \Delta y = 106$ nm. The axial sampling distance between the planes was $\Delta z = 200$ nm. The point spread function (psf) of the objective, which determines the optical resolution, gave a Full Width at Half Maximum of approximately $FWHM = 200$ nm in the lateral direction and 400 nm in the axial direction. Typical image size was 200 x 200 x 100 pixels. Table 1 gives a summary of these values for this imaging system. Figure 1 illustrates the system resolution. An image of a pair of telomeres relatively far apart and an image of a pair close together is shown. It is clear that the telomeres at a distance of 1200 nm can be easily distinguished and telomeres at a distance of 400 nm are just barely separable.

2.3. 3D Image Processing

The 3D digital images were processed to improve the resolution using a constrained iterative maximum likelihood deconvolution[7] which is available in the AxioVision 3.1 (Zeiss) software. This convolution method was chosen for this work because it has been shown to provide the best results.[8] An example of the result of applying this procedure to telomere images is shown in figure 2. After deconvolution the image is interpolated in the axial (z)

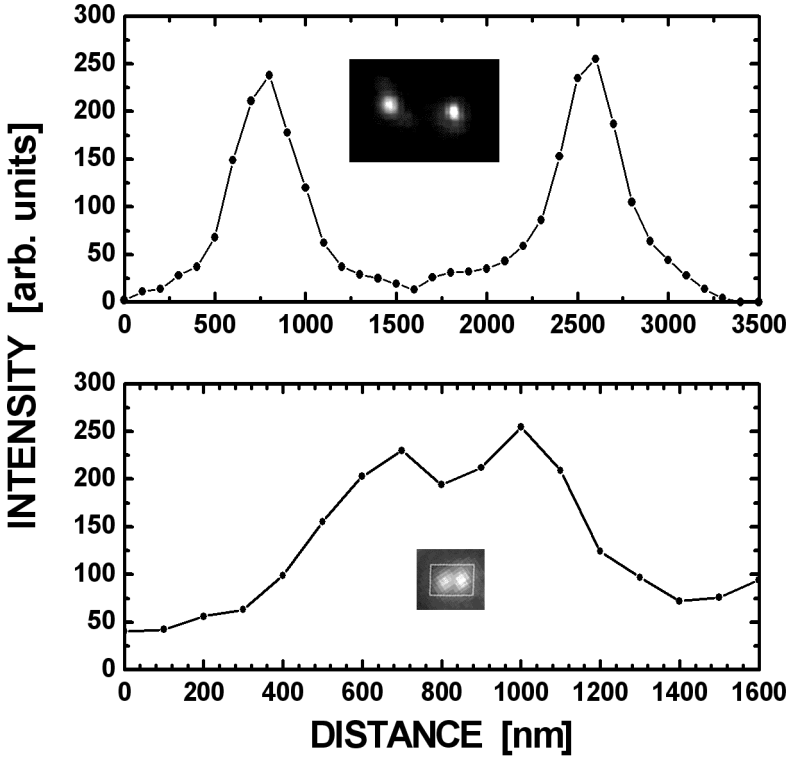

Figure 1. Demonstration of the spatial resolution of our measurements, Two pairs of telomeres are shown: 1200 nm apart (top), which can be easily separated, and 400 nm apart (bottom). The inserts show the original image and the graphs a section through the telomeres.

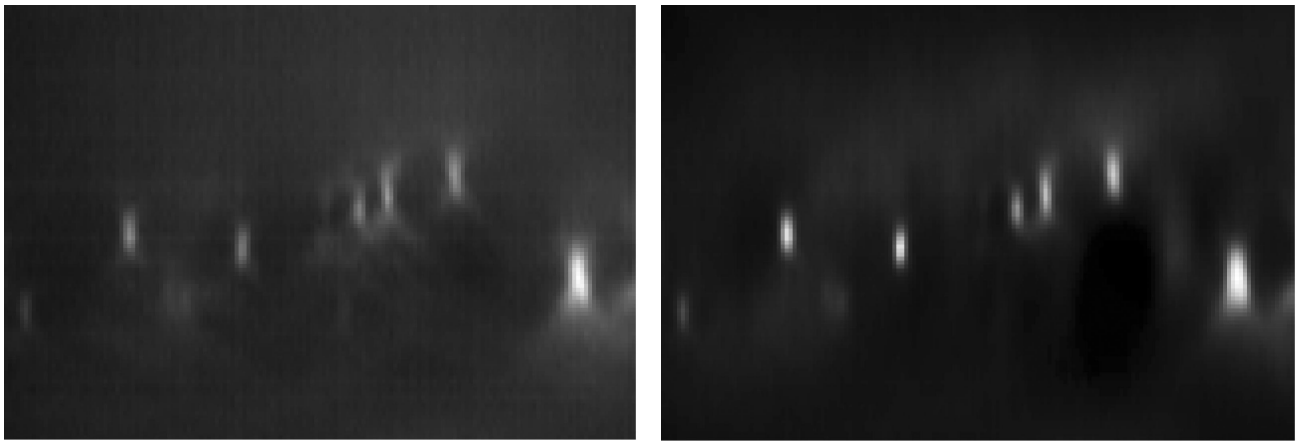

Figure 2. Demonstration of the effect of deconvolution. The left image is before deconvolution and the right is after deconvolution. We clearly see that the left image has more blur and has less contrast than the right image. Both images are shown with a linear intensity contrast stretch.

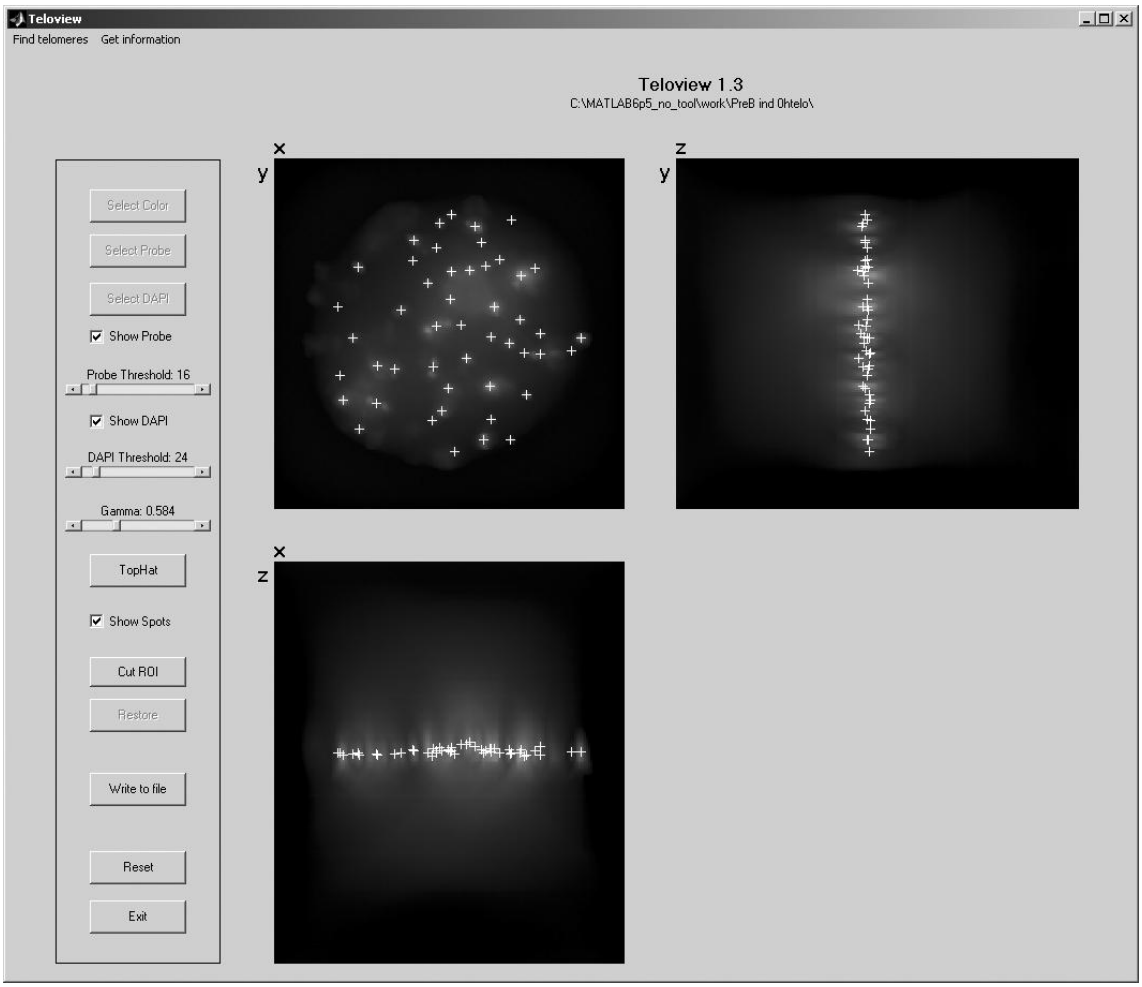

Figure 3. Screenshot of the user interface of TeloView. The screen shows the three displays with maxmimum instensity projections along the three main optical axes. It also shows crosses at the locations where the software identified a telomere. After automatic identification the user can interactivly add or remove crosses.

direction from Δz to $\Delta z'$ so that the sampling distance in all three directions is the same: $\Delta x = \Delta y = \Delta z' = 106$ nm.

2.4. 3D Image Segmentation and Analysis

Image segmentation and analysis of deconvolved 3D images of cells with labeled telomeres have been performed with a sequence of procedures that we have bundled together and named TeloView. The procedures themselves are from our image software library DIPImage which is available as public domain software at <http://www.qi.tnw.tudelft.nl/DIPlib/>. The version of DIPImage used in this development operates under MatLab (The MathWorks, Natick, MA, USA).

TeloView loads the 3D image and displays a maximum projection along the three main optical axes. While thresholds and other parameters can be adjusted for display purposes, the analysis is performed on the original 3D data. After segmentation, the 2D display indicates the location of the automatically found spots for verification. At this point the user can decide to remove falsely labeled spots or add spots that were not found. A screenshot of the user interface can be found in figure 3.

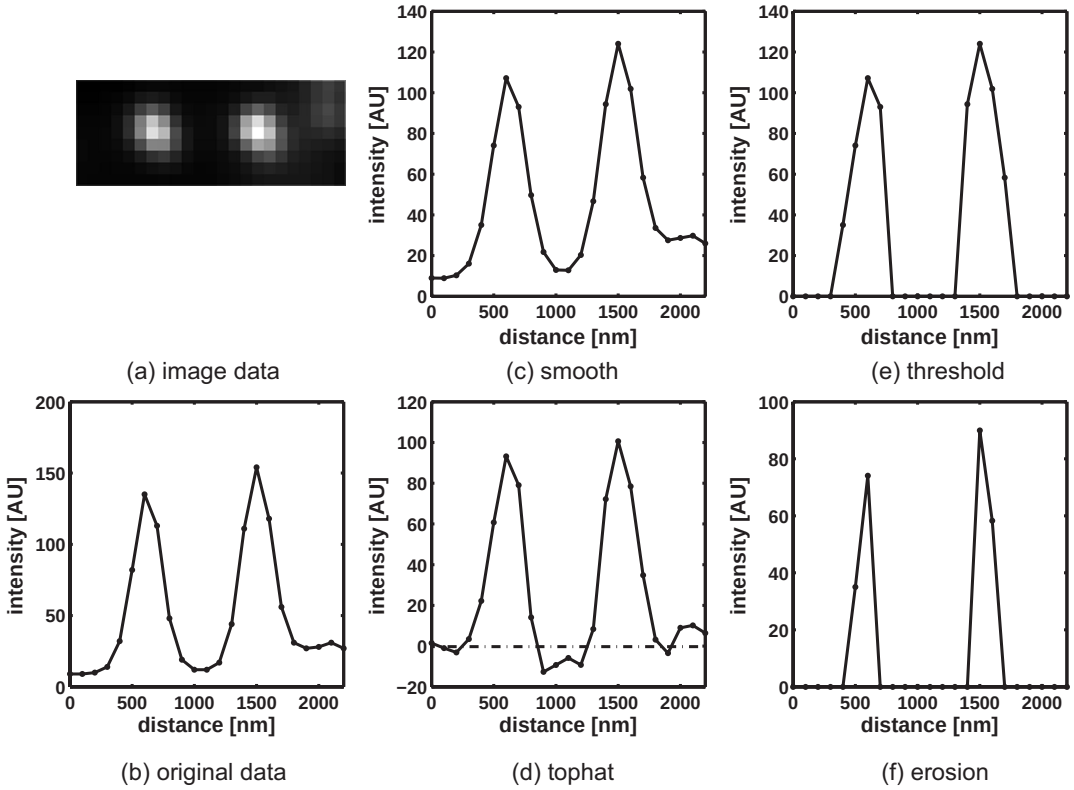

Figure 4. Figure showing the working of the algorithm. First we see the raw 'image data'. A line through the center of this image gives a line section seen in 'original data'. After we 'smooth', we perform a 'tophat' transform. Note that shading is now removed. We 'threshold' and end up with two spots. One last 'erosion' is performed to make sure there are no remaining noise spikes.

2.4.1. Segmentation

Before starting the segmentation we pre-process the data by smoothing with a 3D Gaussian kernel. Figure 4 shows how the data is transformed during the different steps of segmentation. For segmentation of the individual telomeres we have chosen an algorithm based on a morphological TopHat transformation.[9,10] The TopHat transform on an image A with structuring element B is defined as follows[11]; to find objects with high intensity ('light objects'):

$$TopHat(A, B) = A - \max_B(\min_B(A)). \tag{1}$$

To find objects with low intensity ('dark objects'):

$$TopHat(A, B) = \min_B(\max_B(A)) - A. \tag{2}$$

the 'structuring' element, B, can be a quite general three-dimensional, grey-value object but in our case we have chosen for the simple case that B is spherical. The size of B should be bigger than the objects that are being sought but smaller than any shading in the background. For a grey-scale image of telomeres, the telomeres would be our objects and any non-specific binding of Cy3 uniform spread over the nucleus gives us shading. Thus, for our case, this translates to a spherical B with radius 742 nm (7 pixels). After the TopHat transform the resulting image is thresholded with a user chosen value to produce a binary mask. To eliminate noise spikes that may still remain, we conclude with an erosion. This algorithm gives satisfying results for small telomeres. Using

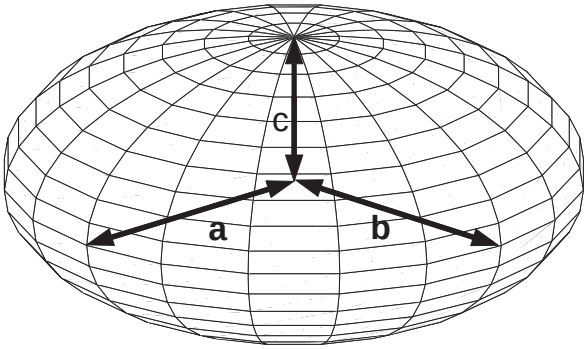

Figure 5. The envelope around the telomeres can be characterized by an oblate spheroid. Here, two main axes have equal length, $a = b$. The third main axis is shorter, c. The ratio $\rho_T = \frac{a}{c}$ gives a measure for the flatness of the spheroid.

the binary image mask from the segmentation, the center of gravity of each dot in the grey-scale image is found. This gives coordinates (x_i, y_i, z_i) for each individual dot, where i is the index number of the dot.

2.4.2. Analysis

We have shown[4,5] that the envelope shape of the telomeres can be characterized as an oblate spheroid. A spheroid is a geometric figure, like an ellipsoid, where the two main axes, a and b, are of equal length and the third main axis, c, has a different length as displayed in figure 5.

To measure the flatness of the spheroid ($c < a$), we define a parameter, ρ_T:

$$\rho_T = \frac{a}{c}. \tag{3}$$

If $\rho_T \approx 1$, the telomeres are distributed more spherical in a nucleus. If, however, $\rho_T > 1$, then the telomeric territory is more disk-like. We determine ρ_T by performing principal component analysis[12] on the spatial spot coordinates $(x_i, y_i, z_i | i = 1, 2, \ldots, N)$. We do this by calculating the singular values (eigenvalues) of the covariance matrix of the data points. The three singular values, $\lambda_1 \geq \lambda_2 \geq \lambda_3$, are real, positive and ordered. They are the variances of the distances from the spots to the new principal axes introduced by the principal component analysis. The standard deviation for each axes is then given by:

$$\sigma_i = \sqrt{\lambda_i}. \tag{4}$$

Now we define ρ_T as

$$\rho_T = \frac{\sqrt{\sigma_1 \sigma_2}}{\sigma_3}. \tag{5}$$

That is, the geometric mean $\sqrt{\sigma_1 \sigma_2}$ is taken as the value for a and σ_3 is the value for c. Given that we work with ordered λ's we have:

$$\rho_T \geq 1. \tag{6}$$

3. CORRECTION OF ρ_T BY MODEL FITTING

For an infinite number of spots in a spheroid, the parameter, ρ_T, would give us the ratio of the length of the principal axes: $\frac{a}{c}$. Because we have a finite number of spots, N, this does not hold anymore. We have done Monte-Carlo simulations to see what the effect is on the observed $\rho_{T,o}$ with respect to the real $\rho_{T,r}$ and N.

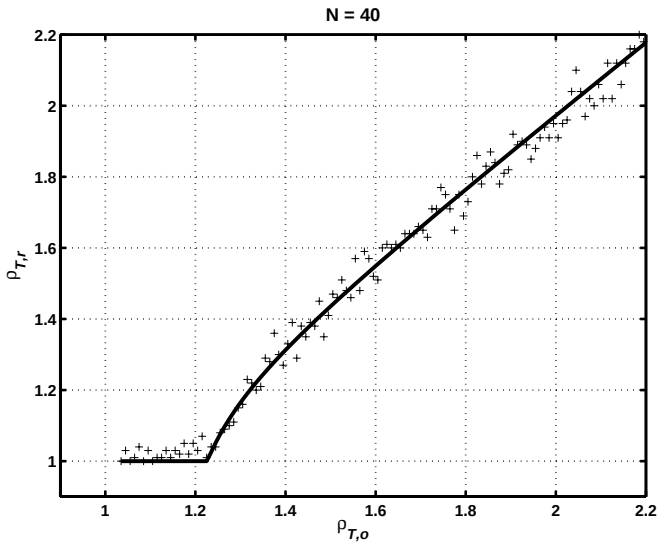

Figure 6. Figure showing maximum likelihood estimators $\hat{\rho}_{T,r}$ plotted against $\rho_{T,o}$ (plusses) and the fitted curve (line).

For this we used MatLab to generate a set of uniformly distributed pseudo-random numbers (x, y, z) in the interval $(-1.0, 1.0)$. Then we calculate the squared distance to the center of the sphere

$$r^2 = x^2 + y^2 + (\rho_{T,r} \cdot z)^2. \tag{7}$$

Here $\rho_{T,r}$ scales the spheroid to a sphere. If

$$r^2 \leq 1, \tag{8}$$

we accept the point, because it is inside the spheroid. This process is repeated until N points are accepted. From this set of coordinates, (x_i, y_i, z_i), $i = 1 \ldots N$, we can compute $\rho_{T,o}$. For every $\rho_{T,r}$ we compute 10000 values of $\rho_{T,o}$. This gives us the probability distribution, $p(\rho_{T,o}|\rho_{T,r})$, for a given N. Using Bayes' rule we can now obtain the distribution, $p(\rho_{T,r}|\rho_{T,o})$, for a given N:

$$p(\rho_{T,r}|\rho_{T,o}) = \frac{p(\rho_{T,r})p(\rho_{T,o}|\rho_{T,r})}{p(\rho_{T,o})}. \tag{9}$$

Because no information is known about the a-priori probability we assume, for practical reasons, $p(\rho_{T,r})$ to be uniform. The $\rho_{T,r}$ which gives us the maximum of $p(\rho_{T,r}|\rho_{T,o})$ is the maximum likelihood estimator, $\hat{\rho}_{T,r}$. In figure 6 we can see $\hat{\rho}_{T,r}$ plotted against $\rho_{T,o}$ for $N = 40$.

The next step is to model a curve for a given N. This is done by cutting the curve in two:

$$\begin{array}{rcll} \widehat{\hat{\rho}_{T,r}} & = & 1 & for \quad \rho_{T,o} < d, \\ \widehat{\hat{\rho}_{T,r}} & = & \rho_{T,o} - \frac{a^c(d-1)}{(\rho_{T,o}-b)^c} & for \quad \rho_{T,o} \geq d. \end{array} \tag{10}$$

Here a, b, c and d are the model parameters, where b is dependent on d and a:

$$b = d - a, \tag{11}$$

to make the function continues. Now we can model the independent parameters a, c and d as a function of N:

$$\begin{array}{rcl} a & = & \gamma_1, \\ c & = & \gamma_2 N + \gamma_3, \\ d & = & \frac{\gamma_4}{N^{\gamma_5}} + 1. \end{array} \tag{12}$$

Table 2. Results of phase sorted cells.

Phase	**Mean** ρ_T	**Stand. dev.** ρ_T
G0/G1	1.03	0.10
S	1.02	0.06
G2	13	3

This gives us the fitting parameters, γ_i $(i = 1 \ldots 5)$, for our model. Next, we fit the model to our data points using a least squares approach, i.e. we minimize the error:

$$\varepsilon = \sum_{\rho_{T,o},N} (\widehat{\hat{\rho}_{T,r}}(\gamma_i) - \hat{\rho}_{T,r})^2, \tag{13}$$

which gives us γ_i. In figure 6 we see the result of the model fitting for $N = 46$.

4. RESULTS

The results of the analysis of the cell-sorted mouse lymphocytes can be found in Table 2. The values of ρ_T measured here are a little bit smaller than in Vermolen et al.[5] This is because of the correction we did in section 3. We still see a clear difference in the measurements between G0/G1 and G2 phase and between S and G2 phase. In other words the telomeres in the G0/G1 and S phases are distributed in a more spherical volume and the telomeres in the G2 phase in a disk-like volume. Typical distributions can be found in figure 7, where we clearly see these spherical and disk-like structures.

5. DISCUSSION

We have developed a three-dimensional digital image analysis tool to segment the telomeres in the nucleus. We also provide an objective parameter, ρ_T, to describe the volume in which the telomeres are distributed; sphere-like or disk-like. In this paper we have mainly focused on a statistical correction for ρ_T. After numerical analysis we concluded that a correction is necessary and a way to do so is provided. A next step is to give a theoretical basis for this correction, but this is not given in this paper.

After doing the same measurements as in Vermolen, et al.,[5] now with the correction of ρ_T, we come to the same conclusions; ρ_T is significantly higher in nuclei in G2 than in G0/G1 or S.

We have observed[4] that telomeres tend to form aggregates, i.e. telomeres are clumped together. Therefore we are currently developing methods to measure the size of telomeres. This enables us to quantify the aggregates and correlate this to cancer progress. Because aggregates are usually bigger than individual telomeres the algorithm for segmentation we have described above is not suitable. We are working on an improved segmentation algorithm. For future studies, we propose to use live cell imaging to follow and measure the cell-cycle dependent movement of telomeres.

ACKNOWLEDGMENTS

This work was partially supported by the Delft Research Initiatives Information and Communication Technology, LifeTech, and Life Science and Technology, by the FOM research program Nano-electrophoresis (99TF13), and by the Bsik research programs MicroNed and Cyttron. Work in Canada was funded by the Canada Foundation for Innovation (CFI), the Canadian Instituted of Health Research (CIHR) and CancerCare Manitoba. T. C.-Y Chuang and A. Y.-C. Chuang were supported through a CIHR Strategic Training Program "Innovative Technologies in Multidisciplinary Health Research Training" fellowship.

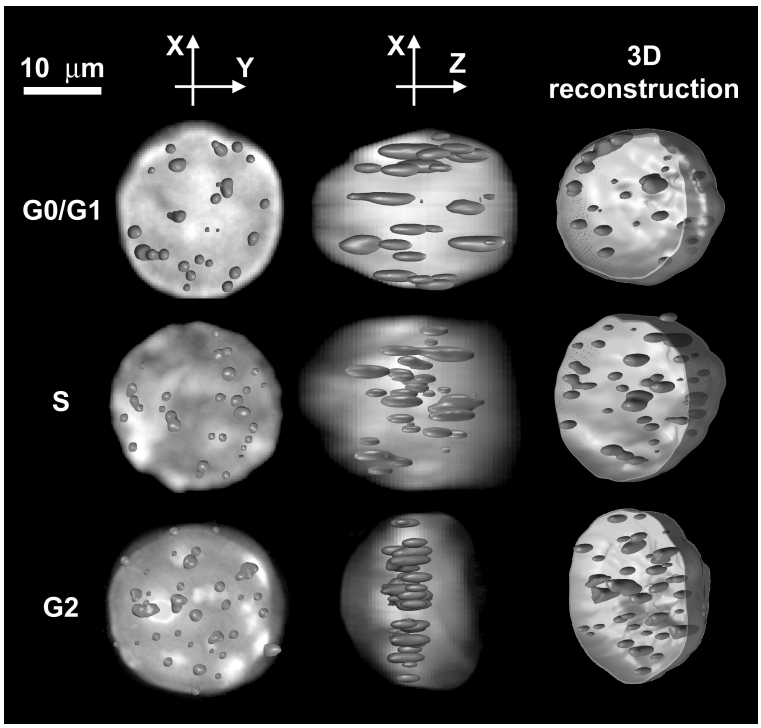

Figure 7. Figure showing the two typical distributions of the telomeres. In the upper row we see a nucleus in G0/G1 phase and in the middle row a nucleus in S phase. We can see the telomeres are distributed in a spherical volume. The lower row shows a nucleus in G2 phase, where the telomeres are distributed in a disk-like volume. The Cy3-stained telomeres are 'embedded' on the DAPI-stained nuclei.

REFERENCES

1. C. Rabl, "Über Zellteilung," in *Morphologisches Jahrbuch*, C. Gegenbauer, ed., **10**, pp. 214–330, 1885.
2. W. Marshall, "Order and Disorder in the Nucleus," *Curr Biol* **12**, pp. R185–R192, 2003.
3. B. McClintock, "The stability of broken ends of chromosomes in Zea Mays," *Genetics* **26**, pp. 234–282, 1941.
4. T. C. Y. Chuang, S. Moshir, Y. Garini, A. Y.-C. Chuang, I. T. Young, B. Vermolen, R. v. d. Doel, V. Mougey, M. Perrin, M. Braun, P. D. Kerr, T. Fest, P. Boukamp, and S. Mai, "The three-dimensional organization of telomeres in the nucleus of mammalian cells," *BMC Biology* **2**(12), 2004.
5. B. J. Vermolen, Y. Garini, S. Mai, V. Mougey, T. Fest, T. C. Y. Chuang, A. Y. C. Chuang, L. Wark, and I. T. Young, "Characterizing the Three-Dimensional Organization of Telomeres," *Cytometry* , 2005 (in print).
6. R. Figueroa, H. Lindenmaier, M. Hergenhahn, K. Nielsen, and P. Boukamp, "Telomeres erosion varies during in vitro aging of normal human fibroblasts from young and adult donors.," *Cancer Res* **60**, pp. 2770–2774, 2000.
7. L. H. Schaefer, D. Schuster, and H. Herz, "Generalized approach for accelerated maximum likelihood based image restoration applied to three-dimensional fluorescence microscopy," *J Microsc* **204**(2), pp. 99–107, 2001.
8. P. J. Verveer, M. J. Gemkow, and T. M. Jovin, "A comparison of image restoration approaches applied to three-dimensional confocal and wide-field fluorescence microscopy," *J Microsc* **193**(1), pp. 50–61, 1999.
9. F. Meyer, "Iterative Image Transformations for an Automatic Screening of Cervical Cancer.," *J Histochem Cytochem* **27**, pp. 128–135, 1979.

10. F. Meyer and S. Beucher, "Morphological Segmentation," *Journal of visual communication and image representation* **1**(1), pp. 21–46, 1990.
11. I. T. Young, J. J. Gerbrands, and L. J. van Vliet, "Fundamentals of Image Processing," in *The digital signal processing handbook*, V. K. Madisetti and D. B. Williams, eds., pp. 51.1–51.81, CRC Press in cooperation with IEEE Press, Boca Raton, Florida, 1998.
12. M. E. Wall, A. Rechtsteiner, and L. M. Rocha, "Singular value decomposition and principal component analysis," in *A Practical Approach to Microarray Data Analysis*, D. Berrar, W. Dubitzky, and M. Granzow, eds., pp. 91–109, Kluwer, Norwell, 2003.